Louvain School of Management

How do macroeconomic indicators influence investor behaviour during periods of market volatility ?

Author : Pierre Pattijn

Supervisor: Bertrand Candelon

Academic year 2023-2025

Declaration

During the preparation of this master's thesis, the author(s) utilized chat GPT for the following purpose:

Text reformulation: AI was used to rewrite some parts of the work to make it more comprehensible and easily understandable. The content of the text is coming from my own reflection.

R code creation help: AI was used to help me create the R code and adapt the methodology into code.

Help in adapting mathematical methodology: AI was used to adapt the mathematical methods I had found to my data.

After using Chat GPT, the author diligently reviewed and edited the content produced by the tool. We take full responsibility for the final content presented in this thesis.

By signing this declaration, we affirm that the content of this master's thesis reflects our original work, augmented by the responsible use of AI.

The 6th August 2024

Acknowledgements

First of all, I'd like to thank my supervisor, Professor Bertrand Candelon, for his irreplaceable advice and unfailing availability. The feedback I received from Mr. Candelon greatly helped me to write this thesis.

Secondly, many thanks to the UCLouvain for the tools made available and the quality of its teaching.

I would also like to thank my parents and my brother who supported and advised me throughout the writing of this thesis. Without their unfailing help I would not have been able to accomplish this research work and more generally my studies.

Table of contents

1. Introduction

Financial markets are complex, dynamic environments that are constantly evolving. Numerous events, monetary policies and behaviors can influence them, and the consequences of these influences are multiple and uncountable.

However, there are periods when investors in the financial markets face greater unpredictability in asset prices. These periods are known as market volatilities. Understanding investor behavior during these particular periods can provide valuable information to economists, researchers and regulators to better understand how markets work.

The aim of our study is to analyze variations in investor behavior during these phases of market volatility. Given that a large number of factors and events can have an impact on investors, we will concentrate our efforts on studying the influence of a panel of macroeconomic factors. During periods of volatility, these factors can amplify or attenuate investor reactions, thus affecting the dynamics and stability of financial markets. In other words, this study will aim to answer the following research question: How can macroeconomic indicators influence investor behaviour during periods of market volatility?

The aim of our research will be to help investors understand how they behave when macroeconomic factors fluctuate during periods of market volatility. By being aware of the consequences of these fluctuations, this work will provide ideas for strategies to adopt depending on the state of the market, and the strategic orientations to achieve in order to navigate through these turbulent periods.

The main aim of this study is to deepen our understanding of the mechanisms that influence investor behaviour during periods of market volatility, offering valuable insights for financial risk management and the development of strategic economic policies. Additionally, the study aims to carry out a mathematical analysis based on sound econometric techniques to enable the most optimal examination of the results.

This paper will be structured as follows: first, we will study the research already carried out on the subject through our literature review. These will enable us to draw up hypothesis. Next, we will set out the mathematical methodology that will be used to test our hypothesis. Finally, we will interpret the mathematical results obtained using our mathematical methodology, draw conclusions and compare them with what has been developed in our literature review.

2.Literature Review

In this literature review, various factors influencing investors will be analysed, but particular attention will be paid to the influence of macroeconomic factors, which will be detailed later in the study.

To carry out a proper analysis, we also need to consider the influence of behavioural biases on investor decisions. Indeed, it is very important in a study such as this not to neglect this, as it shows that macroeconomic factors are not the only things that impact investor behaviour.

The main objective of this literature review will be to map existing knowledge on the interactions between macroeconomic indicators and investor behaviour, focusing on periods of market volatility. The literature review will conclude with a hypothesis that will be evaluated and studied in the empirical part of this research.

2.1.Definition:

Finance is a field in which the complexity of the vocabulary can sometimes lead to ambiguity or even misunderstanding. In order to avoid any problems, this section aims to eliminate them by defining the various terms and concepts that will be used and explaining why they are important for the analysis and research work carried out in this document. In short, the purpose of this section is to lay the foundations for the study (Dahlquist, 2022).

As the research question focuses on investor behaviour during periods of market volatility, it is essential to define and explain these two words separately and their implications for the research question.

Then, as the study is based on the influence of macroeconomic indicators, they will be explained in this section. In addition, as this is a very broad concept, it is important to select those that will be used in the analysis and to explain why they have been chosen.

To explain investor behaviour, let's go back to the definition of behaviour. It is the manner of being, acting or reacting of human beings (Larousse, n. d.). If we combine this with the concept of investor: Agent or organisation that makes financial investments for investment purposes (Larousse, n. d.), we can conclude that investor behaviour is the way in which these individuals make decisions to buy, hold or sell assets.

Investors are obviously influenced by their emotions, values and beliefs. These influences are significant behavioural biases that must be taken into account if we are to carry out the most accurate analysis possible. Among the most striking biases are the disposition effect and overconfidence (Broihanne et al., 2005). These two concept will be defined in the second part of the literature review.

Market volatility refers to the extent to which the price of a financial asset, whether equities or bonds, fluctuates. Volatility is used to quantify the risk associated with the returns and price of the financial asset to which it is linked. Volatility is an indicator that quantifies the level of risk and price fluctuation (Engle, 2013).

Finally, the macroeconomic indicators that will be used in this analysis are: inflation, interest rates, GDP and the unemployment rate. Each factor will influence investor behaviour in different ways, depending on the effects they have.

Inflation is a situation or phenomenon characterized by a generalized and continuous rise in the price level (Larousse, n. d.).

Secondly, the interest rate is the reference rate set by central banks. Variations in interest rates are the result of different monetary policies. For this reason, it is very important to pay attention to changes in these policies depending on the country in which the interest rate is being studied.

GDP is the abbreviation of the following terms: Gross Domestic Product. It is an indicator that enables us to quantify the total value of a country's wealth production. When we talk about GDP, it is important to distinguish between the nominal and the real. The former is a calculated measure distorted by inflation; it is therefore unadjusted and not accurate for use in this study. In practice, to obtain real GDP, nominal GDP is divided by the price deflator. It is interesting to note that in an inflationary environment, nominal GDP is higher than real GDP (Lokeswar Reddy, 2012).

Finally, the unemployment rate is a measure of the percentage of unemployed in the active population *(Definition - Taux de Chômage | Insee*, n. d.). Like GDP, we need to consider two types of unemployment rate, anticipated and non-anticipated, as they have different implications, as we will see later in this research.

2.2 Literature analysis

In this section, we will discuss in more detail the various points that have been defined above. To achieve this, this part will be divided into different sub-sections.

Firstly, we will discuss the various behavioural biases and how these can influence investors' decisions. Despite the fact that the main objective of this research work does not concern behavioural biases, it is nonetheless important to explore this subject of research as it will enable us to understand in a more optimal way the complexity of financial markets and more specifically the decisions that investors make on them.

Next, we will discuss the influence of macroeconomic indicators. To analyse this influence, we will examine how the various factors affect investor behaviour.

Finally, we look at volatility in the financial markets and the difference between volatility and uncertainty. In this final section, we will look more specifically at the relationship between market volatility and the macroeconomic factors defined above.

2.2.1 Behavioral biases and their impact on investors

In this section, we will analyze the various behavioral biases that can influence investors' decisions. To do so, we will draw on the scientific studies of Broihanne et al. (2005).

As we know, an individual's behavior is never perfectly rational (Broihanne et al., 2005). In fact, various deviations of logical thinking from reality can occur and modify a person's decisions: these are known as cognitive or behavioral biases.

We're mainly talking here about the disposition effect and overconfidence.

The first, the disposition effect, is a behavioral condition that is particularly widespread in trading. Research shows that it encourages investors to sell stocks that could yield a higher return far too early, and to hold on to so-called losers, i.e. stocks that make them lose money (Shefrin & Statman, 1985). The result is investor underperformance. This behavior is explained by loss aversion and an asymmetrical approach between gains and losses. This effect indicates a departure from economic rationality, where decisions should be made on the basis of future prospects rather than historical costs or unrealized gains/losses. This effect was demonstrated by Odean in 1998 in an analysis of 10,000 individual investor accounts. Clearly, this effect has a negative impact on the value of the investor's stock portfolio.

Overconfidence is a problematic behavioral bias. The authors (Broihanne et al., 2005) define overconfidence as the tendency of investors to overestimate their ability to understand market movements, as well as their ability to correctly assess the different values of the assets they are analyzing. By overestimating themselves, investors will make mistakes and sub-optimal decisions that will have an impact on the value of their equity portfolios. This impact is due to the omission of certain important information when making decisions.

The overconfidence theory is also confirmed by the research work of Sifouh and Oumal in 2020. Indeed, in their attempt to demonstrate that investor mimicry could play a role in market volatility, they confirmed that investor overconfidence plays a role in stock price movements. To investigate overconfidence, the authors examined the impact of investor confidence on the volatility of the benchmark Moroccan All Shares Index (MASI) over the period 2002-2017. Their main conclusion is that investors have tended to place more importance on private information at the expense of public information, i.e. company and economic fundamentals.

It is important to note that overconfidence is a form of collective irrationality (Sifouh & Oumal, 2020).

By understanding the existence of these two effects, we'll be able to analyze investor behavior more accurately, and we'll be able to bear in mind the existence of these biases when market movements are irrational.

It will also enable us to better understand the effects of variations in macroeconomic indicators and how investors react to them.

2.2.2 The influence of macroeconomic indicators

To analyze the influence of macroeconomic indicators, we need to look individually at the implications each factor has on investor behavior. We'll analyze them one by one, and then draw a general conclusion.

2.2.2.a The interest rate

The first indicator to be analyzed is the interest rate. When talking about the interest rate, it's important to remember that this indicator is intrinsically linked with two of the other indicators that will be analyzed in this work: inflation and the unemployment rate. Indeed, it is not uncommon for interest rates to be modified by different monetary policies, with the aim of combating them. Numerous studies have shown that interest rates influence investor behavior, as will be explained in this section (Nordin, 2014).

Firstly, studies have shown that the relationship between share prices and interest rates is negative (Ali,2014). Taking the example of an increase in interest rates, the impact will be a decrease in share prices. Indeed, from an individual's point of view, investing in shares becomes less attractive, so he or she will turn to alternatives such as bonds, which become more attractive as a result of the rise in interest rates. By the law of supply and demand, we know that if the number of people wishing to acquire shares falls, then the price of those shares will fall (Alam & Uddin, 2009).

This discovery was also made by Thorbecke (1997). He noted that pro-economic monetary policy measures, such as lower interest rates, lead to higher stock market returns. Lower interest rates reduce the cost of borrowing, stimulate economic activity, and increase corporate profits. This makes shares more attractive, resulting in higher demand and, consequently, higher prices. Understanding this dynamic is crucial for investors and policymakers aiming to manage financial risks and develop effective economic strategies.

It's also important to note that in certain specific cases, a significant rise in interest rates can push investors into a flight to quality. This is a phenomenon in which individuals will opt for safer, more liquid investments. This event was observed in the first half of 2022. During this period, the Fed drastically raised rates on emerging markets to combat inflation, but this led to a flight to quality (Kim,2023) and a consequent fall in the price of US corporate equities.

The observations made in the various studies show the impact that exists between the interest rate and the way investors will react to it.

2.2.2.b Inflation

Inflation is an interesting macroeconomic factor to analyze because of its significant influence on investor behavior and financial markets.

As many studies explain, inflation has a negative impact on markets. High inflation increases credit market frictions, leading to greater credit rationing, reduced intermediary activity, less efficient resource allocation and, consequently, a negative influence on long-term economic performance and stock market activity (Boyd et al., 2001).

