Analysing behavioural factors that impact financial stock returns
The case of COVID-19 pandemic in the financial markets

Thamila Madji

https://orcid.org/0000-0001-8840-1068

A thesis submitted in partial fulfilment of the requirements of London South Bank University for the degree of Doctor of Philosophy

October 2023

Preface

This thesis represents a pivotal advancement in the realm of behavioural finance, seamlessly integrating both classical and state-of-the-art models. It navigates the performance and applicability of the Irrational Fractional Brownian Motion (IFBM) model, while also delving into the propagation of investor sentiment, emphasizing the indispensable role of hands-on experiences in understanding, applying, and refining complex financial models.

Financial markets, characterized by 'fat tails' in price change distributions, often challenge traditional models such as the Geometric Brownian Motion (GBM). Addressing this, the research pivots towards the Irrational Fractional Brownian Motion Model (IFBM), a groundbreaking model initially proposed by (Dhesi and Ausloos, 2016) and further enriched in (Dhesi et al., 2019). This model, tailored to encapsulate the 'fat tail' behaviour in asset returns, serves as the linchpin for the first chapter of this thesis.

Under the insightful guidance of Gurjeet Dhesi, a co-author of the IFBM model, we delved into its intricacies and practical applications. The first chapter aspires to evaluate the IFBM's performance in real-world scenarios, enhancing its methodological robustness. To achieve this, a tailored algorithm was crafted for its rigorous testing, alongside the application of a modified Chi-square test for stability assessment. Furthermore, the deployment of Shannon's entropy, from an information theory perspective, offers a nuanced understanding of the model. The S&P500 data is wielded as an empirical testing bed, reflecting real-world financial market dynamics. Upon confirming the model's robustness, the IFBM is then applied to FTSE data during the tumultuous COVID-19 phase. This period, marked by extraordinary market oscillations, serves as an ideal backdrop to assess the IFBM's capability in tracking extreme market shifts.

Transitioning to the second chapter, the focus shifts to the potentially influential realm of investor sentiment, seen as one of the many factors contributing to fat tails' presence in return distributions. Building on insights from (Baker and Wurgler, 2007), we examine the potential impact of political speeches and daily briefings from 10 Downing Street during the COVID-19 crisis on market sentiment. Recognizing the profound market impact of such communications, the chapter seeks correlations between these briefings and market fluctuations.
Employing advanced Natural Language Processing (NLP) techniques, this chapter harnesses the power of the Bidirectional Encoder Representations from Transformers (BERT) algorithm (Devlin et al., 2018) to extract sentiment from governmental communications. By comparing the derived sentiment scores with stock market indices' performance metrics, potential relationships between public communications and market trajectories are unveiled. This approach represents a

melding of traditional finance theory with state-of-the-art machine learning techniques, offering a fresh lens through which the dynamics of market behaviour can be understood in the context of external communications.

In conclusion, this thesis provides an intricate examination of the IFBM model's performance and the influence of investor sentiment, especially under crisis conditions. This exploration not only advances the discourse in behavioural finance but also underscores the pivotal role of sophisticated models in understanding and predicting market trajectories.

Dedication and acknowledgements

I dedicate this work to my dear mother Dalila, who gave up everything for me. Her hard work has always been and will continue to be an inspiration.

I would also like to dedicate this work to my loving husband Adam, who has been a constant source of support and encouragement during the challenging times of this thesis. I love you, and I am truly thankful for having you in my life.

I would also like to give a special thank you to my thesis supervisor Valerio, who stayed in the fight to the very end to complete this work. You have been a true motivation, and without you this would not have been possible.

Lastly, I dedicate this work to my sister Katiya and brother Nazime, whom I love so deeply and who hold a special place in my heart. To my grandparents, who raised me with boundless love and instilled values that have shaped me into the person I am today, my gratitude knows no bounds. Your legacy is an integral part of this journey and this work.

Table of Contents

List of Tables

List of Figures

Figure **Page**

Chapter 1

The Irrational Fractional Brownian Motion and the Shannon Entropy

1.1 Introduction

Traditional finance theories provide invaluable insights into the investment decision-making process by observing and investigating price behaviours. These theories are fundamentally grounded on the rational expectation of investors and the efficient market hypothesis (EMH) (Mill (1874); Fama (1970)). EMH, initially proposed by Fama (1970), asserts that financial markets are always efficient in reflecting all available information, hence no possibility of earning abnormal returns.

A cornerstone of such traditional financial theories, which assume rationality and market efficiency, is the Geometric Brownian Motion (GBM). Introduced by mathematician Louis Bachelier in his 1900 thesis "The Theory of Speculation" (Davis and Etheridge, 2006), GBM is a statistical process that has been widely used in financial modelling, especially in options pricing models, such as the seminal Black-Scholes-Merton model (Black and Scholes, 1973). The GBM model is widely used in finance to simulate the behaviour of asset prices under the assumption that price changes are normally distributed and independent of each other. It assumes that small, continuous changes, which are a product of numerous random events, lead to a Gaussian distribution of price changes, conforming to the principles of the EMH ((Black and Scholes, 1973),(Merton, 1973)).

However, a widely recognized critique of traditional financial theories like EMH concerns the assumption that asset price fluctuations follow a Gaussian or normal distribution characterized by a symmetrical bell-shaped curve. EMH assumes that changes in asset prices should ideally follow a Gaussian or normal distribution, considering that these changes result from the arrival of new, independently and identically distributed information.

However, numerous empirical studies have revealed that these fluctuations often exhibit a behaviour known as leptokurtosis, resulting in 'fatter tails' and a higher peak than a normal distribution. Some notable studies include those by ((Mandelbrot, 1963),(Taleb, 2007),(Rachev et al., 2007)), among others. This observation effectively challenges EMH and similar traditional financial theories. The term "fatter tails" refers to the increased likelihood of large price changes, resulting in more volatile and less predictable markets, in contrast to the thin tails of a Gaussian distribution, which implies that extreme events are exceedingly rare.

In recognition of these limitations, the field of behavioural finance has emerged to provide alternative perspectives on investor behaviour and financial market dynamics. At its core, behavioural finance incorporates insights from psychology and social sciences to better understand how irrationality and cognitive biases can influence investment decisions. Numerous studies have shown that human investors do not always behave rationally, often being influenced by psychological factors such as overconfidence, loss aversion, and herd behaviour ((Shefrin, 2002), (Barberis, 2018)).

For instance, herd behaviour, where investors follow the crowd rather than their own analysis, can contribute to the observed 'fat tails' in asset price distributions by leading to sudden, large-scale buying or selling activities (Bikhchandani et al., 1992). Moreover, behavioural biases can lead to over- or under-reaction to information, resulting in price movements that are more extreme than what would be expected under the EMH.

Therefore, traditional financial models may not fully account for the complexity of market dynamics due to behavioural influences. To tackle this limitation, contemporary models have begun to incorporate behavioural factors to more accurately represent market phenomena (De Bondt and Thaler, 1985; Shleifer and Summers, 1990; Barberis et al., 1998).

An innovative approach to understanding financial market dynamics is the "Irrational Fractional Brownian Motion" (IFBM) model developed by Dhesi and Ausloos (2016). This model seeks to integrate the irrational decision-making of market participants with the impact of time-dependent news on financial markets. Central to the IFBM model is the concept of 'soliton behaviour', a term that has its origins in physics. In physics, a soliton is a self-reinforcing solitary wave that maintains its shape while propagating at a constant velocity. Notably, solitons are resilient, retaining their shape even after colliding with other solitons(Drazin and Johnson, 1989).

The IFBM model employs the soliton concept to mathematically mimic certain characteristics of these solitary waves, illustrating how specific market trends can maintain their form despite new information or market events. While the soliton serves as an analogy, it is grounded in the mathematical function of the IFBM model, making it more than a mere metaphor.

Behavioural finance research indicates that investors often base their decisions not just on raw information or news, but on how they perceive other market participants reacting to it (Barber and Odean, 2008). For instance, suppose there is a significant piece of news about a company that is expected to impact its stock price. Initially, a small group of investors reacted to this news, leading to a minor fluctuation in the stock price. However, as other market participants observe this price movement and interpret it as a reaction to the news, they, too, may decide to buy or sell the stock, leading to further price movement. This cascading effect can continue, leading to a pronounced trend or pattern in the market that, like a 'soliton' wave, maintains its form even in the face of subsequent market events (De Bondt and Thaler, 1985; Shiller et al., 1981; Hong and Stein, 1999). This suggests that the 'soliton' behaviour observed in the IFBM model may be indicative of underlying behavioural phenomena in financial markets, where the response to information can be magnified by the reactions of other market participants.

One of the significant advantages of the IFBM model over traditional Geometric Brownian Motion (GBM) models is its ability to capture the leptokurtic nature of financial returns, a characteristic often observed in real-world markets but missed by GBM models (Mandelbrot, 1963). The connection between irrational behaviour and leptokurtosis is pivotal. When investors overreact to others' actions, it can lead to extreme price movements, resulting in 'fat tails' in the distribution of price changes (Dhesi et al., 2016). The 'soliton' behaviour in the IFBM model, therefore, is not just an analogy but a mathematical representation of certain market dynamics. It sheds light on complex market behaviours, emphasizing the role of irrational decision-making and highlighting the influence of time-dependent news. This makes the IFBM model a valuable tool for a more nuanced understanding of market dynamics, accommodating both irrational behaviours and the leptokurtic nature of financial returns.

It is worth highlighting that the focus of this study is not to identify these 'solitons' or persistent market trends per se but rather to delve deeper into the intricacies of the Irrational Fractional Brownian Motion (IFBM) model. This model distinctively incorporates soliton-like behaviour in its framework. While the IFBM model's accuracy in replicating the actual distribution of financial returns has been previously compared with traditional models like the Geometric Brownian Motion (GBM), our study aims to further this investigation. This study primarily focuses on exploring the stability and robustness of the IFBM model and developing an algorithm to improve its repeatability in financial market analyses. A secondary perspective involves applying tools from information theory to the IFBM model, allowing us to investigate the information content carried in the market dynamics under different conditions.

We aim to critically evaluate the stability and robustness of the IFBM model, using historical S&P500 data as our testing ground. Additionally, the algorithm we develop seeks to address potential limitations and broaden the model's practical applicability.