It is important to note that there is an inverse relationship between inflation, liquidity and trading volume. In other words, as inflation increases, stock market liquidity decreases, as does trading volume. Another particularly interesting observation from this study is that high inflation is accompanied by high instability in stock market returns, explained by high market volatility. The study by Boyd et al (2001) also shows that at an inflation level of 15%, the performance of the financial sector falls drastically. This is a threshold at which the negative impact is much more pronounced. At this threshold, investors' purchasing power is eroded. What's more, such a high level leads to uncertainty on the markets, resulting in increased stock market volatility. It's also important to note that when inflation levels are so high, central banks may raise interest rates in an attempt to control inflation. Higher interest rates increase financing costs for companies, reducing their earnings and, consequently, their stock market valuations.

Other studies indicate that a reduction in the inflation rate would have a positive impact on share price variations (Lokeswar Reddy, 2012). This observation means that if the inflation rate decreases, this means an increase in consumers' purchasing power, which can then boost their spending and improve companies' profit prospects. An increase in profits could, in turn, induce a rise in share prices.

It's also worth noting that when inflation rises, the central bank may react by raising interest rates, which will do all the things seen in the previous section.

In addition, it is crucial to consider the impact of inflation on investors' asset allocation strategies. As we pointed out earlier, inflationary variations directly influence the correlation between different asset classes, thus modifying optimal allocation strategies. In periods of high and unstable inflation, investors may seek out assets offering protection against inflation, such as gold or real estate. Conversely, in an environment of low or stable inflation, investors may favor assets offering a higher real return, such as equities or long-term bonds. These dynamics highlight the importance of adapting investment portfolios to the inflationary environment.

For example, an investment strategy that has worked well in an inflationary regime may not be as effective in a low-inflation regime, and vice versa. Consequently, investors' ability to understand and anticipate these changes is crucial for effective portfolio management and the preservation of their long-term purchasing power (Signori, 2011).

2.2.2.c Unemployment rate

The third factor to analyze is the unemployment rate. It is important to consider that various studies focus on both expected and unanticipated unemployment rates. Indeed, a number of studies have examined these two aspects to see if they both have an impact on stock prices and markets, and by inference on investor behavior.

Firstly, with regard to unanticipated unemployment, the general belief might lead us to think that it would have a significant impact on stock prices. Contrary to this belief, studies reveal that it has no significant effect on stock market values (Gonzalo & Taamouti, 2017). This belief stems from the assumption that all unexpected news or unforeseen economic events have a direct or immediate influence on financial markets and consequently modifies investor behavior. These thoughts on the unanticipated unemployment rate have been refuted in the work of Gonzalo and Taamouti (2017) through non-parametric Granger causality tests and tests based on quantile regression.

Secondly, with regard to the expected unemployment rate, the conclusions of the research on this subject differ from those obtained for the unanticipated unemployment rate. After various studies, it has been proven that the latter does indeed affect stock prices. Indeed, observations show that when there is an increase in the expected unemployment rate, a rise in share prices is observed.

The reason for this is that when unemployment is high, the Federal Reserve cuts interest rates to stimulate the economy. This lowering of interest rates makes equities more attractive than bonds, leading to a rise in equity prices as individuals seek to acquire a greater proportion of them. This dynamic has been analyzed through the Fisher and Phillips equations (Gonzalo & Taamouti, 2017).

This also highlights the fact that the unemployment rate does not directly affect investor behavior, but rather the consequences that flow from it.

2.2.2.d GDP

The relationship between GDP growth and the state of financial markets is a critical area of study in economics. Several researchers have explored this connection, uncovering insights into how economic growth influences financial market behavior.

Shiblee (2009) conducted an in-depth analysis of the relationship between GDP growth and financial markets. The study concluded that an increase in GDP significantly impacts economic growth, which is generally associated with a rise in share prices. This relationship can be attributed to several factors. As GDP grows, businesses often see improvements in their performance. This is because economic growth usually translates to higher consumer spending, increased production, and better overall business conditions. Improved corporate performance, in turn, makes stocks more attractive to investors. Additionally, economic growth boosts investor confidence. When the economy is expanding, investors tend to be more optimistic about future prospects. This optimism leads to increased demand for equities, as investors anticipate higher returns.

Lokeswar Reddy (2012) further elaborates on the mechanism through which GDP growth affects share prices. As economic growth fuels investor confidence, the demand for equities rises. If the supply of shares remains constant while demand increases, share prices naturally go up. This is a straightforward application of the basic economic principle of supply and demand. Investors' choice of investment during periods of GDP growth is influenced by several macroeconomic factors, including interest rates and inflation. When GDP growth is accompanied by low-interest rates and controlled inflation, the environment becomes more conducive for equity investments. Low-interest rates reduce the cost of borrowing, making it easier for companies to expand and for investors to leverage their positions. Controlled inflation ensures that the purchasing power of returns is maintained.

Positive GDP growth sends a strong signal to the markets about the health of the economy. This sentiment can lead to a virtuous cycle where increased investments further stimulate economic growth. Understanding the relationship between GDP growth and financial markets can help policymakers. For example, during times of sluggish economic growth, policies aimed at stimulating GDP can have the added benefit of boosting financial markets, thereby enhancing overall economic stability. For investors, recognizing the influence of GDP growth on financial markets can inform better investment decisions. During periods of robust GDP growth, investors might prioritize equities over other asset classes, anticipating higher returns from a bullish market.

The relationship between GDP growth and financial markets is a complex but well-documented phenomenon. Studies by Shiblee (2009) and insights from Lokeswar Reddy (2012) highlight the positive correlation between economic growth and share prices, driven by improved corporate performance and heightened investor confidence. This understanding is crucial for investors, policymakers, and economists alike, as it provides a framework for anticipating market movements and formulating strategic responses to economic changes.

2.2.3 Market volatility and its interaction with macroeconomic factors

The concept of volatility has already been explained, but it's important to go into greater depth and develop the differences between market volatility and the concept of market uncertainty. This stage of development is necessary in order to fully understand how volatility can impact macroeconomic factors and consequently investors.

2.2.3.a Uncertainty vs. volatility

Firstly, uncertainty refers to a situation where the future state of our environment is unknown. It is crucial to understand that in conditions of uncertainty, predicting even the slightest possibility or probability of future events becomes impossible. In other words, economic actors cannot forecast the likelihood of future occurrences, a concept first articulated (Knight, 1921).

The distinction between market uncertainty and volatility is important to highlight, as volatility is often a direct consequence of uncertainty. When uncertainty pervades the market, it can lead to heightened volatility. This is because investors, faced with uncertain situations such as pandemics or financial crises, tend to react unpredictably. Their behavior under such circumstances often becomes erratic and dissonant, resulting in more significant and frequent price movements in the financial markets. This increased volatility reflects the collective anxiety and indecision of investors, exacerbating the fluctuations in market prices. Understanding this dynamic is essential for both economic actors and policymakers as they navigate the complexities of financial markets.

2.2.3.b Relationship between volatility and macroeconomic factors

Now that the concept of volatility has been explained, it's important to discuss the relationship between market volatility and various macroeconomic factors such as inflation and interest rates.

This is a complex relationship, as its multifaceted nature makes it difficult to study. Short-term fluctuations in financial markets can be strongly influenced by investors' level of uncertainty (Engle, 2013). The latter is induced by economic announcements, changes in monetary policy and geopolitical events (Wu, 2021).

An example of the complexity of the concept of volatility can be seen in its link with interest rates. Indeed, as we know, a cut in interest rates by a central bank can stimulate investment by making credit cheaper, which generally increases investor confidence and reduces volatility. Conversely, a rate hike can dampen investment and increase volatility. Geopolitical events, such as political tensions or conflicts, can create uncertainty, leading to increased volatility and thus rapid and sometimes large fluctuations in asset prices (Al-Thaqeb & Algharabali, 2019). This influence that all external events can have on volatility creates the complexity mentioned above.

Long-term trends, on the other hand, are often linked to broader economic cycles. During periods of economic growth, markets can be less volatile due to increased investor confidence and a positive economic outlook. Arguably, the low level of market uncertainty during such periods means that volatility levels are low. On the other hand, during a recession, economic uncertainty can increase market volatility(Engle, 2013).

The next relationship to be analyzed is that between inflation and volatility. Numerous studies have shown that inflation has a significant impact on stock market volatility (Engle, 2013) (Boyd, 2001). As explained above, inflation reduces the long-term economic performance of stock markets. We also know that if market inflation is high, then monetary policies will be more aggressive, by raising interest rates for example. If the interest rate rises, investment will slow and market uncertainty will increase. As volatility is a consequence of investor uncertainty, it will also increase.

The final relationship is that of the unemployment rate. An analysis of this relationship was carried out, and the conclusions drawn were that these two concepts influence each other negatively. Indeed, during periods of high volatility, companies are more cautious in their recruitment procedures and hiring decisions. This can lead to an increase in the unemployment rate, which is much more pronounced during periods of market downturn (Pan,2019) (Holmes & Maghrebi, 2016).

Finally, we look at the relationship between GDP growth and market volatility. Numerous studies claim that if the latter is high, then investor confidence in the markets will be reduced, leading to lower investment and consumption. This, in turn, will reduce economic growth. In other words, when volatility is high, investment falls, reducing GDP growth (Equity market volatility and global growth expectations, 2018). Despite these observations, the overall effect of market volatility on GDP needs to be analyzed in the light of many other factors, such as monetary policy and global market conditions.

2.3 Hypothesis generation

To conclude this literature review, we'll use the information we've gathered above to compile them and create an hypothesis to answer the research question based on the literature. As a reminder, the research question is as follows: How do macroeconomic factors influence investor behaviour during periods of market volatility?

It goes without saying that in the rest of this work, we will test the veracity of this hypothesis on the basis of mathematical procedures and advanced econometric calculations. This procedure will be developed in the empirical part of this research work, and we will use it to see whether the findings of the literature are consistent with the mathematical results we obtain.

Volatility is a cause of changes in investor behavior, while at the same time being a consequence of the uncertainty generated by macroeconomic factors. In periods of high volatility, it is highly likely that the average investor will opt for a much more cautious investment strategy, favoring more liquid assets, which can slow down economic activity and thus increase the impact of volatility cycles. On the other hand, in periods of low volatility, investors will have more confidence in the markets, and will therefore allow themselves to take greater risks.

Investor behaviour is also influenced by macroeconomic indicators. Our hypothesis is that the negative impacts of macroeconomic indicators are likely to be accentuated during periods of high volatility, and the positive impacts minimized. According to this hypothesis, an increase in interest rates during a period of high volatility would give investors a much greater incentive to turn to safer investments such as bonds, whereas if this increase took place during a period of low volatility, investors would have a weaker tendency to opt for this type of strategy.

We can maintain the same hypothesis for the inflation rate, i.e. that an increase in inflation during periods of high volatility will amplify the level of uncertainty in the markets, and therefore the effects of volatility. We also know that when inflation rises, investors' purchasing power decreases and so does their propensity to invest. Combined with high market volatility, transaction volume will fall more sharply than in periods of low volatility.

The same conclusions can be drawn for the unemployment rate and GDP as for the other two factors.

On the basis of what we discover, our hypothesis is that the impact of macroeconomic indicators on investor behavior is significantly amplified during periods of high market volatility. The negative impacts of macroeconomic indicators are likely to be more severe during periods of high volatility, while their positive impacts will be diminished. The aim of the empirical part of this research is to test this hypothesis using mathematical procedures and advanced econometric calculations to determine whether the literature's findings align with our results. Clearly, this is no more than a hypothesis for the moment, but the aim of the rest of this work will be to test it in order to justify or disprove it.

3.Research methodology

3.1 Data selection

This study will use mainly quantitative data from the S&P 500 index. This index was chosen because it allows us to observe the 500 largest companies listed on American stock exchanges, and therefore offers us a global and diversified view of the market. We will therefore be able to observe the impact of variations in macroeconomic indicators on this index, which will enable us to draw conclusions later in this work in order to answer the research question.

What's more, given the wide reach of the US markets and their importance, detailed data will be easier to find. It should also be noted that the performance of this market has an influence on world markets, which can be positive for our analysis of changes in investor behaviour.

Next, we will define a time period over which to assess variations in the various macroeconomic factors we will be analyzing. We have chosen a period extending from 2000 to 2020. By analyzing a 20-year period, we'll be able to observe different economic cycles and also analyze different underlying trends and structural changes in the markets. This period will also allow us to include key economic events such as financial crises, changes in monetary policy and economic booms and busts.

For the period under analysis, the presence of two major recessions should be considered (the dot-com crisis in the early 2000s and the financial crisis of 2008), as well as periods of stable growth and recovery. As our analysis will stop at the beginning of 2020, we will not be taking into account events linked to the Covid 19 pandemic.

As regards the various macroeconomic factors, we'll be looking for data on the FRED website, which will enable us to obtain the most accurate data possible.

As for the calculation of market volatility over the research period, we will base it on S&P500 returns over the same period.

To facilitate our analysis, we will work on a monthly basis.

All the data we use can be observed graphically in Appendix 2.

3.2 Model definition

Investor behavior can be influenced by many predictable and unpredictable factors, as demonstrated in the previous section of this work. In this section, the methodology of this work will be explained. This analysis aims to understand the impact of variables such as inflation, interest rate, unemployment rate and GDP on investor decisions, particularly during periods of market volatility. To illustrate investor behavior, we will assume that investors can be observed according to the volume of their transactions.

To begin with, it's important to define the different variables that will be used. The relationship between our independent and dependent variables will be explained using multiple linear regression, with the aim of illustrating changes in investor behavior as a function of the different econometric parameters explained above.

Independent variables :

$$Y = transaction\ volume.$$

Dependent variables :

$$z_1 = inflation\ rate,$$

$$z_2 = interest\ rate,$$

$$z_3 = unemployment\ rate,$$

$$z_4 = GDP.$$

The multiple linear regression formula is as follows:

$$Y = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3 + \beta_4 z_4 + \varepsilon.$$

This technique will enable us to quantify the individual influence of each macroeconomic factor on investor behavior, while controlling for the other factors.

Obviously, this is the simplest regression model, and does not take into account the concept of volatility, which we wish to incorporate into our analysis. For this reason, the multiple linear regression equation has been reworked to obtain the following form:

$$Y = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3 + \beta_4 z_4 + \beta_5 \mathrm{V} + \beta_6 (z_1 \times \mathrm{V}) + \beta_7 (z_2 \times \mathrm{V}) + \beta_8 (z_3 \times \mathrm{V}) + \beta_9 (z_4 \times \mathrm{V}) + \varepsilon,$$

$\beta_1, \beta_2, \beta_3, \beta_4$ represents the effect of macroeconomic factors in the absence of volatility considerations,

β_5 represents the direct effect of volatility on trading volume,

$\beta_6, \beta_7, \beta_8, \beta_9$ represents the effect of the interaction between continuous volatility and each of the macroeconomic variables,

V is a continuous measure of volatility based on monthly S&P 500 returns.

$\beta_6, \beta_7, \beta_8, \beta_9$ variables allow us to see how the impact of each macroeconomic factor evolves with volatility.

3.3 Variable stationarity

In our study, it is also important to test the stationarity of our data. Indeed, stationarity is an essential condition for all statistical series, as it ensures that the statistical properties of the series, such as mean, variance do not change over time. In the context of volatility analysis, it is crucial to ensure that the data remain consistent over time. (Daoud, 2017).

Our analysis led us to investigate different methods for analyzing stationarity. First of all, we thought of using a classic Augmented Dickey Fuller(ADF) test to check whether or not our data showed a unit root. However, after analysis, we realized that this test did not take into account any structural breaks that might be present in the data series. In other words, the model did not take into account the impact of major economic events in the data series. Indeed, the classic ADF test could suggest the presence of unit roots in the presence of structural breaks.

We have therefore opted for a variation of this test, the Phillips-Perron test (Ertur, 1998). This choice is particularly significant due to the main feature of the Phillips-Perron test, which is its ability to correct for the effects of heteroscedasticity and autocorrelation in the error term.

The Phillips-Perron test addresses these issues by using non-parametric statistical techniques to modify the Dickey-Fuller test, which is originally designed to test for a unit root in a time series. By doing so, the Phillips-Perron test provides more robust and reliable results, even in the presence of heteroscedasticity and autocorrelation. This makes it particularly suitable for volatility analysis where such irregularities in the data can commonly occur, ensuring that the conclusions drawn from the analysis are valid and dependable.

Phillips-Perron model:

$$\Delta y_t = \alpha + \beta t + \delta y_{t-1} + \varepsilon_t,$$

H_0: The time series contains a unit root $(\delta = 0)$,

H_1: The time series has no unit root and is therefore stationary. $(\delta < 0)$,

Δy_t is the first difference of the series at time t, y_t being the time-series variable observed at time t. This could be the interest rate, unemployment rate, inflation rate or GDP,

α is the model intercept,

βt represents the time trend,

δy_{t-1} is the coefficient of the one-period lagged series (the term tested for the unit root),

ε_t is the error term.

Test statistics :

$$Z(t) = \frac{\hat{\delta}}{SE(\hat{\delta})},$$

where $SE(\hat{\delta})$ is the standard error of the estimate of δ, adjusted for autocorrelation and heteroscedasticity.

With this test statistic, we can conclude that the series is stationary.

If $Z(t)$ is below the critical value of the Student distribution, then the null hypothesis is rejected.

If, on the other hand, $Z(t)$ is greater than the critical value, then the null hypothesis is not rejected.

Rejection of the null hypothesis would mean that the time series is stationary and therefore the variable is usable in the regression model. It is important to note that the stationarity of each macroeconomic factor will have to be tested in this study. This means that we will adapt the model for each of the indicators to be tested (Cem Ertur, 1998).

In the presence of non-stationarity, it will be important to process the data. To do this, we can perform a simple differentiation whose aim is to stabilize the variance and reduce the effect of extreme values (SPSS Modeler Subscription, n.d.). After this transformation, we'll need to re-test stationarity.

We will also test stationarity by means of a KPSS test, which will enable us to visualize another way of analyzing our data. The KPSS test assesses stationarity by starting from the null hypothesis that the time series is stationary around a trend. This test will reveal whether our data series show stable trends or are subject to structural changes.

First, it is important to define that the y_t series decomposes according to the following scheme:

$$y_t = \mu + \tau_t + \varepsilon_t,$$

μ is a constant,

τ_t represents the trend of the data series,

ε_t is the stationary error term,

The test statistic is calculated as follows:

$$KPSS = \frac{\sum_{t=1}^{T} S_t^2}{T^2 \hat{\sigma}^2},$$

$S_t = \sum_{i=1}^{t} \varepsilon_i$ is the cumulative sum of the regression residuals,

$\hat{\sigma}^2$ is the estimate of the long-term variance of ε_t,

In this particular test, rejecting the null hypothesis means that the series is not stationary. This could imply the presence of a unit root, indicating that the time series has a stochastic trend rather than a deterministic one. According to Hobijn et al. (2004), careful consideration should be given to the choice of the long-term variance estimator and the deterministic components included in the regression, as these can significantly influence the test's power and size properties.

Using the KPSS test in volatility analysis is particularly useful because it allows researchers to confirm whether the statistical properties of a time series, such as mean and variance, remain constant over time. If a series is found to be non-stationary, this suggests that shocks to the system have a persistent effect, which is crucial information when analyzing financial time series data for developing robust econometric models.

3.4 Multi-collinearity of variables

Next, the collinearity between interdependent variables needs to be calculated. Indeed, if different variables are related to each other, problems of multi-collinearity could arise. To test for multi-collinearity, the VIF (Variance Inflation Factor) will be calculated. This analysis method will detect the presence of multi-collinearity between the independent variables in the regression model. If this is detected, it may lead to an increase in the variance of the estimated regression coefficients. An increase in variance will make the coefficients less reliable, and may therefore lead to errors in the results obtained (Stock & Watson, 2006).

A VIF = 1 means that there is no correlation between the different variables. If the VIF is between 1 and 5, this will be considered as an acceptable level of multi-collinearity. On the other hand, if the VIF is greater than 5, multi-collinearity may become problematic and should be examined more closely (Appendix2).

The VIF can be calculated using the following formula:

$$VIF_i = \frac{1}{1 - {R_i}^2},$$

${R_i}^2$ is the coefficient of determination of the independent variable to be analyzed z_i.

In the event that we detect multi-collinearity between different variables, several treatment methods can be applied. If the VIF is too high for certain variables, we might consider removing them from our model. We can also use regression methods such as Ridge regression. The aim of the latter is to limit the size of the coefficients, by constraining their size. By doing so, we can reduce multi-collinearity. By reducing coefficients, Ridge regression can reduce the variance of estimates, increasing model stability.

In the precise case where the multicollinearity between two variables is too high, we'll have to eliminate the one with the highest correlation level. Indeed, ridge regression allows us to reduce the VIF, but the most appropriate solution in a case of severe multicollinearity is to eliminate the variable with the highest correlation level.

3.5 Estimation of coefficients

Having done all this, it's important to note that the β coefficients of the linear regression formula need to be estimated. For this, the least squares method is the most optimal. Thanks to this method, the numerical values expressing the relationship between each independent variable and the dependent variable can be obtained.

The ordinary least squares formula for calculating the coefficients is :

$$\widehat{\beta} = (X'X)^{-1}X'Y,$$

X is the matrix of independent variables,

Y is the vector of the dependent variable,

$\widehat{\beta}$ is the vector of estimated coefficients. This formula finds the values of $\widehat{\beta}$ that minimize the sum of squared residuals between predicted and observed values of Y.

<u>3.6 Significance test</u>

Once the coefficients have been estimated, t and F significance tests must be performed. The t-test will be performed for the individual coefficients, and its aim is to assess whether the regression coefficients have a value other than zero. In other words, it will check whether or not each of the independent variables has a significant effect on the dependent variable.

$$t = \frac{\hat{B}_i - theorical\ value\ of\ \beta_i}{standard\ error\ of\ \hat{B}_i},$$

$\hat{B}_i$ is the OLS estimate of the coefficient.

The theoretical value of β_i is generally zero in the hypothesis test to check whether the coefficient is significant.

The standard error of $\hat{B}_i$ measures the accuracy of the $\hat{B}_i$. The smaller the standard error, the better the estimate of $\hat{B}_i$ is precise.

The null hypothesis is that $\beta_i = 0$, In the above equation, we therefore replace the theoretical value by 0 to test the null hypothesis and check whether it is true. If the value of t is large enough (in absolute value) and exceeds the critical value of the t distribution for the chosen confidence level, then we reject the null hypothesis in favor of the alternative hypothesis. If the null hypothesis is true, this would mean that the independent variable tested has no impact on the dependent variable.

The alternative hypothesis would simply be $\beta_i \neq 0$. It would suggest that the independent variable has a statistically significant effect on the dependent variable.

After analyzing and verifying that each of the independent variables is individually significant, it is necessary to consider them as a whole and perform a test to see whether the independent variables taken together have a statistically significant effect on the dependent variable. The F-test is the test that will analyze this significance. The F-test is a means of assessing the overall significance of the regression model (Stock & Watson, 2006).