To gain a different perspective on the IFBM model, we will employ tools from information theory, such as Shannon entropy. Shannon's entropy quantifies the amount of 'information' or 'surprise' a particular event brings (Shannon, 1948). By using this measure, we aim to comprehend the information content carried in the market dynamics when the IFBM model is applied. This information-theoretic viewpoint allows us to understand the model's behaviour and implications better, providing a novel lens through which to examine market dynamics. With this foundation, we then apply the IFBM model to the COVID-19 crisis period using

FTSE data. The choice of the FTSE market, a different market from the S&P500, allows us to validate the model's versatility and adaptability in various market contexts. During such periods of market volatility, investor irrationality and information dynamics can significantly influence market movements (Kahneman and Tversky, 2013; Shiller et al., 1981; Tversky and Kahneman, 1992).
This comprehensive investigation, therefore, offers a fresh perspective on investor behaviour, merging concepts from behavioural finance, mathematical finance, and information theory. The objective is to foster a deeper understanding of financial markets, transcending the limitations of traditional finance theories by accounting for behavioural factors and harnessing insights from information theory.

In the following sections of this chapter, we embark on an in-depth examination of the underpinnings of these financial theories. Our investigation spans the traditional financial paradigms rooted in rational expectations and market efficiency, as well as modern frameworks that recognize the profound influence of behavioural biases and irrationality in investor decision-making. Subsequent to the background analysis, we will then outline the motivation underpinning this research, laying out the research objectives this chapter of the thesis seeks to address. Here, we will articulate why the IFBM model and the unique perspective it provides on market dynamics are worthy of detailed investigation and our contribution to it.
In subsequent sections, we transition into an in-depth examination of the origins and evolution of the IFBM Model, which constitutes the focal point of this study.
Further, we will describe the datasets to be used in our study and outline the methods employed for the analysis. The final chapters of the thesis will present the results of our analysis. Ultimately, the thesis will conclude by discussing the key findings, their implications, and potential directions for future research.

1.2 Background study

Traditional financial theories, such as the Efficient Market Hypothesis (EMH) and the Capital Asset Pricing Model (CAPM), have been instrumental in shaping our understanding of financial markets. However, these theories also have inherent limitations, predominantly stemming from their underlying assumptions.
One of these key assumptions is that asset price changes follow a Gaussian or normal distribution characterized by its symmetrical bell-shaped curve. In such a distribution, extreme price changes – those in the 'tails' of the distribution – are extremely rare. This assumption underlies various aspects of traditional finance theories, such as the idea that financial markets efficiently incorporate all available information into asset prices (EMH) or that investors can optimize their portfolios by balancing expected returns against the standard deviation of returns (CAPM)

(Fama (1970); Sharpe (1964)).
However, empirical evidence often contradicts this assumption. Asset price changes frequently display a distribution pattern known as leptokurtosis, characterized by 'fatter tails' and a sharper peak compared to a normal distribution. This pattern indicates a higher occurrence of extreme price changes, leading to more volatile and less predictable markets (Mandelbrot (1963); Campbell et al. (1997)).
Benoit Mandelbrot was among the pioneers who challenged the assumption of a Gaussian distribution in financial market returns. In his 1963 paper, Mandelbrot argued for the presence of 'fat tails' in financial returns and introduced the idea of using fractal geometry and stable Paretian distributions to better describe these phenomena (Mandelbrot (1963)). His work inspired subsequent research into modelling asset price changes with non-Gaussian distributions. In "The Econometrics of Financial Markets," Campbell, Lo, and MacKinlay (Campbell et al. (1997)) highlighted the need for alternative models that can better capture the 'fat tails' and volatility clustering observed in financial markets. Similarly, Taleb in "The Black Swan" discusses the impact of highly improbable but high-impact events, which are often overlooked in standard Gaussian models but are essential in understanding financial markets (Taleb (2007)).
In this context, the multifractal nature of asset returns has garnered attention. The multifractal concept suggests that asset returns are not merely driven by a single scaling factor but can be better understood through multiple scaling behaviours across different time frames. (Calvet and Fisher, 2002) has offered a comprehensive analysis of its implications, emphasizing the versatility of multifractals in capturing the intricacies of financial returns. Furthermore, the non-normalities in these return distributions, especially their fat-tailed and skewed behaviours, have been discussed by (Rachev et al., 2005). They provide valuable insights into the challenges of simulating realistic financial data, emphasizing the complexities and deviations from the standard Gaussian distribution.
Financial econophysicists have also pointed to the need for complex models that capture the adaptive behaviours and 'fat-tail' phenomena in financial markets (Potters et al. (1998); Sornette and Malevergne (2001)). Such efforts show a clear trend towards developing financial models that break away from the Gaussian assumption and better capture the true distribution of asset price changes.
As we consider traditional quantitative models, we must also turn our attention to the domain of behavioural finance. Financial decisions are not always driven by rational calculations or predictions based on past data. In reality, investors, driven by emotions, psychological biases, and a plethora of other non-quantitative factors, often act in ways that defy traditional economic theories (Shiller, 2003). Various models have been proposed to better capture investor behaviour in financial markets. These models often incorporate elements of behavioural finance, emphasizing the role of psychological biases and irrationality in shaping investor behaviour.
The Prospect Theory by Kahneman and Tversky is a seminal work that illustrates this shift. It

suggests that people make decisions based on the potential value of losses and gains, rather than the final outcome. The theory highlights the asymmetry between gains and losses, where losses cause more emotional impact than equivalent gains—a phenomenon called “loss aversion” (Kahneman and Tversky (1979)).

Similarly, the Irrational Exuberance model, proposed by Shiller, suggests that investor behaviour can sometimes deviate significantly from rationality, leading to asset price bubbles and subsequent crashes (Shiller (2000)). Shiller won the Nobel Prize for his work demonstrating that stock market prices are more volatile than can be explained by the efficient market hypothesis.

In discussions about investor behaviour, it’s essential to understand the idea of ’normal agents’—individuals who don’t always act rationally in the economic sense. Instead of the theoretical rational agents that traditional financial theories often assume, these ’normal agents’ may exhibit tendencies like ’unconscious herding behaviour’, where they mimic the decisions of the majority. Such behaviours can challenge traditional assumptions about the predictability and normality of financial returns Prechter Jr (2001).

Moreover, researchers have found that introducing “agents who trade in a random way” into financial market models can reproduce some well-known phenomena observed in real market trading paths (Biondo et al. (2015)). This reinforces the concept that the presence of irrationality or ’normal’ behaviours in market participants can be an essential factor in understanding financial market dynamics.

In charting the evolution of behavioural finance, it’s crucial to appreciate that traditional models’ perceived shortcomings laid the groundwork for more nuanced models, incorporating psychological and behavioural factors for a fuller understanding of market phenomena.

To this end, De Bondt and Thaler (1985) initiated this shift by applying psychological theories to financial markets. Their work probed the stock market’s potential for overreaction, highlighting the psychological underpinnings of this phenomenon. This study’s findings not only challenged the traditional efficient markets hypothesis but also paved the way for future research in behavioural finance.

Building upon this foundational work, Shleifer and Summers (1990) introduced the noise trader approach. Their research underscores the influence of noise traders, who base decisions not on fundamental data but on extraneous ’noise’. This noise, according to their findings, generates unpredictable fluctuations that significantly affect market prices.

Moving into the late 90s, Barberis et al. (1998) developed a model of investor sentiment and its impact on asset pricing. Around the same time, Daniel et al. (1998) studied how investor overconfidence and biased self-attribution could cause underreactions and overreactions to information. Both studies advanced the understanding of investor behaviour’s complex role in financial market dynamics.

Further progress was made by Hong and Stein (1999), who presented a model unifying underreaction, momentum trading, and overreaction. Their unified framework underscores the intricate

interplay between these behavioural elements in asset pricing.
The Adaptive Market Hypothesis, proposed by Lo, merges the principles of behavioural finance with traditional models. According to this hypothesis, the degree of market efficiency is not static but varies over time, influenced by environmental conditions, the number of competitors, the magnitude of profit opportunities, etc. This model represents a more dynamic and realistic picture of financial markets, taking into account the adaptive behaviours of investors (Lo (2004)).
A new dimension was added to the field when Bakshi and Wu (2006) utilized options market data to study investor behaviour and risk perceptions during the NASDAQ bubble. Their study demonstrated the potential for unique insights to be gleaned from financial derivatives.
Around the same period, Hirshleifer et al. (2006) provided a model wherein irrational investors, through their trading activities, could influence underlying cash flows and secure abnormal profits. This work further accentuated the significant impact irrationality can exert in the financial markets.
Moving into the next decade, Manapat et al. (2013) offered an evolutionary perspective on trust, interpreting it as a manifestation of investor irrationality in economic interactions. Their research underscores trust's critical role in market transactions.
Adding further depth to our understanding of irrational behaviour, Dhesi and Ausloos (2016) introduced an innovative model incorporating irrational agent behaviour in response to time-dependent news. This research underscores the role of news in influencing stock prices through its impact on irrational agents. Bordalo et al. (2018) presented a model of credit cycles arising from diagnostic expectations, adding a new dimension to our understanding of financial markets. This study highlights how cognitive biases can shape broader economic phenomena, including credit cycles.
More recently, (Chen et al., 2020) observed a pervasive human tendency towards irrational stockpiling during crises such as the COVID-19 pandemic. This behaviour, also termed hoarding or panic buying, arises from limited, distorted, or exaggerated information that affects people's judgment. Such actions, though seemingly irrational, can be decoded using economics and psychology frameworks, with repercussions on the economy, society, and local communities. Parallelly, a study by (Ye et al., 2020), focusing on China's A-share market, unveiled that contrary to the efficient market hypothesis, investors often display herding behaviour. This tendency to mimic the majority, influenced by cognitive psychology, is especially pronounced in severe bear markets. Factors like loss aversion and extrapolative expectations, which anchor on historical stock price trends, were identified as key drivers behind such behaviour.
In summary, the field has seen a gradual but significant shift from traditional models towards a more behavioural approach, and as underscored by the multifractal nature of asset returns, towards understanding the intricate structures and dynamics of financial markets. This body of work underscores the importance of understanding investor behaviour to comprehend more

fully the complex dynamics of financial markets.

1.3 Motivation of the study

The "Irrational Fractional Brownian Motion" (IFBM) model, as presented by (Dhesi and Ausloos, 2016), is a pioneering approach that delves into the effects of irrational agent behaviour in response to time-dependent news. This innovative model posits the concept of a psychological Soliton, suggesting that forecast errors can initiate a ripple-like effect in financial markets akin to a Soliton. In fluid dynamics, a soliton is a solitary wave that retains its shape as it travels at a constant velocity (Drazin and Johnson, 1989). Translating this to financial markets, it suggests that the impact of a forecast error moves through the market, influencing prices and behaviours over time without quickly dissipating. Such a prolonged effect stands in contrast to the Efficient Market Hypothesis (EMH), which asserts that markets rapidly adjust to new information, leaving little room for sustained anomalies (Fama, 1970). The soliton-like behaviour, thus, presents a challenge to the EMH, implying that markets might not always be as efficient as traditionally believed. The introduction of a "news factor" offers a unique avenue to consider real-world stimuli that substantially influence financial markets, making it a departure from conventional models that often overlook such effects. The response to the news, especially during periods of financial turmoil or extreme market events, marks an area where irrationality might exert a significant influence, thus underscoring the relevance of the IFBM model.