The formula for the F-test in the context of multiple linear regression is:

$$F = \frac{MSR}{MSE} = \frac{\frac{SCE}{k}}{\frac{SCR}{(n-k-1)}},$$

where MSR (Mean Square Regression) is the sum of squares due to the regression divided by the number of degrees of freedom of the model (number of explanatory variables),

MSE (Mean Square Error) is the sum of squares of the residuals divided by the degrees of freedom of the residuals (number of observations minus the number of estimated parameters).

SCE is the sum of squares explained by the regression,

SCR is the residual sum of squares,

n is the total number of observations,

k is the number of explanatory variables.

If the calculated value of F is greater than the critical value of the F distribution for a given confidence level (usually 95% or 99%), the null hypothesis is rejected, indicating that the overall regression model is significant.

On the other hand, if the calculated value of F is below the critical value, the null hypothesis is not rejected, suggesting that the overall model is not significant.

3.7 Analysis of residual homoscedasticity

When it comes to regression analysis, it is crucial that the variance of the residuals is constant at all levels of the independent variables. If we find ourselves in a scenario where this assumption is false, then we are in the presence of heteroscedasticity. This concept refers to a situation where the variability of the error term is not constant but varies with the independent variables.

In the situation where the errors are homoscedastic, i.e. the variance is constant, then the least-squares estimators are optimal. In the presence of heteroscedasticity, we'll find ourselves in a state of estimator inefficiency, which will adversely affect our hypothesis testing and thus prevent us from drawing suitable conclusions.

To test for heteroscedasticity, we use the Breusch-Pagan test. This test examines whether the residuals of the regression model show a systematic relationship with the explanatory variables. If the test rejects the null hypothesis of homoscedasticity, this suggests that heteroscedasticity is present (Halunga et al., 2017).

3.8 Autocorrelation of residuals

When carrying out an econometric analysis such as ours, it is also very important to test the autocorrelation of residuals. Indeed, this is an essential condition for guaranteeing the reliability of coefficient estimates and the validity of statistical tests. Analyzing this autocorrelation is all the more important when it comes to interpreting the results of our model. Autocorrelation could lead to inappropriate and incorrect conclusions. To test for autocorrelation, we use the Breusch-Godfrey test. This works by regressing the residuals of the regression model on the lagged values of these residuals as well as on the original explanatory variables. If the test rejects the null hypothesis of no autocorrelation, this indicates that autocorrelation is present and that adjustments need to be made, such as using autocorrelation-corrected regression models or adding autoregressive terms (Shukur, 2000).

In the event of autocorrelation or heteroscedasticity problems, we will opt for a HAC correction to resolve these issues. The aim of this method is to correct standard errors in regression models when the assumptions of homoscedasticity and absence of autocorrelation are not met. By adjusting standard errors appropriately, it enables economists and data analysts to obtain more reliable results and draw more accurate conclusions from their econometric models.

4.Empirical finding

4.1 Preliminary Tests

4.1.1Multi-Collinearity test

As explained in the previous section, it is important to test the correlation between different variables, using a variety of tests. In this situation, we opted for a variation inflation factor (VIF) test to assess the multicollinearity between the 5 key economic indicators. The results of these tests are presented in Appendix 3.

In order to interpret the results, we know that values between 5 and 10 remain acceptable, but above that the collinearity becomes too high to consider the variable in the analysis.

In the Appendix 3, we can observe the level of VIF for all our dependent variables. The VIF levels are correct for all variables except for the level for "inflation and GDP. These values are respectively equal to 142.01 and 148.43. Needless to say, these numbers are far too high and we need to do something about it.

With this in mind, we can assume that our values for inflation and GDP are highly correlated. This suggests a strong interdependence between these variables, which would bias the regression estimates. Knowing the values for VIF, we will remove the GDP variable from our analysis, as it has the highest level of multicollinearity and therefore influences our model negatively to a greater extent than the inflation variable.

To determine the exact level of correlation between these two macroeconomic factors, we created a correlation matrix, as shown in Appendix 4.

Looking at this matrix, we can see that it confirms the previous results, but using a different method. Here, too, we can see the strong correlation between GDP and inflation. This high correlation can be caused by a number of things.

Firstly, we know that in economies with strong growth in gross domestic product, an increase in income can be observed, enabling consumers to spend more. This increase in consumption can push up prices if supply cannot keep pace with growing demand.

Secondly, an increase in GDP can sometimes be accompanied by a rise in production costs (raw materials, wages, etc.), which are then passed on to consumers in the form of higher prices.

As far as the relationship between interest rates and unemployment is concerned, we can consider the correlation to be moderately important, but it may be worth retaining these two variables in our model, as they provide unique and important information despite a certain multi-collinearity.

4.1.2 Test of stationarity of variables

As explained above, working with stationary variables is imperative to ensure the reliability and validity of the analyses that will be carried out. For this reason, we will test it using two separate tests. First, we'll analyze the results of the Phillips-Perron test and compare them with the results of the KPSS test.

In our test, we will consider the critical value at a significance level of 5%, which in our situation is equivalent to a value of -2.87365.

The results of our Phillips-Perron test are shown in the table in Appendix 5.

As we can see, before differentiation, the test statistics obtained were higher than our critical value of -2.87365. We can interpret these results as a non-rejection of the null hypothesis, indicating non-stationarity for all our variables. However, this interpretation does not apply to the volatility variable, which is stationary without the need for differentiation.

To make our other variables stationary, we had to perform a differentiation. After this operation, the results fell below the critical value, enabling us to state that, according to the Phillips-Perron method, our variables are stationary after differentiation.

We then applied the same analytical methodology using the KPSS test to compare the results. The findings from this test are presented in Appendix 6.

At a 5% critical value, we obtained a result of 0.463. Compared to the PP test, the KPSS test produced consistent results with the PP test for four out of the five variables. For inflation, the interest rate, trading volume, and volatility, differentiation was necessary to achieve stationarity in the time series. For the unemployment rate, two differentiations were required.

The combined use of these two tests provides a comprehensive view of the stationarity of the time series. The PP (Phillips-Perron) test, by rejecting the null hypothesis of non-stationarity after differentiation, indicates that an appropriate transformation has been applied to model the series correctly. The KPSS (Kwiatkowski-Phillips-Schmidt-Shin) test confirms these transformations by validating the absence of residual stochastic structure.

Using different methods to analyze stationarity allows us to avoid erroneous conclusions that might arise if our analysis were based solely on the results of a single test. The PP test is particularly useful for identifying the presence of unit roots and determining if a series is non-stationary and requires differentiation. On the other hand, the KPSS test checks for stationarity directly by testing the null hypothesis that an observable series is stationary around a deterministic trend. This dual approach ensures that our transformations are not only necessary but also sufficient for achieving stationarity.

By integrating the insights from both the PP and KPSS tests, we enhance the robustness of our analysis. This multi-faceted approach ensures that we have thoroughly examined the time series data, leading to more reliable and accurate modelling. The confirmation provided by both tests strengthens our confidence in the stationarity of the transformed series and underlines the importance of using multiple tests in time series analysis.

4.1.3 Test of the heteroscedasticity of the residuals

To test for heteroscedasticity in our model, we employed the Breusch-Pagan test. This statistical test is designed to detect the presence of heteroscedasticity, which occurs when the variance of the residuals from a regression model is not constant across all levels of the independent variables. The null hypothesis in the Breusch-Pagan test is that of homoscedasticity, meaning that the residuals have constant variance.

The procedure for the Breusch-Pagan test involves regressing the squared residuals from the original regression model on the original independent variables. The test statistic is derived from this auxiliary regression, and a p-value is calculated to determine the significance of the result.

The test results show a p-value equal to 0.074291 (Appendix 7). This result is obviously higher than our threshold of 0.05, so we can interpret it as a non-rejection of our null hypothesis of homoscedasticity.

In simpler terms, because the p-value exceeds 0.05, we conclude that there is no statistically significant evidence to suggest that heteroscedasticity is present in the residuals of our regression model. This implies that the residuals exhibit constant variance, which is an assumption of the classical linear regression model. Ensuring homoscedasticity is important because it validates the use of ordinary least squares (OLS) estimators, which rely on the assumption of constant error variance to be efficient and unbiased.

Overall, the non-rejection of the null hypothesis of homoscedasticity strengthens the reliability of our regression results, suggesting that our model does not suffer from the potential inefficiencies and biases that heteroscedasticity could introduce. This finding allows us to proceed with further analyses and interpretations with greater confidence in the robustness of our model.

4.1.4 Test for autocorrelation of residuals

In this model, we have opted for the Breusch-Godfrey test to analyze the residuals.

The results of this test can be seen in Appendix 8. As can be seen in our situation, the p-value of the model for order 1 is equal to 0.004966. As this value is less than 0.05, we are obliged to reject the null hypothesis, the latter being the absence of autocorrelation in the model's residuals. In other words, our model (Appendix 9) is autocorrelated.

To correct these autocorrelation problems, we performed a HAR (Heteroscedasticity Autocorrelation Regression) correction. This technique is mainly used to adjust standard error estimates of coefficients in regression models, with the aim of making these estimates robust to heteroscedasticity and autocorrelation. In our situation, we opted for the Newey-West method because it adjusts standard errors to be robust to both heteroscedasticity and autocorrelation. We chose this method over White's, as the latter was unable to correct for autocorrelation. The results are presented in Appendix 10.

The main difference between the pre-correction and post-correction models is the increase in standard errors for most coefficients. This increase indicates that the initial model underestimated the true variability of the coefficient estimates due to the presence of autocorrelation. By applying the Newey-West correction, we obtain standard errors that better reflect the underlying data structure, which includes both heteroscedasticity and autocorrelation.

Furthermore, this adjustment results in a more realistic measure of uncertainty around the estimates, providing a clearer picture of the confidence we can have in our coefficient values. This is particularly important for hypothesis testing and confidence interval construction, as underestimated standard errors can lead to overestimating the significance of predictors.

We can also see a change in the significance of the variables through the value of their p-value. In particular, the value of volatility becomes less significant after correction.

Overall, these corrections make our model more reliable and robust by providing a more accurate representation of the data's underlying variability. This improved reliability enhances the credibility of our model's predictions and inferences, ensuring that our statistical conclusions are well-founded.

4.2 Results interpretation

In this section, we'll set out the mathematical conclusions that our model and these results have enabled us to draw. In each of the sub-sections, we'll look at variations in trading volume with one of the different factors we chose earlier. First, we'll compare the variations in trading volumes with each factor, ignoring volatility. Then, we'll analyze the same variations, but taking volatility into account.

4.2.1 Relationship between transaction volume and interest rate

To begin with, we'll look at the relationship between interest rate and trading volume without considering volatility. The results we can observe in Appendix 10 show us that the relationship between these two elements is significant due to its very low p-value. In fact, with a value equal to 0.0004228, the variables are well below our confidence interval of 0.05. The results also show that a one-unit increase in the interest rate would generate an increase of 2.0137 billion units of volume.

Looking at interest rates alone without the consideration of the volatility variable, a rise in interest rates can be interpreted as a sign of confidence in the economy. Central banks often raise interest rates when the economy is doing well, to avoid overheating.

In analyzing the results of the interaction between interest rate and volatility, we note results that are the opposite of what was previously observed when volatility did not come into play. Indeed, the interaction shows a negative effect, indicating that an increase in interest rates and volatility leads to a decrease in transaction volumes. These results, opposite to those observed previously, are highly significant, with a p-value of 0.0006138.