The IFBM model presents a groundbreaking shift from traditional financial models, such as the Geometric Brownian Motion (GBM). By emphasizing investor irrationality and the leptokurtic nature of financial returns ((Mandelbrot, 1963), (Taleb, 2007)), the IFBM model offers a more nuanced portrayal of financial markets. These features are critical given the frequently observed 'fat tails' and pronounced peaks in financial return distributions, suggesting a higher likelihood of large price changes and more volatile market conditions. The IFBM model, characterized by its parameters k and c, has been proposed as a more general and comprehensive alternative to traditional models like the Geometric Brownian Motion (GBM) (Black and Scholes, 1973).
The GBM, a cornerstone in financial mathematics, is commonly used to depict the stochastic progression of stock prices, presuming that percentage changes in these prices are normally distributed. However, its assumption of constant volatility and normally distributed returns often results in shortcomings, especially during financial market anomalies (Hull, 2012).
One of the standout features of the IFBM model is its adeptness at accurately capturing the fat tails and pronounced leptokurtosis observed in financial returns distributions. Collectively, k and c modulate the distribution of returns to resemble a leptokurtic nature, which is characterized by pronounced peaks and fatter tails than a normal distribution. This leptokurtic behaviour is often

observed in real-world financial returns, especially during turbulent times, making the IFBM model a potentially more accurate reflection of market dynamics. Such a precise representation is pivotal for practitioners, especially in the realm of Value at Risk (VaR) management Dhesi et al. (2019). By offering a more accurate forecast of the fat-tailed frequency distribution for returns, the IFBM model equips financial professionals with a robust tool that can enhance risk management and foster more informed investment decisions.

Despite the evident advantages of the IFBM model, the stability of its to-be-estimated defining parameters, k and c, remains largely untested in the literature. Given this gap, the study proceeded to conduct extensive simulations of the IFBM model's parameters to rigorously assess its stability, especially in the paradigmatic case of S&P500, assessing its historical data. This intensive testing phase sought to verify the resilience and reliability of the IFBM model under real-world market conditions. Furthermore, earlier investigations into the model primarily focused on its performance in relation to the GBM, utilizing the Chi-Square criterion as the primary metric for comparison (Dhesi and Ausloos, 2016). While this approach offers valuable insights, its robustness under varying conditions became a point of contemplation. This encouraged a broader examination, paving the way for considering alternative methods that might offer more consistent and reliable results.

Recognizing the potential depth and multifaceted nature of the IFBM model, our research aims to expand this perspective. We propose to investigate the model's attributes through alternative analytical tools, notably Shannon entropy from information theory. Shannon's entropy quantifies the 'surprise' or 'information' content that a specific event yields. In the context of financial markets, this can be considered the amount of unexpected or new information an event brings to the market, such as an unexpected earnings announcement or a sudden change in interest rates (Shannon, 1948).

Our exploration of financial time series through the lens of information theory is rooted in a rich tapestry of research that spans several decades. The foundational concept of entropy, particularly Shannon entropy, was established as a measure of information content within datasets. Pincus (1991) was among the early pioneers who sought to quantify the complexity of time series data. Introducing the concept of approximate entropy, Pincus provided a methodology to assess system complexity, an idea that, while not strictly confined to financial datasets at the time, laid the groundwork for future applications in finance. Building upon these foundational principles, Darbellay and Wuertz (2000) ventured into the realm of finance, employing entropy measures to detect and quantify statistical dependencies in financial time series. Their findings illuminated the potential of entropy in capturing intricate market dynamics, setting the stage for further investigations into the relationship between entropy and financial markets. Further advancing this line of inquiry, Zunino et al. (2010) presented the complexity-entropy causality plane. This novel approach was designed to measure stock market efficiency by elucidating the intricate relationship between time series predictability (entropy) and complexity.

In a similar vein, Rosso et al. (2007) leveraged permutation entropy to differentiate between chaotic dynamics and stochastic noise within financial datasets. Their methodology underscored the adaptability and versatility of entropy-based measures in discerning the underlying processes driving financial data.

Given this historical context, our research seeks to harness the power of Shannon entropy in understanding the IFBM model. In our study, we will apply Shannon entropy to historical price data. We will calculate the entropy of price changes to elucidate the evolution of uncertainty and the information content of market events over time; we aim to dissect the intricacies of financial data and ascertain the accuracy of models like IFBM within the broader framework established by the aforementioned studies. Specifically, we plan to use the IFBM model to simulate market reactions to these events. Subsequently, the Shannon entropy of these simulated price changes will be calculated. Our ultimate goal is to compare the entropy of actual price changes with that of the simulated price changes under the IFBM model. This comparison allows us to assess how effectively the IFBM model captures the complexity of real-world market dynamics.

Furthermore, while the theoretical foundation of the IFBM model is robust, in the literature, there is a lack of a published best practice to use it. A formally designed algorithm would not only streamline the application of the IFBM model but also ensure its consistent and repeatable use in financial analyses. Addressing this gap, our research endeavours to develop a comprehensive algorithm that encapsulates the core principles of the IFBM model, making it more accessible and usable for both researchers and practitioners alike.

The IFBM model, characterized by two parameters (K and c), necessitated a procedure to accurately estimate these parameters across biennial periods, thereby obtaining a corresponding distribution of returns. An optimized Matlab algorithm was specifically developed for this purpose, ensuring precision in replication. The core of this procedure hinged on selecting optimal K and c values that, adhering to the Chi-Square criterion, minimized the divergence between the historical and simulated frequency distributions.

However, a concern emerged regarding the lack of convergence when escalating the number of simulations from 10 to $10,000$. This observation hinted at potential limitations within the Chi-square-based approach, thereby sparking curiosity towards exploring alternative methods. Shannon's entropy method emerged as a promising alternative, veering towards quantifying the variability inherent in given probability distributions. Moreover, to further scrutinize the robustness and applicability of the IFBM model, it was deemed imperative to test it under different market conditions. This line of thought naturally led to the application of the IFBM model to FTSE 100 data, with a keen eye on observing the stability of parameters K and c therein.

Moreover, the computational demands associated with the Chi-Square criterion, especially evident when handling large datasets or conducting multiple simulations, underscored the need for more computationally efficient metrics. Unlike the Chi-Square test, Shannon entropy directly employs observed probabilities, thus obviating the need for auxiliary computations, making it a valuable alternative in scenarios necessitating extensive simulations. Its utility in financial analyses has been underscored in studies by Zunino et al. (2010); Darbellay and Wuertz (2000), both highlighting its versatility and computational efficacy.

Building on these foundational analyses, the research then introduces Shannon's entropy as an advanced analytical tool (Shannon, 1948). Shannon's entropy quantifies the "informational content" in a dataset, thereby assessing its randomness or unpredictability. In the financial realm, understanding this unpredictability is paramount. High entropy values can indicate a vast array of possible outcomes, suggesting higher unpredictability and more information in the evaluated period. Conversely, low entropy values hint at more predictable outcomes, implying reduced risk (Cover, 1999) and a less informative time span. In times of financial crisis, markets often exhibit heightened volatility and unpredictability, which can amplify the effects of any model inaccuracies or misjudgments. Understanding and quantifying this unpredictability becomes even more essential during these periods as market participants seek to navigate turbulent waters and protect their investments. The IFBM, with its incorporation of a "news factor" and focus on investor irrationality, is particularly pertinent in this context. It offers a more dynamic and responsive framework that can potentially capture the abrupt and often extreme market reactions to unexpected news or events, which are characteristic of crises. Traditional models might fall short in such scenarios due to their inherent assumptions, which might not hold during turbulent times Mandelbrot and Hudson (2010); Shiller et al. (1981); Lo (2004). Thus, employing tools like Shannon's entropy in conjunction with the IFBM model can offer a more holistic and accurate assessment of the market, especially during periods of heightened uncertainty (Cont, 2001). The author posits that deviations from standard models during crises arise due to the complex interactions and dependencies among various market factors. The IFBM, by accounting for some of these complexities, positions itself as a potentially superior tool for understanding and navigating markets in crisis scenarios. By evaluating its predictions and behaviours through Shannon's entropy, we aim to test its mettle in truly challenging market conditions, offering insights that could reshape our approach to risk management and strategic decision-making during future crises.

Finally, the robustness and versatility of the developed algorithm for the IFBM model will be further evaluated by testing it on different market data, specifically the FTSE during the COVID-19 crisis period. This period, characterized by high volatility and potential behavioural biases Baker et al. (2020a); Al-Awadhi et al. (2020); Goodell (2020); Zaremba et al. (2020), presents an ideal testing ground to evaluate the model's ability to capture market dynamics under extreme conditions. This practical test will offer insights into the model's adaptability across diverse market contexts.

By integrating insights from the Chi-Square approach, stability assessments, Shannon's entropy, and different market data, this thesis aims to offer a comprehensive, multi-faceted perspective on the dynamics and associated risks of the IFBM model. In essence, our research seeks to journey through the relatively uncharted territories of the IFBM model, aiming to provide a comprehensive understanding of its dynamics, its practical applicability, and its value in capturing the multifaceted intricacies of financial markets.

1.4 Research objectives

Qualitative: How does the IFBM model, with its incorporation of investor irrationality and leptokurtosis, offer a more accurate representation of financial markets compared to the GBM model?

The Irrational Fractional Brownian Motion (IFBM) model, proposed by Dhesi (2016), diverges from traditional financial models like the Geometric Brownian Motion (GBM) by incorporating investor irrationality and leptokurtosis, an empirical reality often overlooked in classical models. The IFBM acknowledges that asset price fluctuations often exhibit a leptokurtic behaviour ((Mandelbrot, 1963), (Taleb, 2007)), characterized by 'fatter tails' and a higher peak than a normal distribution. This implies an increased likelihood of large price changes, resulting in more volatile and less predictable markets. Investigating the merits of this alternative model can provide a fresh perspective on financial market dynamics, shedding light on how irrational investor behaviours contribute to observed patterns in financial markets. In this research, we aim to revisit and elaborate on the IFBM model, highlighting its core principles and detailing its methodology. Our intention is to deepen understanding and provide a clear interpretation of the IFBM, making it more accessible to the broader financial academic community. We respect the foundational work on the model and seek to enhance, rather than critique, its presentation, especially considering its innovative approach.

Empirical: How does the IFBM model perform in terms of stability and robustness when tested against historical data?

Prior research has compared the GBM and IFBM models (Dhesi and Ausloos, 2016), focusing on how these models represent asset price dynamics. However, the stability and robustness of the IFBM model remain largely unexplored. The stability of the k and c parameters is a critical aspect of the IFBM model's applicability and reliability. This research question aims to rigorously test these parameters using both Chi-Square and Shannon entropy metrics. By employing these tests on historical S&P500 and FTSE100 data, we strive to validate the stability and robustness of the IFBM model, filling a crucial gap in the existing literature.