These results allow us to conclude that in periods of high volatility, interest rate hikes amplify market uncertainty, leading to a reduction in economic activity. More specifically, our analysis indicates that an increase in volatility and interest rates exerts a significant negative effect on the volume of economic transactions.

According to our model, a one-unit increase in volatility and interest rates translates into a 150.9 billion-unit decrease in trading volume. This relationship suggests that, when volatility is high and interest rates rise, companies and investors become more reluctant to engage in transactions. Increased uncertainty and higher borrowing costs prompt them to delay or reduce investment and spending.

The combination of rising interest rates and high volatility increases the risk perceived by investors, prompting them to reduce their trading and adopt a more defensive stance to protect their assets.

4.2.2 Relationship between transaction volume and unemployment rate

As with the interest rate, the relationship between transaction volume and the unemployment rate (without taking volatility into account) is positive and significant. Indeed, with a p-value of 0.0049866, we can easily conclude that the estimator is significant, since this value is less than 0.05. The results also show that a one-unit increase in the unemployment rate would generate an increase of 748.4 million units of volume. This could indicate that higher levels of unemployment are associated with stimulus policies or economic changes that influence transaction volumes.

However, this can be said only if we don't take into account the volatility variable, because when we observe the relationship with the inclusion of the volatility variable, the results are completely opposite. In fact, the relationship becomes negative when we include the concept of volatility, but remains significant with a p-value equal to 0.0067057. The effect of the relationship is negative with a value of -53.75 billion. This may reflect an economic climate where unemployment rises, coupled with volatility, exacerbate economic problems, leading to a decline in confidence and activity. Every time when we include the concept of volatility it induces that there is market uncertainty and it can explain why there is such a difference between the two different model.

In conclusion, although the initial results indicate a positive relationship between the unemployment rate and the volume of transactions, the introduction of volatility into the analysis reveals a more complex dynamic. Increases in unemployment, when accompanied by increased volatility, appear to have a negative effect on the volume of economic transactions. This duality underscores the importance of integrating several contextual variables into economic analyses to obtain a more nuanced understanding of the interactions between different indicators. These results also remind policy-makers of the need to consider the joint effects of these variables when devising effective economic strategies.

4.2.3 Relationship between transaction volume and inflation

Firstly, the relationship between trading volume and inflation, independently of market volatility, is negative. In effect, this means that for an increase in inflation, trading volume will decrease. As we can see in Appendix 10, a one-unit increase in inflation will reduce transaction volume by 270.79 million units.

With a p-value of 0.0001173, we can clearly state that the effect of inflation on transaction volume is statistically significant at the 0.1% threshold. This reinforces the validity of the model.

In contrast to what has just been explained, the results when we include the volatility variable are different. Instead of a negative relationship, as might have been expected, we find ourselves in the opposite situation, with an estimated coefficient equal to 1.7761 billion. This would mean that for a one-unit increase in inflation, we would have an increase of the same value as the estimator for transaction volume.

With a p-value of 0.0043364, we know that the interaction is also statistically significant at the 0.5% threshold.

The interaction between inflation and volatility suggests that periods of high volatility can modify the impact of inflation on the economy. In periods of high volatility, the effect of inflation may be amplified or attenuated, making the analysis of economic dynamics more complex. Further analysis is therefore required to fully understand how these two factors interact to influence markets and economic activity.

The interaction between inflation and volatility requires a nuanced approach. Understanding how these two factors combine to influence markets and economic activity is essential to designing effective economic policies tailored to market realities.

4.3 Model result

In this section, we'll discuss the different results we've obtained for the model as a whole. We will also analyze the quality of the model through the adjusted R-squared and the p-value of the model.

First of all, we have previously discussed the influences that the volatility variable had with the other macroeconomic factors in our model, but we have not discussed this variable independently of the others. We didn't do this because, looking at Appendix 10, we can quickly see that the p-value of this indicator is well above our 5% confidence level, meaning that this variable alone is not statistically significant. It is therefore not interesting to discuss the results linked to its estimator.

Next, it's interesting to analyze the value of the adjusted R-squared we can observe in Appendix 9. Indeed, with a value of 0.1554, it allows us to deduce the following conclusion: the model explains around 15.54% of the variability in transaction volume. This value is small, but we can justify it by the relatively small number of independent variables we have considered in our model. Indeed, in order to have a more complete analysis of transaction volume variance, we could have taken into account many other macroeconomic factors, such as the exchange rate, unexpected events, etc. We believe it would be interesting to deepen our analysis in the future by considering other macroeconomic factors.

Moreover, as we have seen in the literature review, behavioral biases can affect investor behavior. Their actions also have an impact on trading volume, and it is difficult to represent them mathematically. It is therefore important to note that it will never be possible to explain trading volume 100%, as there will always be unpredictable elements affecting it.

Finally, it's important to note that our model is statistically significant due to its p-value of $6.71 * 10^{-8}$, which is obviously below our 5% threshold.

4.4 Comparison between theory and model results

The aim of this final section of our research will be to analyze the various results we have obtained using our mathematical methods, comparing them with what the theory has taught us in the literature review. This comparison will enable us to verify our results and determine whether or not they are consistent.

The interaction of interest rates with financial markets has been extensively analyzed and explained in numerous research studies. The conclusions that emerge from these various studies are unanimous. An increase in the interest rate leads to a reduction in the number of transactions on the markets, as investors prefer to turn to bonds that have become more attractive as a result of the increase. Our empirical results confirm this, but only when we consider market volatility. Indeed, in the absence of volatility, our results show an increase in trading volume when the interest rate rises, which is rather contradictory.

We believe that if the results do not confirm the influence of the interest rate on investor behavior when volatility is not considered, this could indicate that normal market conditions present an equilibrium in which interest rate adjustments are gradually incorporated into asset prices without causing immediate or significant behavioral changes.

Despite these concerns, we can confirm the hypothesis formulated at the end of our literature review. As a reminder, the latter stated that during periods of volatility, the effect of a change in macroeconomic factors would be much more pronounced than in traditional periods.

With regard to the interaction of unemployment rates with financial markets, we had discovered through our research in the literature that when unemployment rates were high, the Fed tended to lower interest rates, which ultimately made equities more attractive and thus increased trading volume. This was confirmed in our empirical results, but only when we consider the unemployment rate factor without taking into account interactions with volatility. When volatility comes into play, the relationship reverses and becomes negative, i.e. an increase in the unemployment rate leads to a decrease in trading volume.

This change in behavior can be explained by investors' risk perception. Indeed, despite low interest rates, in times of high volatility, stock price movements can be more extreme, which may discourage investors. Moreover, high volatility can reduce investors' confidence in the future stability of the market, leading to risk-averse behavior and a minimal desire to invest in risky assets. During these periods, investors prioritize security over returns. Additionally, it is important not to overlook the behavioral biases that can come into play during these times when emotional reactions are stronger.

To conclude this section, we will discuss the interactions between inflation and financial markets. Our literature review showed that the relationship between them was negative. Indeed, numerous studies have demonstrated that when inflation increases, it increases friction in the market and consequently reduces market liquidity, leading to a decline in market performance and transaction volume.

Our empirical results allowed us to reach the same conclusion but only when we look at the relationship without considering volatility. In this specific situation, the relationship is positive. This rather radical change can be explained by a shift in investors' investment preferences. Indeed, in times of high volatility, investors may wish to acquire assets with increased resistance to inflation, such as stocks. It is also possible that some investors might anticipate changes in monetary policies by central banks to counter high inflation during periods of increased volatility. These anticipations could encourage speculative purchases of assets, thereby increasing transaction volume.

Our model showed that it is important to consider global market conditions by taking into account the interactions between different macroeconomic variables to fully understand their impact on financial markets. Indeed, we observed that considering or not considering these interactions had diverse influences on the results.

5. Conclusion

To conclude this work, it is important to emphasize once again the high degree of complexity that governs investor behavior when faced with the vagaries of macroeconomic indicators. With interest rates, inflation, GDP and unemployment rates fluctuating wildly, investors don't know which way to turn. Far from being mere abstractions, these elements are powerful forces that are constantly shaping and reshaping the financial landscape.

Thanks to painstaking data collection, rigorous data processing and the meticulous application of econometric methodology, this study has enabled us to shed light on the significant relationships linking macroeconomic variables, market volatility and investor market behavior. As our study has shown, each variation in a factor can have multiple impacts on investors' buying and selling decisions, as well as on market volatility and other factors. As we have shown, these consequences lead to changes in transaction volumes and asset prices. Added to this complexity are behavioral biases, which also have consequences for investors.

As they are difficult to quantify, it is difficult to predict every single change they will bring to the decisions of the average investor. In light of this, it's clear that understanding the markets is not just a matter of analyzing figures and curves. However, it is very important to understand their mathematical dimension. The insights gained from this study offer valuable perspectives for investors seeking stability in the storm, for policy-makers keen to regulate markets effectively, and for financial regulators aspiring to better risk management.

While this work has enabled us to gain a better understanding of investor behavior, it does not close the door to more in-depth analysis in the future. Indeed, we may extend this analysis to other macroeconomic factors. This will enable us to understand markets even better.

Incorporating global indicators like exchange rates and trade volumes can also offer insights into international influences on local markets. Utilizing advanced techniques like machine learning and scenario analysis can help predict market movements and assess the impact of various economic conditions.

These extended analysis can inform policymakers and investors. Policymakers can design better interventions by understanding the interactions between macroeconomic factors and market behavior. Investors can enhance their strategies by considering a wider range of economic indicators.

By broadening our analysis and using advanced methods, we can achieve a more nuanced understanding of markets and investor behavior, leading to more informed decisions and promoting economic stability and growth.

6. Bibliography

Alam, M. M., & Uddin, G. (2009). *Relationship between Interest Rate and Stock Price : Empirical Evidence from Developed and Developing Countries*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2941281

Ali, H. (2014). Impact of interest rate on stock market ; Evidence from Pakistani market. *IOSR Journal of Business and Management*, *16*(1), 64-69. https://doi.org/10.9790/487x-16176469

Al-Thaqeb, S. A., & Algharabali, B. G. (2019). Economic policy uncertainty : A literature review. *Journal Of Economic Asymmetries*, *20*, e00133. https://doi.org/10.1016/j.jeca.2019.e00133

Boyd, J. H., Levine, R., & Smith, B. D. (2001). The impact of inflation on financial sector performance. *Journal of Monetary Economics*, *47*(2), 221-248. https://doi.org/10.1016/s0304-3932(01)00049-6

Broihanne, M., Merli, M., & Roger, P. (2005). Le comportement des investisseurs individuels. *Revue française de gestion*, *31*(157), 145-168. https://doi.org/10.3166/rfg.157.145-168

Cem Ertur. Méthodologies de test de la racine unitaire. [Rapport de recherche] Laboratoire d'analyse et de techniques économiques(LATEC). 1998, 36 p., Table, ref. bib. : 54 ref. ffhal-01527262f

Dahlquist, J. (2022, 24 mars). *1.6 Microeconomic and Macroeconomic Matters - Principles of Finance | OpenStax*. https://openstax.org/books/principles-finance/pages/1-6-microeconomic-and-macroeconomic-matters#:~:text=Macroeconomics,firm%20had%20to%20consider

Daoud, J. I. (2017b). Multicollinearity and regression analysis. *Journal of physics*, *949*, 012009. https://doi.org/10.1088/1742-6596/949/1/012009

Définition - Taux de chômage | Insee. (s. d.). https://www.insee.fr/fr/metadonnees/definition/c1687

Equity market volatility and global growth expectations. (2018, 24 avril). CEPR. https://cepr.org/voxeu/columns/equity-market-volatility-and-global-growth-expectations

Engle, R. F., Ghysels, É., & Sohn, B. (2013). Stock market volatility and macroeconomic fundamentals. *The Review of Economics and Statistics*, *95*(3), 776-797. https://doi.org/10.1162/rest_a_00300

Federal Reserve Economic Data | FRED | St. Louis Fed. (s. d.). https://fred.stlouisfed.org/

Gonzalo, J., & Taamouti, A. (2017). The reaction of stock market returns to unemployment. *Studies in Nonlinear Dynamics and Econometrics*, *21*(4). https://doi.org/10.1515/snde-2015-0078

Halunga, A. G., Orme, C. D., & Yamagata, T. (2017). A heteroskedasticity robust Breusch–Pagan test for Contemporaneous correlation in dynamic panel data models. *Journal Of Econometrics*, *198*(2), 209-230. https://doi.org/10.1016/j.jeconom.2016.12.005

Hobijn, B., Franses, P. H., & Ooms, M. (2004). Generalizations of the KPSS-test for stationarity. *Statistica Neerlandica/Statistica Neerlandica*, *58*(4), 483-502. https://doi.org/10.1111/j.1467-9574.2004.00272.x

Holmes, M. J., & Maghrebi, N. (2016). Financial Market impact on the real Economy : An assessment of asymmetries and volatility linkages between the stock market and unemployment rate. *The Journal of Economic Asymmetries*, *13*, 1-7. https://doi.org/10.1016/j.jeca.2015.10.003

Kim, J. (2023). Stock market reaction to US interest rate hike : Evidence from an emerging market. *Heliyon*, *9*(5), e15758. https://doi.org/10.1016/j.heliyon.2023.e15758

Knight, F. H. (1921). *Risk, Uncertainty and Profit*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1496192

Larousse, É. (s. d.). *Définitions : comportement - Dictionnaire de français Larousse*. https://www.larousse.fr/dictionnaires/francais/comportement/17728

Lokeswar Reddy, D. V. (2012). IMPACT OF INFLATION AND GDP ON STOCK MARKET RETURNS IN INDIA. *International Journal of Advanced Research in Management and Social Sciences*.

Nordin, N. (2014, 1 décembre). *The impact of commodity prices, interest rate and exchange rate on stock market performance : An empirical analysis from Malaysia*. https://e-journal.uum.edu.my/index.php/mmj/article/view/9015

Odean, T. (1998). Are investors reluctant to realize their losses ? *The Journal of Finance*, *53*(5), 1775-1798. https://doi.org/10.1111/0022-1082.00072

Pan, W. (2018). Does the stock market really cause unemployment ? A cross-country analysis. *The North American Journal of Economics and Finance*, *44*, 34-43. https://doi.org/10.1016/j.najef.2017.11.002

Shefrin, H., & Statman, M. (1985). The disposition to sell winners too early and ride losers too long : theory and evidence. *The Journal of Finance*, *40*(3), 777. https://doi.org/10.2307/2327802

Shiblee, L. S. (2009). The impact of inflation, GDP, unemployment, and money supply on stock prices. *Social Science Research Network*. https://doi.org/10.2139/ssrn.1529254

Shukur, G. (2000). The robustness of the systemwise breusch-godfrey autocorrelation test for non-normal distributed error terms. *Communications In Statistics. Simulation And Computation*, *29*(2), 419-448. https://doi.org/10.1080/03610910008813620

Sifouh, N., & Oubal, K. (2020). Peut-On Expliquer la Volatilité du Marché Boursier Marocain Par un Comportement Mimétique des Investisseurs? Global Journal of Management and Business Research, 20(1).

Signori, O. (2011, 12 septembre). *Inflation-Hedging Portfolios : Economic regimes matter*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1945818

SPSS Modeler Subscription. (s. d.). https://www.ibm.com/docs/fr/spss-modeler/saas?topic=data-series-transformations

Stock, J. H. and Watson, M. W. (2006). Introduction to Econometrics. Second edition, Addison Wesley

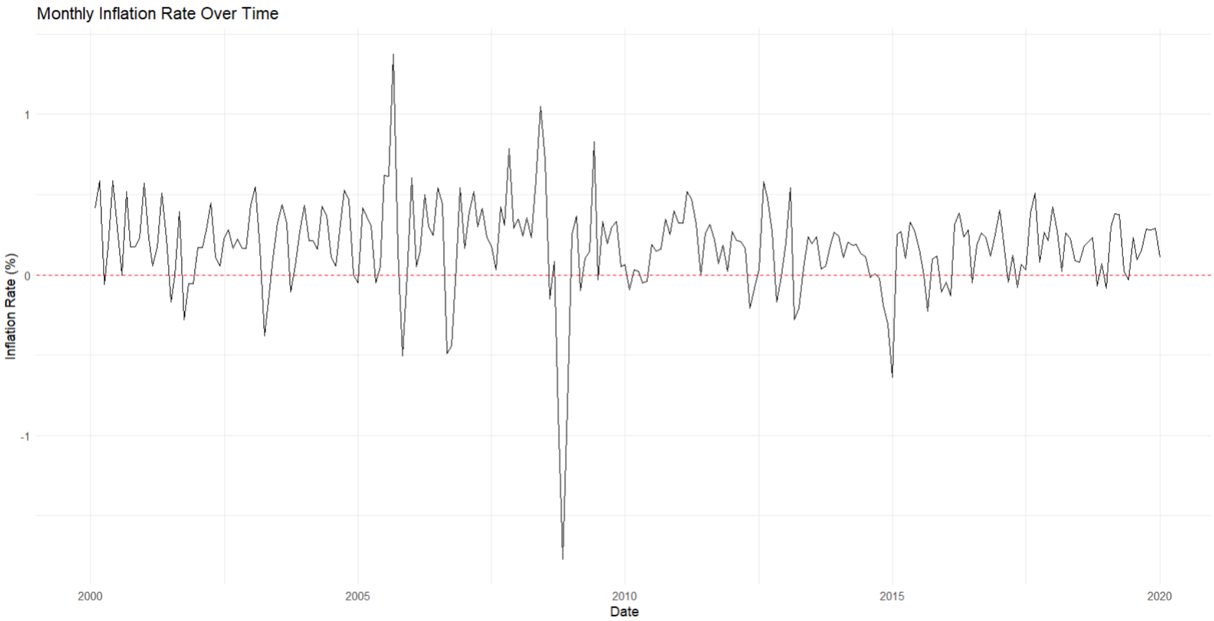
Thorbecke, Willem, 1997. "On Stock Market Returns and Monetary Policy," Journal of Finance, American Finance Association, vol. 52(2), pages 635-654, June.

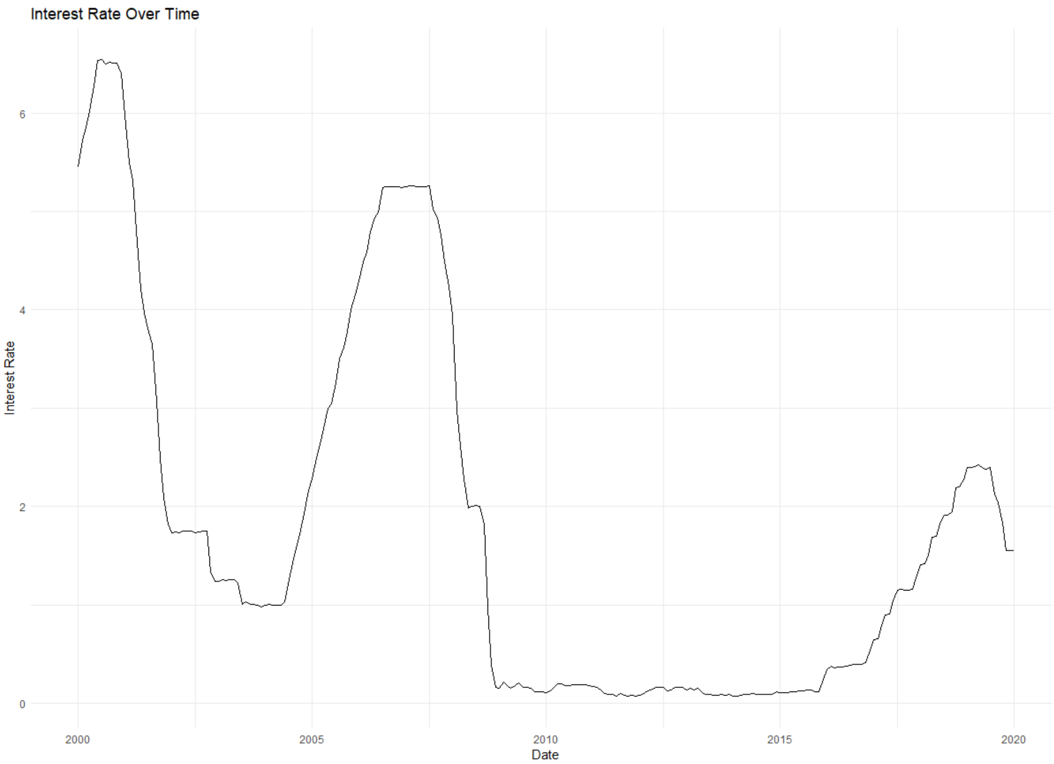
Tsay, R. S. (2010). *Analysis of Financial Time Series*. John Wiley & Sons.

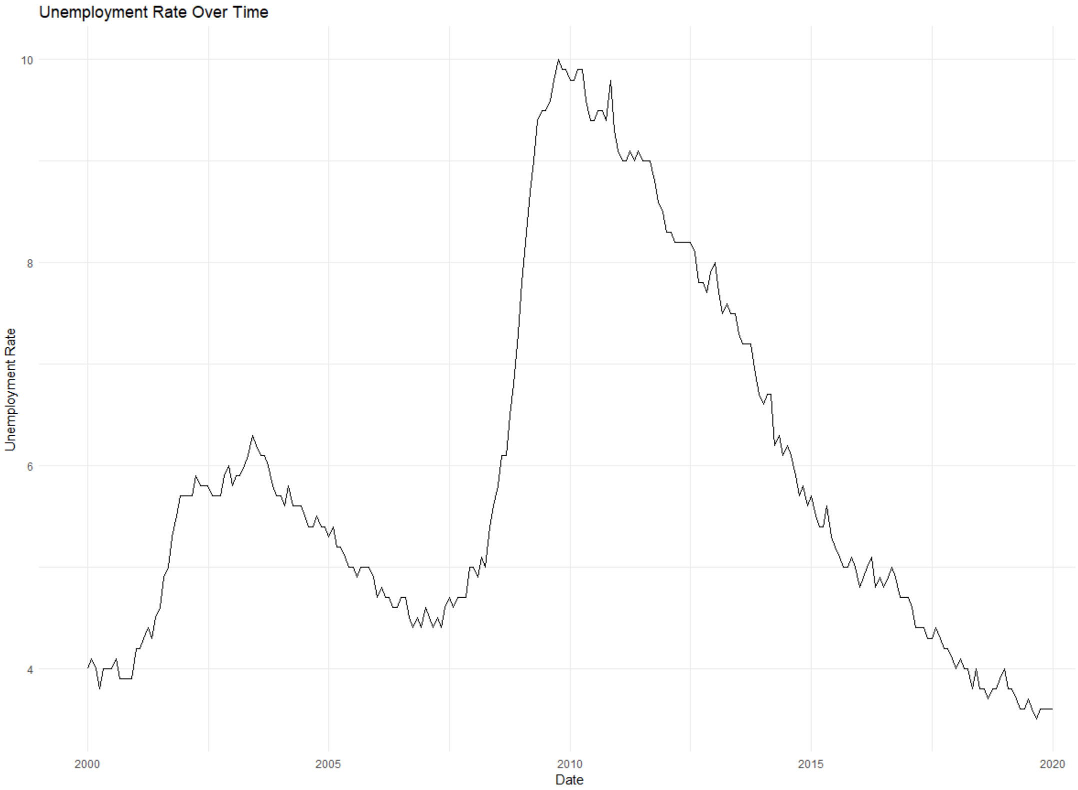
Wu, Z., & Chong, T. T. L. (2021). Does the macroeconomy matter to market volatility? Evidence from US industries. Empirical Economics, 61(6), 2931 2962. https://doi.org/10.1007/s00181-020-02001-3