Methodological: Can we develop a repeatable and robust algorithm to facilitate the application of the IFBM model in financial market analyses?

While the IFBM model offers a more nuanced understanding of financial market dynamics, its application can be complex and not readily repeatable. To overcome this hurdle, this research aims to develop a robust algorithm for implementing the IFBM model in financial market analyses. This effort will ensure the model's practical applicability, making it a versatile tool for practitioners and researchers alike.

Practical: How does the developed algorithm for the IFBM model perform when applied to different market data, specifically the FTSE, during the COVID-19 crisis period?

The robustness and versatility of the developed algorithm for the IFBM model will be further evaluated by testing it on different market data, specifically the FTSE during the COVID-19 crisis period. This period, characterized by heightened volatility and pronounced behavioural biases (Baker et al., 2020a; Goodell, 2020), presents an ideal testing ground to evaluate the model's ability to capture market dynamics under extreme conditions. This practical test will offer insights into the model's adaptability across diverse market contexts.

Empirical: Using tools from information theory like Shannon entropy, what insights can we glean about the information content inherent in market dynamics as captured by the IFBM model?

Finally, to further our understanding of the IFBM model, we will employ tools from information theory, like Shannon entropy, to study the information content inherent in market dynamics as captured by the model. Shannon's entropy quantifies the amount of 'information' or 'surprise' a particular event brings ((Shannon, 1948)). By applying this measure to the IFBM model, we aim to explore the model's ability to capture the complex interactions and dependencies in financial markets, potentially offering a richer understanding of market behaviour and risks, especially during turbulent times (Cont, 2001).

1.5 Origin and Evolution of the IFBM Model

In financial research, accurately modelling financial returns is a persistent challenge. A significant advancement in this domain is the introduction of the IFBM model, designed to capture intricate patterns of financial returns. The origins of the IFBM can be traced back to the paper titled "Modified Brownian Motion Approach to Modelling Returns Distribution" (Dhesi et al., 2016). This pioneering research, which introduced the concept of the MBMM, serves as the foundation for the more evolved IFBM model.

1.5.1 Paper 1: Modified Brownian Motion Model

The authors of the MBMM sought to bridge the gap between traditional financial theories, rooted in the efficient market hypothesis (EMH) (Fama, 1970), and empirical findings. The

EMH posits that financial returns, driven by rational investor expectations, are normally distributed. However, empirical evidence frequently contradicts this, showing that financial returns often display leptokurtosis (Mandelbrot, 1963). This misalignment signals a need for a model that better represents the distribution of financial returns. The study's impetus came from an experimental paper on a semi-closed stock market (Dhesi et al., 2011), which incorporated the Geometric Brownian Motion (GBM) model with additional demand and supply factors.

In this section, we delve into the mathematical intricacies of the models under consideration, focusing on the Geometric Brownian Motion (GBM) and its innovative extension, the Modified Brownian Motion Model (MBMM).

1.5.1.1 Geometric Brownian Motion (GBM)

The cornerstone of classical quantitative finance can be expressed by the equation:

$$\ln\left(\frac{P_t}{P_{t-1}}\right) = \mu + \epsilon_t \tag{1.1}$$

where $\epsilon_t \sim \mathrm{NID}(0, \sigma^2)$. This equation, in continuous time, is represented as:

$$\frac{dP}{P} = \mu dt + \sigma dZ \tag{1.2}$$

In the classical GBM (Black and Scholes, 1973), the stochastic process is influenced solely by a random shock. This often falls short in capturing intricate patterns in financial returns, particularly during volatile market conditions (Dhesi and Ausloos (2016); Dhesi et al. (2016, 2019)).

1.5.1.2 Modified Brownian Motion Model (MBMM)

Building on the GBM, the research introduced an innovative extension to the model. This extension incorporates a weighting factor and a stochastic function, enhancing its ability to reflect real-world financial return distributions more accurately. The resulting formulation is:

$$P_{t+\delta} = P_t \exp\left(\mu\delta + \sigma\sqrt{\delta} Z_t + K f(Z_t)\mu\delta\right) \tag{1.3}$$

The function $f(Z)$ is defined as:

$$f(Z) = (2\exp(-c\frac{Z^2}{2}) - 1)\arctan(Z) \tag{1.4}$$

This function, in conjunction with the parameters K and c, enables the model to encapsulate deviations from rational behaviour commonly observed in financial markets. The behaviour of $f(Z)$ and the influence of parameters K and c are detailed in the subsequent sections.

Function $f(Z)$ and Parameters K and c

This section explores the behaviour of the function $f(Z)$ and the roles of parameters K and c in the MBMM.

The function $f(Z)$, in conjunction with the parameter K, enables the model to encapsulate deviations from rational behaviour commonly observed in financial markets.

The function $f(Z)$

- When Z_t is small: Such instances can be interpreted as minimal news or slight market fluctuations. Here, the collective investor sentiment might lean towards a perceived market stagnation, potentially driving them to sell their assets. Consequently, the MBMM's predicted price surge is subdued compared to the GBM's projection.
- When Z_t is substantially positive: Such scenarios can be seen as significant positive news or robust market trends. Here, the market sentiment might be overwhelmingly bullish, triggering a herding effect with investors scrambling to invest. The resulting demand overshadows supply, leading to a steeper price ascent in the MBMM compared to the GBM.

Parameter K

- K functions as a weighting factor, magnifying the impact of the $f(Z)$ function within the MBMM.
- At $K = 0$, the MBMM reverts to the traditional GBM, underscoring its role in accommodating deviations from this classic model.
- K's values modulate market reactions to varying news magnitudes and orientations. The paper emphasizes that the optimal K values differ across datasets, showcasing the model's adaptability to specific market nuances.
- K amplifies corrections on the GBM. In simpler terms, it accentuates price changes, making them more pronounced. A high positive value of K indicates anticipation of higher prices, pushing the distribution of returns to have "fatter tails". These fatter tails imply a higher probability of extreme price changes, which becomes particularly crucial during market crises when unexpected price fluctuations are rampant.

Parameter c

- c serves as a controlling variable in the $f(Z)$ function, pinpointing where tail "flattening" commences relative to standard deviations from the mean.

- The magnitude of c designates the proximity to the mean where this flattening phenomenon initiates.

- c equips the model with the flexibility to recalibrate based on empirical data, ensuring congruence between the model's distribution and real-world observations. Analogous to K, the ideal c value is dataset-dependent, accentuating the model's versatility.

- c introduces phases wherein the MBMM's influence on the GBM is subdued, working towards stabilizing the prices. This means that with c in play, the MBMM model tries to concentrate more prices around the average or mean. Such a concentration makes the distribution peak, indicating that most price changes are close to the mean, and fewer are farther away.

Collectively, K and c modulate the distribution of returns to resemble a leptokurtic nature, which is characterized by pronounced peaks and fatter tails than a normal distribution. This leptokurtic behaviour is often observed in real-world financial returns, especially during turbulent times, making the MBMM model a potentially more accurate reflection of market dynamics.

The research findings highlighted the superiority of the Modified Brownian Motion Model (MBMM) in several ways. Upon optimization, the MBMM showcased an exemplary fit to the historical histogram of returns, distinctly outperforming the traditional Geometric Brownian Motion (GBM). This supremacy was further solidified through chi-squared goodness of fit tests. One of the standout traits of the MBMM is its proficiency in capturing the leptokurtic tendencies intrinsic to financial return distributions, a feature that is often overlooked by many traditional models. The introduction of the parameter K and the function $f(Z)$ furnishes the MBMM with the requisite tools to encapsulate the often irrational behaviours observed in financial markets, marking a significant advancement in the field. Furthermore, the model's ability to adeptly represent leptokurtosis, a hallmark of empirical return data, emerged as a salient feature.
Incorporating the function $f(Z)$ into the MBMM provides a nuanced mechanism to modulate the influence of random shocks based on real-time market observations. This refinement offers a truer representation of financial return distributions, adeptly capturing both the central peak and the expansive tails. Such tails and peaks in the returns distribution are the manifestations of volatile asset price movements, which are frequently steered by factors like investor sentiment, herd behaviour, or unforeseen external shocks. Through $f(Z)$ and the parameters K and c, the model discerningly adjusts the influence of these random shocks. In essence, the MBMM effectively bridges the chasm between theoretical finance constructs and empirical market behaviours, offering invaluable insights into the idiosyncratic nature of financial markets (Dhesi et al., 2016).

1.5.2 Paper 2: Irrational Fractional Brownian Motion Model

Having delved deep into the intricacies of the MBMM model and its implications for understanding market behaviours in the previous paper, it's imperative now to turn our attention to a subsequent piece of research that further refines this perspective, introducing the Irrational Fractional Brownian Motion (IFBM) model and shedding light on the nuanced behaviours of market agents.

The primary motivation for Dhesi and Ausloose al.'s subsequent research paper appears to be the identification and quantification of the irrational behaviours exhibited by agents in financial markets. Building on their previous model, the authors sought to delve deeper into the implications of irrational agent behaviour, especially in response to time-dependent news on log returns, and its effect on financial market evolution. The authors were particularly intrigued by the observable kink-like effect reminiscent of soliton behaviour, suggesting the role of analysts' forecast errors in adjusting stock prices, thereby aiming to propose a measure of the irrational force in a market.

The initial paper by Dhesi et al. introduced a novel model that enhanced the Geometric Brownian Motion (GBM) by integrating an additional stochastic function, capturing the erratic and irrational movements observed in financial markets more accurately. This subsequent research builds upon the foundational model, but it focuses more intently on the psychological and behavioural aspects of agents in response to financial news or economic information. While the first paper laid the groundwork for the model and its parameters, this continuation seeks to apply, interpret, and further explore the implications of the model, especially in terms of the "stochastic psychological soliton".