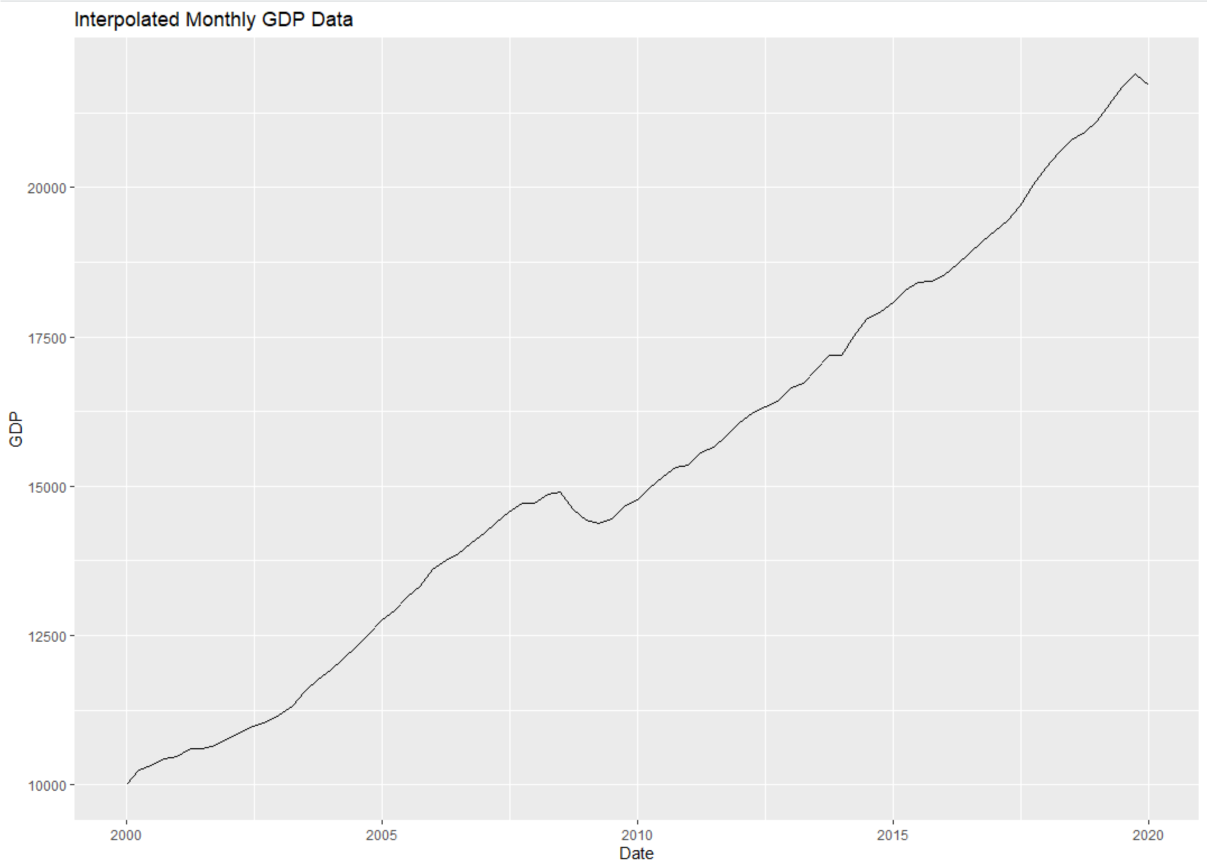
7.Appendices

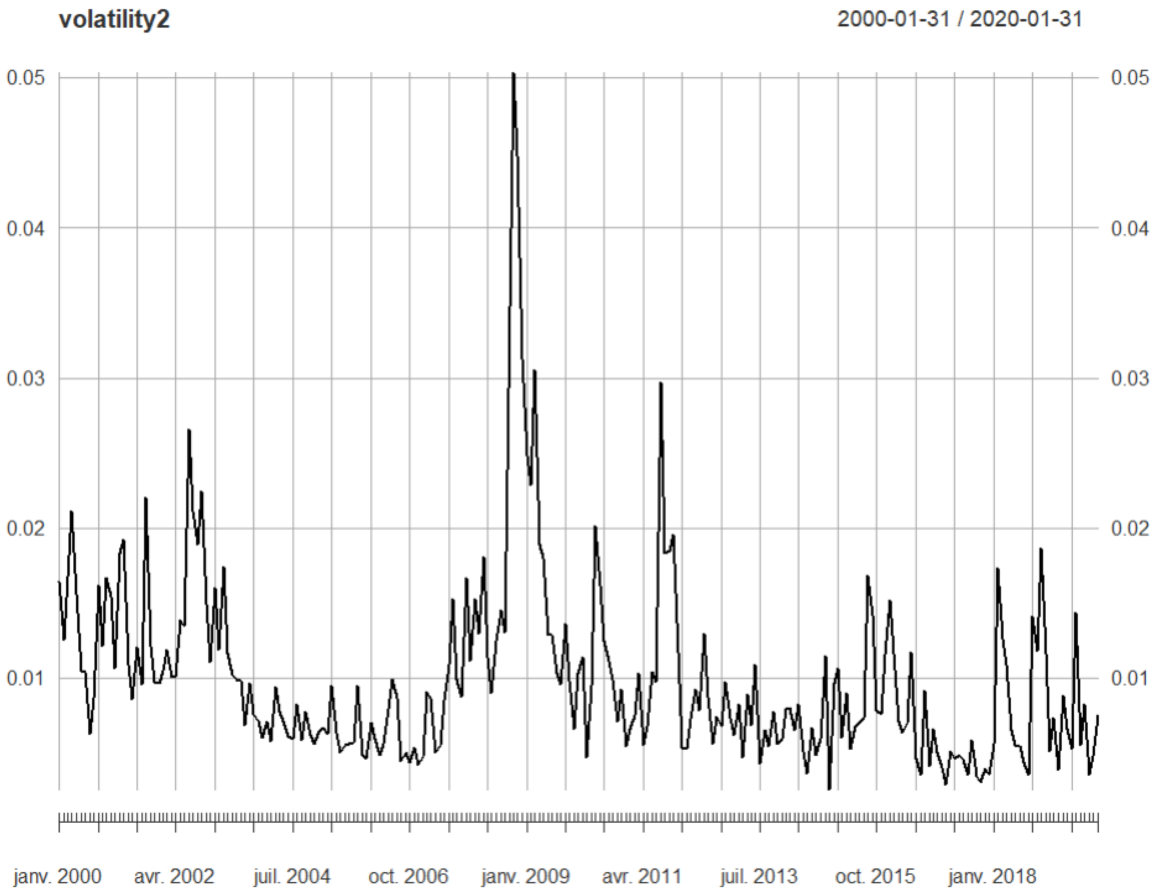
Appendix1:inflation rate, interest rate, unemployment rate, GDP, volatility & S&P500 charts:






Unemployment Rate Over Time
Unemployment Rate
10
8
6
4
2000
2005
2010
2015
2020
Date


Interpolated Monthly GDP Data
GDP
20000
17500
15000
12500
10000
2000
2005
2010
2015
2020
Date



Source : FRED, 2024

Appendix 2 : VIF value meaning

VIF -value	conclusion
$VIF = 1$	Not correlated
$1 < VIF \leq 5$	Moderately correlated
$VIF > 5$	Highly correlated

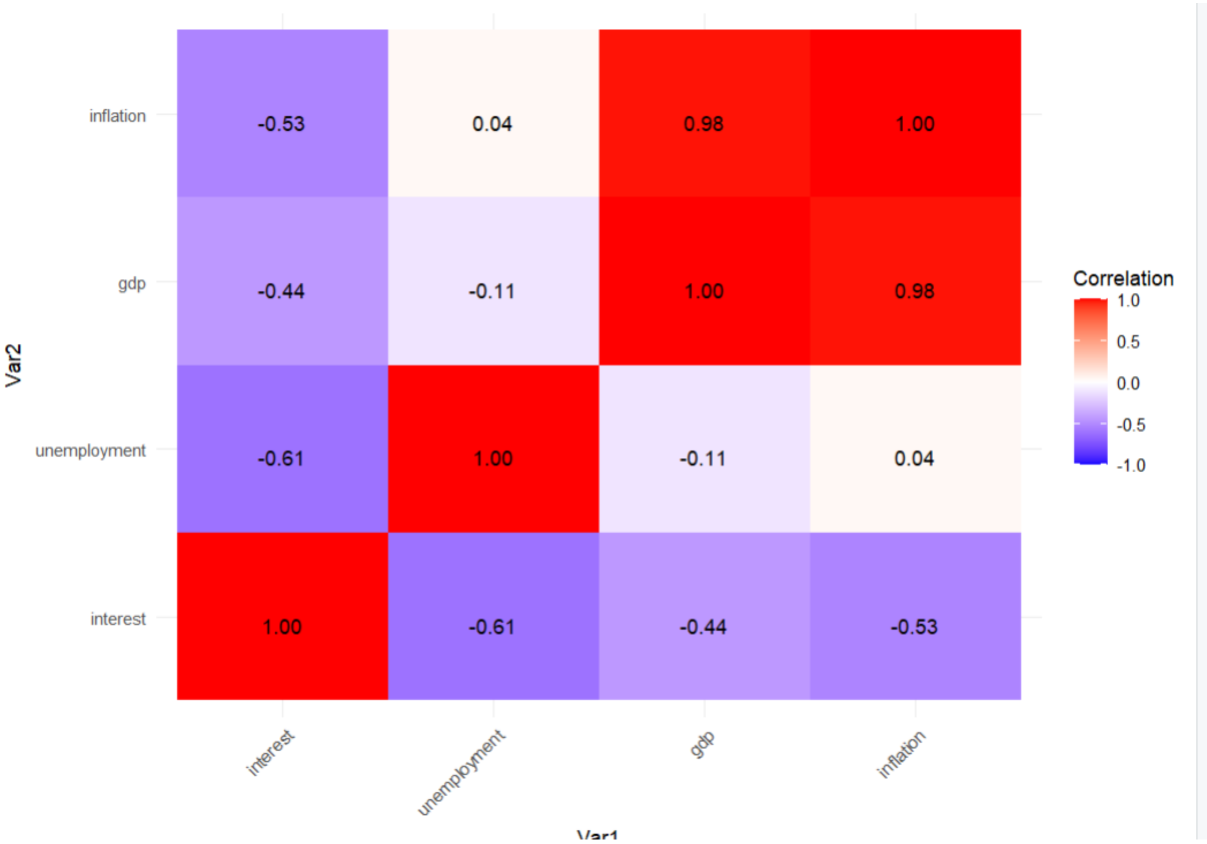
Source : Stock & Watson, 2006

Appendix 3 :Test of multicollinearity

Dataset	VIF
Inflation rate	142.01
Interest rate	2.82

Dataset	VIF
Unemployment rate	5.94
GDP	148.43
Volatility	1.15

Appendix 4 :Correlation matrix between the macroeconomics factors :



Appendix 5 :Result of Phillips-Perron test:

Variable	At Levels	At First Differentiation
Trading Volumes	-2.2231	-22.5089
Volatility	-5.8584	/
Inflation	-0.6583	-9.5039
Interest rate	-1.9718	-6.3409

Unemployment rate	-0.8325	-13.1191

Appendix 6: Result of Kpss test:

Variable	At Levels	At First Differentiation	At Second Differentiation
Trading Volumes	2.4777	0.1003	/
Volatility	0.6059	0.0162	/
Inflation	4.8724	0.0724	/
Interest rate	1.775	0.2603	/
Unemployment rate	0.8287	0.6253	0.0169

Appendix7: Result of our heteroscedasticity test:

```
> ncvTest(model_with_v)#hetéroscédasticité est OK
Non-constant Variance Score Test
Variance formula: ~ fitted.values
Chisquare = 3.185578, Df = 1, p = 0.074291
```

Appendix8 : Result of Breusch-Godfrey test:

Order	Lagrange multiplier	P-value
1	7.8919	0.004966
2	10.723	0.004694

Appendix 9: Summary of our model:

```
> summary(model_with_v)

Call:
lm(formula = volumes ~ interest + unemployment + inflation +
    volatility2 + volatility2 * inflation + volatility2 * unemployment +
    volatility2 * interest, data = donnees)

Residuals:
       Min         1Q     Median         3Q        Max
-1.614e+09 -1.711e+08  5.200e+06  1.647e+08  1.084e+09

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)              -2.956e+07  6.908e+07  -0.428 0.669098
interest                  2.014e+09  4.334e+08   4.647 5.65e-06 ***
unemployment              7.485e+08  3.259e+08   2.297 0.022535 *
inflation                -2.708e+08  7.366e+07  -3.676 0.000294 ***
volatility2               6.550e+09  6.034e+09   1.086 0.278792
inflation:volatility2     1.776e+10  4.390e+09   4.045 7.11e-05 ***
unemployment:volatility2 -5.376e+10  2.203e+10  -2.440 0.015449 *
interest:volatility2     -1.509e+11  2.681e+10  -5.628 5.22e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 420600000 on 232 degrees of freedom
Multiple R-squared:  0.1802,    Adjusted R-squared:  0.1554
F-statistic: 7.283 on 7 and 232 DF,  p-value: 6.761e-08
```

Appendix 10 : HAC correction method of Newey-West:

```
> coeftest(model_with_v,  vcov = NeweyWest(model_with_v))

t test of coefficients:

                            Estimate  Std. Error t value  Pr(>|t|)
(Intercept)              -2.9560e+07  4.1812e+07 -0.7070 0.4802958
interest                  2.0137e+09  5.6293e+08  3.5771 0.0004228 ***
unemployment              7.4848e+08  2.6401e+08  2.8350 0.0049866 **
inflation                -2.7079e+08  6.9104e+07 -3.9185 0.0001173 ***
volatility2               6.5500e+09  5.0752e+09  1.2906 0.1981300
inflation:volatility2     1.7761e+10  6.1651e+09  2.8809 0.0043364 **
unemployment:volatility2 -5.3757e+10  1.9650e+10 -2.7357 0.0067057 **
interest:volatility2     -1.5091e+11  4.3452e+10 -3.4730 0.0006138 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Appendix 11 :code for the data importation and code for the volatility calculation

```
#### Définissez votre clé API pour FRED####
fredr_set_key("c44d049b94b5aa776746d422191a3053")

#### Téléchargement des données du S&P 500####
getSymbols("^GSPC", src = "yahoo", from = "1999-12-01", to = "2020-01-01")
# Accéder aux données de clôture journalière
volume_data <- GSPC[, "GSPC.Volume"]
# Convertir les données journalières en moyenne mensuelle
sp500_monthly_data <- apply.monthly(volume_data, mean)

#### Téléchargement des données des varibales dépendantes####
inflation_data <- fredr(series_id = "CPIAUCSL", observation_start = as.Date("2000-01-01"), observation_end = as.Date("2020-01-01"))
interest_data <- fredr(series_id = "FEDFUNDS", observation_start = as.Date("2000-01-01"), observation_end = as.Date("2020-01-01"))
unemployment_data <- fredr(series_id = "UNRATE", observation_start = as.Date("2000-01-01"), observation_end = as.Date("2020-01-01"))
gdp_data <- fredr(series_id = "GDP", frequency = "q", observation_start = as.Date("2000-01-01"), observation_end = as.Date("2020-01-01"))

### Conversion des données FRED en série temporelle
interest_ts <- zoo(interest_data$value, order.by = interest_data$date)
unemployment_ts <- zoo(unemployment_data$value, order.by = unemployment_data$date)

#### Interpolation linéaire pour le GDP####
# Création de l'objet zoo avec les données d'inflation
ts_gdp <- zoo(gdp_data$value, order.by = as.Date(gdp_data$date))

# Création d'une séquence de dates mensuelles à partir de la première à la dernière date
date_seq <- seq(start(ts_gdp), end(ts_gdp), by = "month")

# Interpolation linéaire sur la grille mensuelle
ts_gdp_monthly <- na.approx(ts_gdp, xout = date_seq)

#### Interpolation linéaire pour l'inflation####
# Création de l'objet zoo avec les données d'inflation
ts_inflation <- zoo(inflation_data$value, order.by = as.Date(inflation_data$date))

# Création d'une séquence de dates mensuelles à partir de la première à la dernière date
date_seq <- seq(start(ts_inflation), end(ts_inflation), by = "month")

# Interpolation linéaire sur la grille mensuelle
ts_inflation_monthly <- na.approx(ts_inflation, xout = date_seq)

#donnée utilisable
sp500_data2<- sp500_monthly_data
inflation_data2<-ts_inflation_monthly
gdp_data2<-ts_gdp_monthly
interest_data2<-interest_ts
unemployment_data2<-unemployment_ts
```

```
####volatilité####
getSymbols("^GSPC", src = "yahoo", from = "1999-12-01", to = "2020-01-01")
prices2 <- Cl(GSPC)

log_returns <- ROC(prices2, type = "continuous")
# Calcul de l'écart-type mensuel des rendements (volatilité)
volatility_by_month <- apply.monthly(log_returns, sd)
# Nettoyer les valeurs NA (premier mois pourrait contenir NA)
volatility_by_month <- na.omit(volatility_by_month)

result_volatility <- ur.pp(volatility_by_month, type="Z-tau")
volatility2<-c(volatility_by_month)
```

Appendix 12 :code the Phillips-Perron test before differentiation

```
####testPP####
result_inflation <- ur.pp(inflation_data, type="Z-tau")
result_interest <- ur.pp(interest_data, type="Z-tau")
result_unemployment <- ur.pp(unemployment_data, type="Z-tau")
result_sp500_monthly_data <-ur.pp(volumes, type="Z-tau")
```

Appendix 13 :code the Phillips-Perron test after differentiation

```
####différentiation####
inflation_data_diff <- diff(inflation_data2)
interest_data_diff <- diff(interest_data2)
unemployment_data_diff <- diff(unemployment_data2)
volumes_diff<-diff(sp500_monthly_data)
volumes_diff<-na.omit(volumes_diff)
####testPP####
result_inflation_diff <- ur.pp(inflation_data_diff, type="Z-tau")
result_interest_diff <- ur.pp(interest_data_diff, type="Z-tau")
result_unemployment_diff <- ur.pp(unemployment_data_diff, type="Z-tau")
result_sp500_monthly_data <-ur.pp(volumes_diff, type="Z-tau")
```

Appendix 14 :Code for the model explanation and others test

```
#### Création dataframe####
donnees <- data.frame(Volumes = volumes, Inflation = inflation, Interest = interest, Unemployment = unemployment, volatility = volatility2)

model_with_v <- lm(volumes  ~ interest + unemployment + inflation + volatility2 + volatility2:inflation + volatility2:unemployment + volatility2:interest, data = donnees)

summary(model_with_v)
####test autocorrélation####
bgtest(model_with_v,order=1)
####
ncvTest(model_with_v)
#### test heteroscedasticity###
bptest(model_with_v)
####HAR correction####
coeftest(model_with_v,  vcov = NeweyWest(model_with_v))
```

Appendix 15: code for VIF

```
# Calcul du VIF pour tester la multicolinéarité
vif_results <- vif(model_with_v)
print(vif_results)
```

Appendix16: code for the KPSS test before differentiation

```
#donnée utilisable
sp500_data2<- sp500_monthly_data
inflation_data2<-ts_inflation_monthly
gdp_data2<-ts_gdp_monthly
interest_data2<-interest_ts
unemployment_data2<-unemployment_ts

# Test KPSS pour le trading volume
kpss_result_sp500 <- ur.kpss(sp500_data2)
summary(kpss_result_sp500)

# Test KPSS pour l'inflation
kpss_result_inflation <- ur.kpss(inflation_data2)
summary(kpss_result_inflation)

# Test KPSS pour le taux d'intérêt
kpss_result_interest <- ur.kpss(interest_data2)
summary(kpss_result_interest)

# Test KPSS pour le taux de chômage
kpss_result_unemployment <- ur.kpss(unemployment_data2)
summary(kpss_result_unemployment)

# Test KPSS pour le volatilité
kpss_result_volatility <- ur.kpss(volatility2)
summary(kpss_result_volatility)
```

Appendix17: code for KPSS test after differentiation

```
#différentiation
sp500diff<-diff(sp500_data2)
inflation_data_diff <- diff(inflation_data2)
interest_data_diff <- diff(interest_data2)
unemployment_data_diff <- diff(diff(unemployment_data2))
volatility_diff<-diff(volatility2)

#test kpss postdiff
result_sp500_diff<-ur.kpss(sp500diff)
summary(result_sp500_diff)

result_volatility_diff<-ur.kpss(volatility_diff)
summary(result_volatility_diff)

result_inflation_diff <- ur.kpss(inflation_data_diff)
summary(result_inflation_diff)

result_interest_diff <- ur.kpss(interest_data_diff)
summary(result_interest_diff)

result_unemployment_diff <- ur.kpss(unemployment_data_diff)
summary(result_unemployment_diff)
```

Appendix 18: code for the creation of the correlation matrix

```r
# Assurez-vous que les noms de colonnes sont corrects et essayez à nouveau
cor_matrix <- cor(data[, c("interest", "unemployment","gdp", "inflation")])

# Affichage de la matrice de corrélation
print(cor_matrix)


# Pour une meilleure visualisation, vous pouvez utiliser la bibliothèque ggplot2 pour créer une image graphique de la matrice de corrélation
library(ggplot2)
library(reshape2)  # pour la fonction melt()

# Transformation de la matrice en format long pour ggplot
cor_data <- melt(cor_matrix)

# Création du graphique de corrélation
ggplot(data = cor_data, aes(x = Var1, y = Var2, fill = value)) +
  geom_tile() +  # Créer un remplissage de tuiles pour la matrice
  geom_text(aes(label = sprintf("%.2f", value)), vjust = 1) +  # Ajouter les coefficients de corrélation
  scale_fill_gradient2(low = "blue", high = "red", mid = "white", midpoint = 0, limit = c(-1,1), space = "Lab", name="Correlation") +
  theme_minimal() +  # Utiliser un thème minimaliste
  theme(axis.text.x = element_text(angle = 45, vjust = 1, hjust=1))  # Incliner les étiquettes de l'axe des x pour une meilleure lisibilité
```