1.5.2.1 Discussion and Details of the Model, Parameters, and Regions

The new paper reintroduces the model, now termed the Irrational Fractional Brownian Motion (IFBM) model, emphasizing its distinction from the traditional GBM. The change in nomenclature from MBMM to IFBM underscores the model's focus on capturing irrational behaviours in the market, emphasizing the fractional nature of the Brownian motion considered.
The IFBM model utilizes the feedback function $f(Z)$, which is a function of the random number Z that's assumed to represent financial news or economic information at a given time. The function $f(Z)$ is critical as it measures and models the irrational feedback behaviour of agents, especially when $K < 0$. The paper identifies six distinct regions based on the value of Z, which describe various market scenarios and the corresponding irrational behaviours.
These regions serve as a comprehensive framework to understand how agents might react to different market news, either amplifying or dampening their reactions based on the nature of

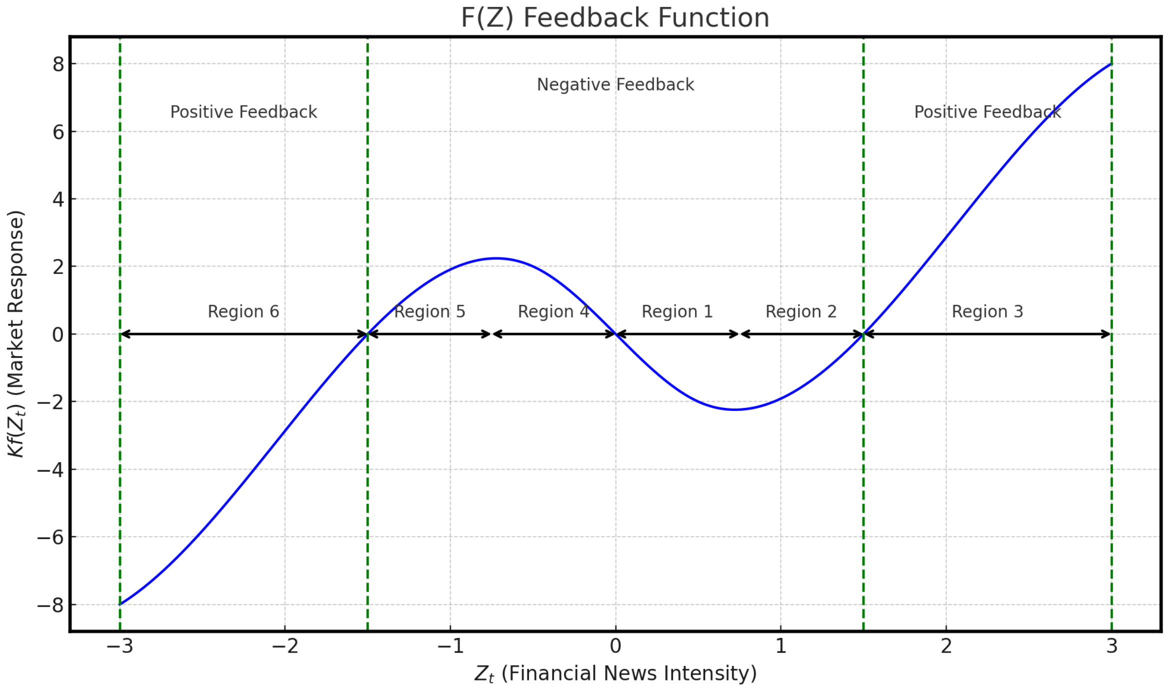


Figure 1.1: The plot showcases the $f(Z)$ feedback function, which represents the market's response to financial news intensity Z at time t. This function is integral to capturing deviations from the classic Geometric Brownian Motion (GBM) model by considering irrational behaviours in financial markets. The six regions on the graph highlight various market scenarios, ranging from extreme optimism to pronounced pessimism, further elucidating the market dynamics in response to varying intensities of financial news.

the news.

Understanding the different Regions on the Graph of the Feedback Function $f(Z)$:

The feedback function $f(Z)$, when graphed against the variable Z, provides insights into how the market might respond to varying degrees of financial news or economic information:

Positive Feedback:
This refers to a scenario where the feedback function amplifies or increases the original behaviour of the GBM. In the context of the graph, it's when $f(Z)$ has the same sign as Z, thus amplifying the effect of the news on the GBM.
For instance, in regions where the market is flooded with overwhelmingly positive news (high positive Z), the response (represented by $f(Z)$) is also positive and greater than the GBM's prediction. Similarly, when the news is extremely negative (very negative Z), the market response is also negative and more pronounced than what GBM would suggest. This exaggeration or amplification in response to the news is termed as positive feedback.

Negative Feedback:
In this context, the term "negative feedback" refers to the regions where the feedback function $f(Z)$ acts in a way that the modified Brownian motion model approaches the behaviour of the classic GBM. Here, the influence of $f(Z)$ is such that the market's response is more in line with the predictions of the classic GBM, even when considering the effects of financial news intensity Z.
Essentially, in these negative feedback regions, the deviations or modifications introduced by the $f(Z)$ function are such that the market's behaviour is closer to what would be expected in a purely rational market, as described by the GBM.

The feedback is about the behaviour of the market in response to the news, captured by the feedback function $f(Z)$, and how this behaviour deviates from the classic GBM. It's about how market agents collectively respond to the news and how this collective response deviates from what would be expected in a purely rational market as described by GBM.

Region 1:
Even with a hint of positive news, market agents react with caution. Instead of capitalizing on the good news, they might be selling assets, possibly out of impatience or scepticism. It's as if they doubt the credibility or longevity of this positive information.

Region 2:
As the news becomes more distinctly positive, the market's cautious or even negative response intensifies. The scepticism grows stronger, leading agents to possibly continue their sell-off, doubting the sustainability of the good news.

Region 3:
When the market is flooded with overwhelmingly positive news, there's a dramatic shift in behaviour. The market reacts with euphoria or over-enthusiasm, leading to buying sprees. It's as if the positivity is too strong to ignore or doubt.

Region 4:
On the flip side, with just a touch of negative news, the market surprisingly responds with a bit of optimism. Instead of panicking, agents might see this as a buying opportunity, hoping it's just a temporary dip.

Region 5:
As the negative news becomes more apparent, the market's optimism or hopeful behaviour

grows stronger. Agents could be acting contrarily, hoping for a turnaround or considering the more pronounced negative situation as a significant buying opportunity.

Region 6:
However, when faced with extremely negative news, the market finally succumbs to fear. There's a strong negative reaction, with the market likely in a state of panic, leading to rapid sell-offs.

1.5.2.2 The Soliton Aspect and behaviour

A significant highlight of this paper is the introduction of the concept of a "stochastic psychological soliton". The term "soliton" traditionally describes a self-reinforcing solitary wave or pulse that maintains its shape while it propagates at a constant velocity (Drazin and Johnson, 1989). In the context of this paper (Dhesi and Ausloos, 2016), the soliton represents the consistent, travelling kink in market behaviour due to the irrational actions of investors. It's a visual and mathematical manifestation of the market's irrational tendencies, providing both a theoretical and practical framework for understanding and measuring these behaviours.

The paper's findings underscore the relevance and efficacy of the IFBM model in capturing the leptokurtic distribution of asset returns and measuring irrational behaviour in financial markets. The research suggests that this model, with its incorporation of the "stochastic psychological soliton", offers a more foundational approach than the GBM for financial contexts. The IFBM not only captures the irrational tendencies in asset pricing but also provides a tangible measure for this irrational behaviour. The authors conclude by proposing several avenues for future research, emphasizing the broader applications and implications of the IFBM model.

1.5.3 Paper 3: IFBM Modelling and Forecasting the Kurtosis of Financial Markets

The ever-evolving landscape of financial markets continually challenges existing forecasting models. Building on foundational concepts introduced in the MBMM and IFBM papers, this latest research (Dhesi et al., 2019) embarks on a deeper exploration into the intricacies of financial returns distributions. It delves into the relationship between the kurtosis of the distribution and the parameters k and c through autoregressive (AR) processes. Kurtosis, as a statistical measure, offers insight into the distribution of observed data. A mesokurtic distribution, characterized by a kurtosis value of three, mirrors a normal distribution. In the realm of financial returns, distributions often exhibit leptokurtosis, characterized by a kurtosis exceeding three. Such distributions indicate fat tails, suggesting a higher propensity for outliers and extreme market events (DeCarlo, 1997). In the context of this research, the exploration of leptokurtic behaviour provides a lens to understand and predict significant market movements, forming a core aspect of the paper's methodology. This research illuminates not only the intricate relationship between kurtosis and the model parameters but also the dynamics between the ratio $\frac{k}{c}$ and the parameters themselves.

Methods

The research employs AR processes, a statistical tool utilized in time series analysis. In an AR process, the value of a variable at a given time is linearly dependent on its prior values, allowing it to capture inherent temporal dependencies (Wulff, 2017). Within the context of this research, AR processes were instrumental in dissecting the dynamics between k, c, and their subsequent influence on kurtosis, leveraging a rich dataset of financial returns from January 1, 1950, to December 31, 2015. This dataset was pivotal for validating the AR models and the ensuing findings.

Paper Findings

The research's findings are multifaceted:

- The relationship between kurtosis and the parameters k and c is captured by:

  $$\ln(\mathrm{Kurt}_t) = 1.66 + 0.078 \ln(k_t) - 0.081 \ln(c_t)$$

  This equation underscores the temporal dependencies of kurtosis on the parameters. By modelling the relationship between kurtosis and these parameters, the equation provides insights into the propensity for extreme market events. This understanding is vital for risk management; by predicting the likelihood of these "fat tail" events, risk managers can better strategize to mitigate potential losses.

- The dynamics of the $\frac{k}{c}$ ratio, when analysed through AR processes, unveiled significant outliers corresponding to major market events, such as the Cuban missile crisis in 1962. This suggests the potential of the $\frac{k}{c}$ ratio as an early warning system for market disruptions.

- The refined model has profound implications for risk management, especially in the calculation of Value at Risk (VaR). By accurately forecasting the leptokurtic nature of financial returns, the model enables practitioners to better estimate the potential downside of an investment portfolio over a specific timeframe (Jorion, 2007).

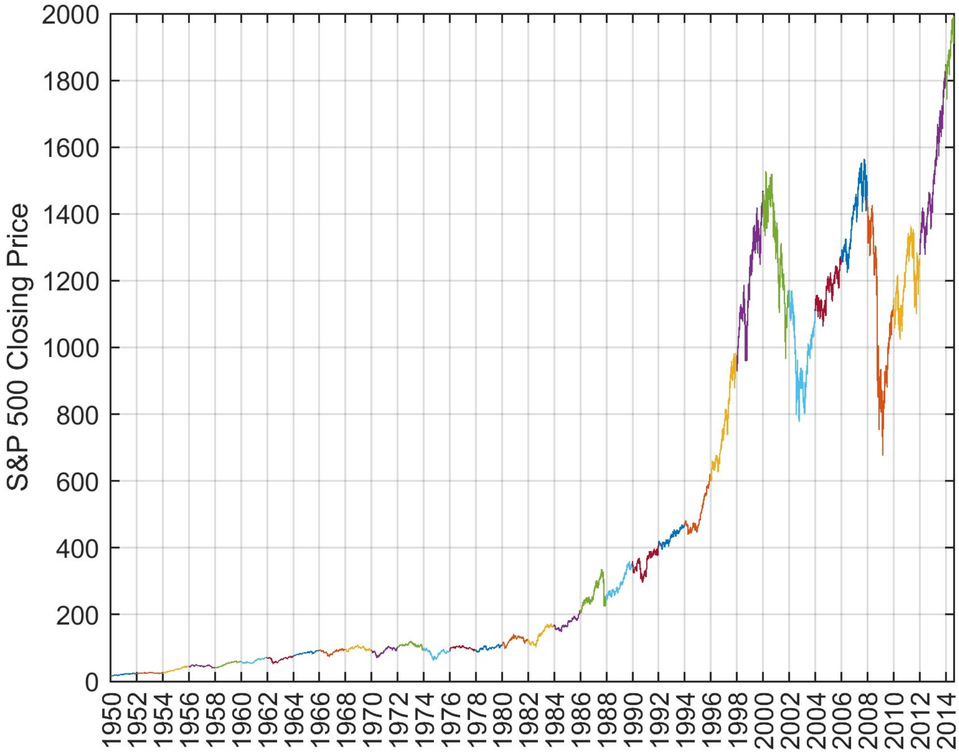
1.6 Financial data gathering

To explore the IFBM model's performance and adaptability, this study leverages two pivotal datasets that encapsulate the multifaceted dynamics of real-world financial markets. Two seminal datasets are employed: the S&P 500 and the FTSE 100. The former, an emblematic index of American equities, offers a panoramic view of the U.S. market, reflecting its historical highs and lows, its reactions to global events, and its inherent volatility. Given its extensive historical data, the S&P 500 serves as an empirical crucible, testing and validating theoretical frameworks in a setting that mirrors real-world financial dynamics (Hamilton, 1983; Malkiel, 2003; Carlson, 2007).
On the other hand, the FTSE 100, which represents the top 100 publicly traded companies in the UK, reflects the economic and financial conditions of the United Kingdom. This study finds the FTSE 100 data during the COVID-19 pandemic particularly relevant due to the period's significant market volatility and uncertainty. The dataset, characterized by its relevance and the distinct challenges brought by the pandemic, offers a suitable platform to examine the performance of models like the IFBM in capturing extreme market movements (Crafts and Mills, 2017; Emmerson et al., 2016). Note that all data presented in the table was downloaded from the official Bloomberg terminal.

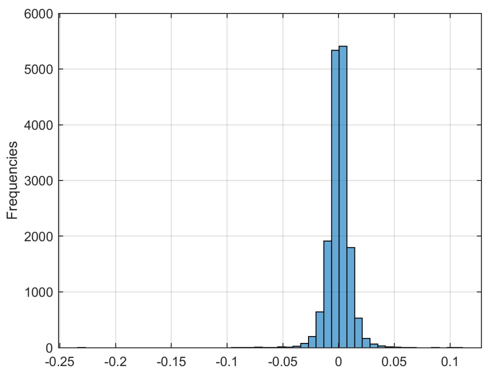
1.6.1 S&P 500 Dataset

Figure 1.2a showcases the historical progression of the S&P 500's daily closing prices over an extensive time frame, stretching from 03/01/1950 to 22/08/2014. The graphical representation employs alternating colours at two-year intervals, streamlining the data segmentation strategy of this study.
Prominent market events can be discerned by observing the data. The speculative bubble of the late 1960s and early 1970s, the "Nifty Fifty" era, possibly caused noticeable market behaviours. The oil crisis of 1973-1974, which led to a significant stock market downturn, might be reflected as well. The recession and high inflation of the early 1980s, Black Monday in 1987, the Dot-com bubble of the late 1990s, the aftermath of the 9/11 attacks in 2001, and the Global Financial Crisis between 2007-2009 are all pivotal moments in financial history that likely influenced the trends and volatility observed in the figure (Hamilton, 1983; Malkiel, 2003; Carlson, 2007).
Furthermore, the natural logarithm of the daily closing prices (Figure 1.2c is also depicted. The use of natural logarithm transformations is common in financial econometrics to stabilize the variance and to make patterns more discernible, especially when dealing with financial time series data (Tsay, 2005).
Figure 1.2b exhibits the empirical distribution of returns for the S&P 500. This visual representation offers insights into the frequency and magnitude of returns, allowing researchers and investors to glean a better understanding of the index's volatility and risk dynamics over the

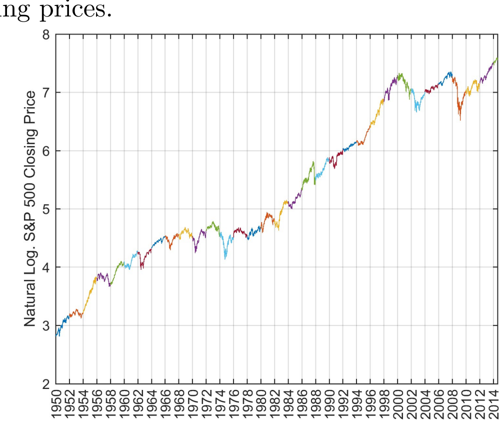
years.
The data source, having also been used by (Dhesi and Ausloos, 2016; Dhesi et al., 2016, 2019) publications, ensures continuity in research methodologies and allows for comparative analysis with previous findings.



(a) S&P 500 closing prices.



(b) Histogram of the S&P 500's daily returns.



(c) Natural Logarithm of the S&P 500 closing prices.

Figure 1.2: S&P 500 closing prices and returns historical data. The colours change every two years to match the dataset breakdown used in this study.

Table 1.1 presents a comprehensive statistical summary of the S&P 500 index over various biennial periods from 1950 to 2015, focusing on closing prices and returns. The skewness and kurtosis values provide insights into the symmetry and tail heaviness of the distribution, respectively. For instance, the skewness values fluctuate around zero, suggesting a mix of left and right-skewed distributions over the periods, while the kurtosis values, mostly below 3, signify a platykurtic distribution indicative of lighter tails compared to a normal distribution. The latter part of the table, focusing on returns, elucidates the performance of the S&P 500 index, which is paramount for investors and market analysts. It's evident that the returns have varied widely over the years, with some periods witnessing significant positive returns while others experienced notable negative returns. Particularly, during the financial crisis period of 2008-2009, the data exhibited heightened volatility with a noticeable decrease in the mean closing price. The Sharpe ratio also declined during this period, indicating a decreased risk-adjusted performance of the S&P 500 index.

Table 1.2 delineates the start and end dates for each biennial period from 1950 to 2014, serving as a chronological guide to the data analysis presented in the previous table. By establishing these temporal boundaries, the table facilitates a structured and organized examination of the S&P 500 market dynamics across distinct historical phases. Each row represents a two-year segment, beginning on the first trading day of the year and concluding on the last trading day of the subsequent year. This segmentation allows for a granular analysis of market behaviours and trends over time. It is worth noting that some pairs of years have fewer observations due to either a lack of data or the presence of a different number of non-trading days in each year.

type	N	max	min	μ	σ	μ/σ	skew.	kurt.
Clo. Pr. 1950-51	498	23.85	16.66	20.3596	2.1832	9.3256	-0.11001	1.6392
Clo. Pr. 1952-53	501	26.66	22.71	24.6096	0.83798	29.3679	0.41511	2.5594
Clo. Pr. 1954-55	504	46.41	24.8	35.1115	6.2706	5.5994	0.1303	1.7927
Clo. Pr. 1956-57	503	49.64	38.98	45.5292	2.4756	18.391	-0.68495	2.9134
Clo. Pr. 1958-59	505	60.71	40.33	51.8219	6.4558	8.0272	-0.43821	1.6615
Clo. Pr. 1960-61	502	72.64	52.2	61.0353	5.7908	10.54	0.33773	1.6582
Clo. Pr. 1962-63	503	75.02	52.32	66.0826	5.8747	11.2486	-0.5348	2.0266
Clo. Pr. 1964-65	505	92.63	75.43	84.7584	4.3467	19.4994	-0.091414	2.2609
Clo. Pr. 1966-67	503	97.59	73.2	88.562	5.7691	15.3512	-0.61923	2.3814
Clo. Pr. 1968-69	476	108.37	87.72	98.0596	4.7969	20.4423	-0.084633	2.2315
Clo. Pr. 1970-71	507	104.77	69.29	90.7196	8.9111	10.1805	-0.41008	2.0208
Clo. Pr. 1972-73	503	120.24	92.16	108.2849	5.2305	20.7028	-0.28473	3.7632
Clo. Pr. 1974-75	506	99.8	62.28	84.4811	9.196	9.1867	-0.63944	2.3399
Clo. Pr. 1976-77	505	107.83	90.71	100.1124	3.5173	28.4625	-0.47142	2.7866
Clo. Pr. 1978-79	505	111.27	86.9	99.5633	5.7248	17.3916	-0.30985	2.4772
Clo. Pr. 1980-81	506	140.52	98.22	123.374	9.895	12.4683	-0.62929	2.4807
Clo. Pr. 1982-83	506	172.65	102.42	140.0894	22.6124	6.1952	-0.12449	1.4184
Clo. Pr. 1984-85	505	212.02	147.82	173.6198	15.2975	11.3495	0.34846	2.1901
Clo. Pr. 1986-87	506	336.77	203.49	261.6942	33.6349	7.7804	0.53819	2.1311
Clo. Pr. 1988-89	505	359.8	242.63	294.4092	33.7079	8.7341	0.48966	1.7694
Clo. Pr. 1990-91	506	417.09	295.46	355.4092	27.5915	12.8811	-0.23065	1.8093
Clo. Pr. 1992-93	507	470.94	394.5	433.6455	20.2202	21.4461	0.12237	1.6311
Clo. Pr. 1994-95	504	621.69	438.92	501.0678	52.3581	9.57	0.80351	2.2038
Clo. Pr. 1996-97	507	983.79	598.48	771.7612	117.941	6.5436	0.3804	1.6843
Clo. Pr. 1998-99	504	1469.25	927.69	1206.4164	136.0961	8.8644	-0.091741	1.8023
Clo. Pr. 2000-01	500	1527.46	965.8	1311.6322	137.5035	9.5389	-0.31962	1.9271
Clo. Pr. 2002-03	504	1172.51	776.76	979.5812	98.9107	9.9037	0.1129	1.9231
Clo. Pr. 2004-05	504	1272.74	1063.23	1168.9394	49.0787	23.8176	0.05069	2.0311
Clo. Pr. 2006-07	502	1565.15	1223.69	1393.8228	96.3255	14.4699	-0.030881	1.6142
Clo. Pr. 2008-09	505	1447.16	676.53	1084.3135	208.5795	5.1986	0.1306	1.6585
Clo. Pr. 2010-11	504	1363.61	1022.58	1203.8022	87.2458	13.7978	0.019123	1.8647
Clo. Pr. 2012-13	502	1848.36	1277.06	1512.1033	153.3999	9.8573	0.41859	1.9194
Clo. Pr. 2014-15	162	1992.37	1741.89	1890.4495	59.6813	31.6758	-0.057852	2.2133
Ret. 1950-51	497	0.026546	-0.055316	0.00071512	0.0082253	0.086941	-1.1773	8.9124
Ret. 1952-53	500	0.016692	-0.03142	8.3122e-05	0.0055507	0.014975	-0.80731	6.2984
Ret. 1954-55	503	0.035116	-0.068476	0.0011936	0.0080511	0.14826	-1.2932	15.6885
Ret. 1956-57	502	0.043916	-0.029695	-0.0002422	0.0080041	-0.030259	0.2795	5.5254
Ret. 1958-59	504	0.021859	-0.0211	0.00078455	0.0057989	0.13529	-0.22207	3.8557
Ret. 1960-61	501	0.034213	-0.036778	0.0003544	0.0064786	0.054703	-0.21085	6.7828
Ret. 1962-63	502	0.045438	-0.069089	0.00011083	0.0083484	0.013276	-0.79358	16.248
Ret. 1964-65	504	0.020538	-0.018543	0.00040327	0.0038742	0.10409	-0.67827	6.7526
Ret. 1966-67	502	0.028037	-0.024912	9.0615e-05	0.0064832	0.013977	-0.14848	4.7018
Ret. 1968-69	475	0.024963	-0.0203	-9.0638e-05	0.0061406	-0.01476	0.10073	3.8759
Ret. 1970-71	506	0.049003	-0.028072	0.0001843	0.0081164	0.022707	0.48718	6.4175
Ret. 1972-73	502	0.030113	-0.030991	-8.2405e-05	0.0079202	-0.010404	-0.012977	4.4705
Ret. 1974-75	505	0.044934	-0.037403	-0.00015798	0.011945	-0.013226	0.30672	3.6962
Ret. 1976-77	504	0.018435	-0.018137	8.9621e-05	0.0064062	0.01399	0.059687	2.8908
Ret. 1978-79	504	0.038952	-0.030024	0.00027817	0.007374	0.037723	0.1295	4.9089
Ret. 1980-81	505	0.035727	-0.030529	0.00029178	0.0094515	0.030871	-0.11723	3.4958
Ret. 1982-83	505	0.046459	-0.040498	0.00058506	0.010038	0.058282	0.41818	4.6627
Ret. 1984-85	504	0.027223	-0.018374	0.00050213	0.0072512	0.069248	0.64442	4.0292
Ret. 1986-87	505	0.087089	-0.229	0.00032586	0.016397	0.019873	-5.5467	80.6733
Ret. 1988-89	504	0.033907	-0.070082	0.00064019	0.0095121	0.067303	-1.4471	13.4975
Ret. 1990-91	505	0.036642	-0.037272	0.00029319	0.0095265	0.030776	-0.045572	4.2283
Ret. 1992-93	506	0.019094	-0.024293	0.00022024	0.005769	0.038176	-0.043394	4.1181
Ret. 1994-95	503	0.021123	-0.022936	0.00055696	0.0056296	0.098934	-0.29578	4.4817
Ret. 1996-97	506	0.049887	-0.071127	0.00088309	0.009645	0.091559	-0.67927	10.2059
Ret. 1998-99	503	0.049646	-0.070438	0.00081517	0.012114	0.067294	-0.33588	5.9224
Ret. 2000-01	499	0.048884	-0.060045	-0.00047508	0.013789	-0.034453	0.0071553	4.4182
Ret. 2002-03	503	0.055744	-0.042423	-7.5003e-05	0.013867	-0.0054088	0.27792	4.1436
Ret. 2004-05	503	0.019544	-0.016862	0.00023615	0.006737	0.035053	-0.067225	2.8668
Ret. 2006-07	501	0.02879	-0.035343	0.00029157	0.0083881	0.03476	-0.43967	5.3257
Ret. 2008-09	504	0.10957	-0.094695	-0.00051718	0.021983	-0.023526	-0.11768	7.2996
Ret. 2010-11	503	0.046317	-0.068958	0.00020745	0.013134	0.015794	-0.44035	6.0113
Ret. 2012-13	501	0.025086	-0.025328	0.000738	0.0074983	0.098422	-0.1421	4.1087
Ret. 2014-15	161	0.015152	-0.023097	0.0005089	0.0066107	0.076981	-0.87811	4.5575

Table 1.1: Summary statistics of the S&P 500 variables here considered. N indicates the number of observations in the period.

years	Start	End
1950-51	03-Jan-1950	31-Dec-1951
1952-53	02-Jan-1952	31-Dec-1953
1954-55	04-Jan-1954	30-Dec-1955
1956-57	03-Jan-1956	31-Dec-1957
1958-59	02-Jan-1958	31-Dec-1959
1960-61	04-Jan-1960	29-Dec-1961
1962-63	02-Jan-1962	31-Dec-1963
1964-65	02-Jan-1964	31-Dec-1965
1966-67	03-Jan-1966	29-Dec-1967
1968-69	02-Jan-1968	31-Dec-1969
1970-71	02-Jan-1970	31-Dec-1971
1972-73	03-Jan-1972	31-Dec-1973
1974-75	02-Jan-1974	31-Dec-1975
1976-77	02-Jan-1976	30-Dec-1977
1978-79	03-Jan-1978	31-Dec-1979
1980-81	02-Jan-1980	31-Dec-1981
1982-83	04-Jan-1982	30-Dec-1983
1984-85	03-Jan-1984	31-Dec-1985
1986-87	02-Jan-1986	31-Dec-1987
1988-89	04-Jan-1988	29-Dec-1989
1990-91	02-Jan-1990	31-Dec-1991
1992-93	02-Jan-1992	31-Dec-1993
1994-95	03-Jan-1994	29-Dec-1995
1996-97	02-Jan-1996	31-Dec-1997
1998-99	02-Jan-1998	31-Dec-1999
2000-01	03-Jan-2000	31-Dec-2001
2002-03	02-Jan-2002	31-Dec-2003
2004-05	02-Jan-2004	30-Dec-2005
2006-07	03-Jan-2006	31-Dec-2007
2008-09	02-Jan-2008	31-Dec-2009
2010-11	04-Jan-2010	30-Dec-2011
2012-13	03-Jan-2012	31-Dec-2013
2014-15	02-Jan-2014	22-Aug-2014

Table 1.2: S&P 500 – Summary of the years divided into pairs

1.6.2 FTSE100 Dataset

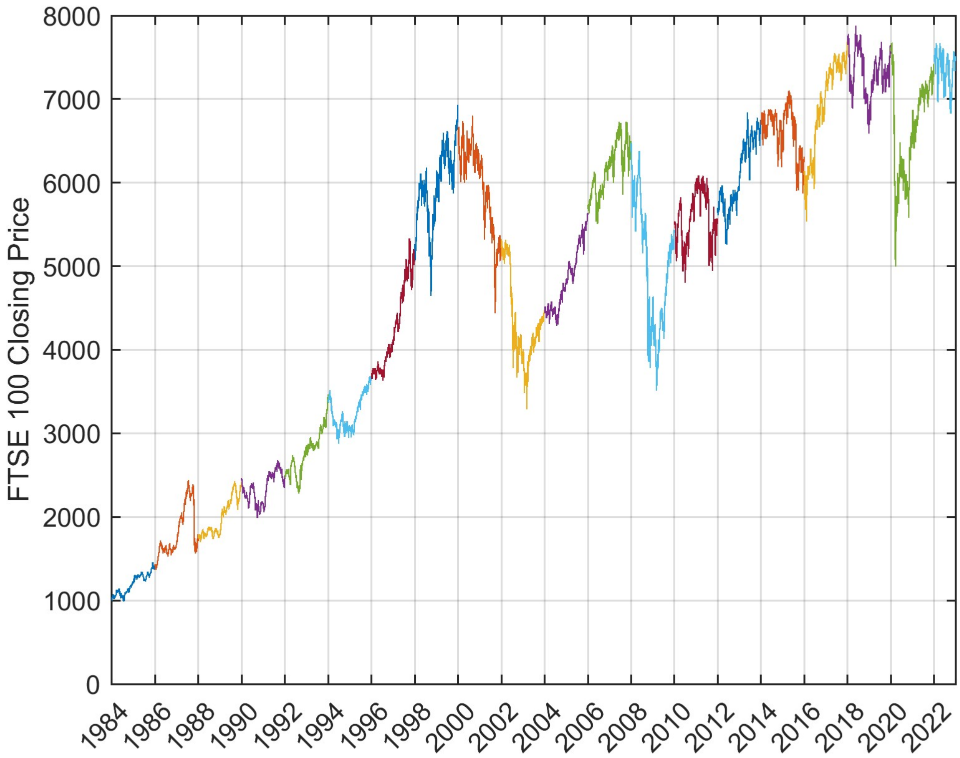
Figure 1.3a illustrates the historical trajectory of the FTSE 100's daily closing prices from 1984 to 2022. The data is visually segmented at two-year intervals, employing alternating colours in alignment with the data segmentation strategy of this study, facilitating a clearer understanding of the FTSE 100's performance across different time periods.

Several market events might have potentially influenced the trends observed in the figure. For instance, the Black Monday in 1987, the Dot-com bubble around the year 2000, the financial crisis of 2007-2008, and the more recent COVID-19 pandemic-induced market fluctuations in 2020-2021 are all significant events that could have impacted the market behaviours.

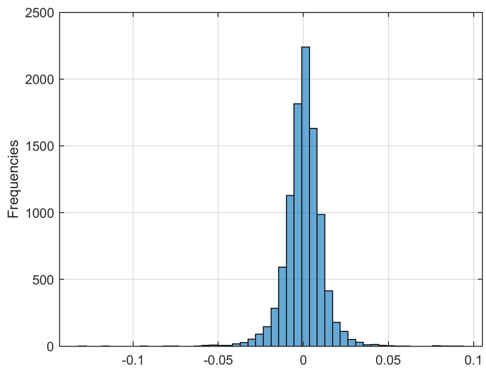
In Figure 1.3c, the natural logarithm of the daily closing prices is depicted. Similar to the S&P 500 analysis, the use of natural logarithm transformations is common in financial econometrics to stabilize the variance and to make patterns more discernible, especially when dealing with financial time series data.

Figure 1.3b portrays the empirical distribution of returns for the FTSE 100. This graphical representation provides insights into the frequency and magnitude of returns, aiding researchers and investors in better understanding the index's volatility and risk dynamics over the years. It further allows for a clearer analysis of how market events have potentially impacted the distribution of returns, enhancing our comprehension of the FTSE 100's performance.

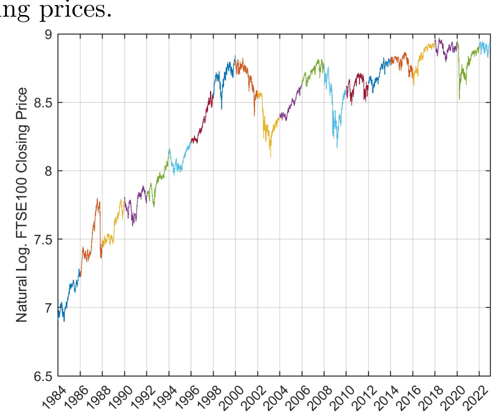
Table 1.3 presents a comprehensive statistical summary of the FTSE 100 index over various biennial periods from 1984 to 2022, focusing on closing prices and returns. The skewness and kurtosis values provide insights into the symmetry and tail heaviness of the distribution, respectively. For instance, the skewness values fluctuate around zero, suggesting a mix of left and right-skewed distributions over the periods, while the kurtosis values, mostly above 3, signify a leptokurtic distribution indicative of heavy tails and, therefore, potential outliers or extreme values. The latter part of the table, focusing on returns, elucidates the performance of the FTSE 100 index, which is paramount for investors and market analysts. It's evident that the returns have varied widely over the years, with some periods witnessing significant positive returns while others experienced notable negative returns. Notably, during the COVID-19 pandemic period of 2020-2021, the data exhibited heightened volatility with a noticeable decrease in the mean closing price. The Sharpe ratio also declined during this period, indicating a decreased risk-adjusted performance of the FTSE 100 index. The negative skewness suggests a tendency for the distribution to exhibit a longer left tail, potentially due to a few instances of sharp declines in the index value.



(a) FTSE-100 closing prices.



(b) Histogram of the FTSE-100's daily returns.



(c) Natural Logarithm of the FTSE-100 closing prices.

Figure 1.3: FTSE-100 closing prices and returns historical data. The colours change every two years to match the dataset breakdown used in this study.

type	N	max	min	μ	σ	μ/σ	skew.	kurt.
Clo. Pr. 1984-85	506	1455.5	986.9	1200.4318	119.7839	10.0216	-0.0038948	1.7743
Clo. Pr. 1986-87	520	2443.4	1370.1	1812.5663	293.0029	6.1862	0.70715	2.112
Clo. Pr. 1988-89	505	2426	1694.5	1989.0196	214.6952	9.2644	0.50415	1.8159
Clo. Pr. 1990-91	506	2679.6	1990.3	2345.6482	178.532	13.1385	-0.098636	1.822
Clo. Pr. 1992-93	507	3462	2281	2760.3598	248.4548	11.1101	0.2641	2.5375
Clo. Pr. 1994-95	504	3689.3	2876.6	3245.854	209.8769	15.4655	0.4026	1.8624
Clo. Pr. 1996-97	507	5330.8	3632.3	4261.558	504.414	8.4485	0.49474	1.8383
Clo. Pr. 1998-99	504	6930.2	4648.7	5955.505	438.4843	13.582	-0.43736	2.779
Clo. Pr. 2000-01	505	6798.1	4433.7	5964.5129	516.2268	11.5541	-0.6825	2.378
Clo. Pr. 2002-03	505	5323.8	3287	4318.7159	511.6661	8.4405	0.67255	2.3281
Clo. Pr. 2004-05	506	5638.3	4287	4839.7067	373.3482	12.963	0.39898	1.9015
Clo. Pr. 2006-07	505	6732.4	5506.8	6162.3663	300.5508	20.5036	-0.02554	1.998
Clo. Pr. 2008-09	507	6479.4	3512.09	4966.6566	739.9251	6.7124	0.063949	1.7912
Clo. Pr. 2010-11	504	6091.33	4805.75	5574.2038	310.7519	17.9378	-0.16006	1.9244
Clo. Pr. 2012-13	505	6840.27	5260.19	6108.3601	402.1224	15.1903	-0.006932	1.7051
Clo. Pr. 2014-15	506	7103.98	5874.06	6635.5853	259.3159	25.5888	-0.73572	2.9409
Clo. Pr. 2016-17	505	7687.77	5536.97	6926.2401	542.8201	12.7597	-0.67717	2.1223
Clo. Pr. 2018-19	506	7877.45	6584.68	7319.6854	258.1794	28.3512	-0.19547	2.513
Clo. Pr. 2020-21	507	7674.56	4993.89	6638.6532	584.6195	11.3555	-0.36376	2.1525
Clo. Pr. 2022-23	250	7672.4	6826.15	7357.4331	203.0286	36.2384	-0.73548	2.5719
Ret. 1984-85	505	0.033516	-0.028331	0.00068898	0.0086123	0.08	-0.18301	3.605
Ret. 1986-87	519	0.07597	-0.13029	0.00037117	0.013734	0.027027	-2.936	30.2422
Ret. 1988-89	504	0.025057	-0.032113	0.00064821	0.0080791	0.080233	-0.36709	3.6934
Ret. 1990-91	505	0.034976	-0.03692	4.7425e-05	0.0088724	0.0053453	0.097474	4.1607
Ret. 1992-93	506	0.054396	-0.041399	0.00062404	0.0082531	0.075614	0.742	9.0502
Ret. 1994-95	503	0.021778	-0.022617	0.00015738	0.0074362	0.021165	-0.20875	2.9656
Ret. 1996-97	506	0.031251	-0.031027	0.00065439	0.0079563	0.082248	-0.19592	4.4598
Ret. 1998-99	503	0.043451	-0.036595	0.00057352	0.012338	0.046484	-0.040964	3.6335
Ret. 2000-01	504	0.039839	-0.058853	-0.00048612	0.01279	-0.038008	-0.22075	3.8665
Ret. 2002-03	504	0.059038	-0.055888	-0.00030405	0.015017	-0.020247	-0.037844	5.033
Ret. 2004-05	505	0.019733	-0.023209	0.0004352	0.006027	0.072208	-0.32831	3.9846
Ret. 2006-07	504	0.034441	-0.04185	0.00025384	0.0095781	0.026502	-0.3983	5.0047
Ret. 2008-09	506	0.093843	-0.092656	-0.00033621	0.019713	-0.017056	-0.011834	7.2124
Ret. 2010-11	503	0.050322	-0.047792	2.5834e-05	0.012246	0.0021095	-0.15054	4.6125
Ret. 2012-13	504	0.030323	-0.030272	0.00033523	0.0081968	0.040898	-0.14892	3.9452
Ret. 2014-15	505	0.034976	-0.047795	-0.0001454	0.0092176	-0.015774	-0.29861	5.6192
Ret. 2016-17	504	0.03515	-0.035192	0.00046115	0.0083388	0.055302	0.040784	5.749
Ret. 2018-19	505	0.023255	-0.032839	-2.7548e-05	0.0077219	-0.0035675	-0.38024	4.5273
Ret. 2020-21	506	0.086668	-0.11512	-5.7955e-05	0.014335	-0.004043	-1.2006	15.8202
Ret. 2022-23	249	0.038391	-0.039555	-2.8682e-05	0.010348	-0.0027718	-0.34439	5.1673

Table 1.3: Summary statistics of the FTSE-100 variables here considered. N indicates the number of observations in the period.

1.7 IFBM procedure

In the domain of financial modelling, the widely held assumption concerning the normality of returns distributions is often challenged, especially among practitioners. Despite such objections, the Geometric Brownian Motion (GBM) continues to be a linchpin for simulating stock returns. As described in Section 1.5, the GBM is encapsulated by the equation:

$$P_{t+\delta t} = P_t \exp\left(\mu\delta t + \sigma Z_t\sqrt{\delta t}\right) \tag{1.5}$$

where P_t denotes the stock price at time t, μ and σ symbolize the mean and standard deviation of observed returns over a specified timeframe, δt is the designated time step, and $Z_t \sim N(0, 1)$.

The utilization of GBM for simulating prices and returns frequently misses the mark in accurately representing the observed distribution of returns. This shortfall is majorly attributed to the leptokurtic nature of returns distributions. Therefore, the inclusion of higher moments, such as skewness (third moment) and kurtosis (fourth moment) is deemed necessary. It is observed that stock returns often exhibit a left skewness, indicating a tendency towards negative returns during specific periods (Critchley and Jones, 2008; Groeneveld and Meeden, 1984).

The imperative for precise forecasting amplifies the importance of modelling empirical distributions to capture skewness and/or kurtosis Rubio and Steel (2015). Alterations of known symmetric distributions by introducing parameters to govern skewness are termed "skewing mechanisms" Ley and Paindaveine (2010), while the inclusion of a kurtosis parameter is referred to as "elongations" Fischer and Klein (2004), given their impact on the distribution's shoulders and tails. Significant collections of flexible tail distributions are outlined in Jones (2014) and Ley (2015), while other methodologies like semi-parametric models Ferreira et al. (2009) or fully non-parametric models provide alternative avenues for achieving model flexibility.

The limitations of GBM prompted the exploration of non-Gaussian processes and Fractal Brownian Motion to better portray the behaviour of financial assets (Nunzio Mantegna, 1991; Castellano et al., 2020; Di Matteo et al., 2005). A notably innovative approach is the Irrational Fractional Brownian Motion (IFBM) Dhesi and Ausloos (2016); Dhesi et al. (2016, 2019). The IFBM enhances GBM as shown in:

$$P_{t+\delta t} = P_t \exp\left(\mu\delta t + \sigma Z_t\sqrt{\delta t} - \mu K f(Z_t)\delta t\right) \tag{1.6}$$

where,

$$f(Z_t) = \left[2\exp\left(-c\frac{Z_t^2}{2}\right) - 1\right]\arctan\left(Z_t\right) \tag{1.7}$$

As a reminder from Section 1.5, the terms P_t, μ, σ, δt, and Z_t maintain their meanings from the GBM formulation, while K, and c are new parameters contributing to the "feedback function"

as outlined in Dhesi and Ausloos (2016), referred by Eq. (1.7).
This discourse elucidates the interaction between the component $-\mu K f(Z_t)\delta t$ and the conventional GBM, spotlighting the characteristics of parameters K and c.

A: When $-\mu K f(Z_t)\delta t < 0$, the GBM component $\mu\delta t + \sigma Z_t\sqrt{\delta t}$ is diminished in comparison to its standard form.

B: Conversely, when $-\mu K f(Z_t)\delta t > 0$, the GBM component $\mu\delta t + \sigma Z_t\sqrt{\delta t}$ is amplified relative to its standard form.

1.7.1 Parametric Influence on Feedback Function and GBM Behaviour

In the analysis that follows, we delve into the impact of varying parameters K and c on a mathematical mechanism designed to apply corrections to the Geometric Brownian Motion (GBM), a common model used to describe the stochastic evolution of financial prices. The corrections are aimed at moderating the price movements generated by the GBM, especially around certain critical points, improving the predictability and stability of the modelled financial system. The behaviour of this corrective mechanism under different settings of parameters K and c is explored through graphical illustrations in the referenced subfigures below.

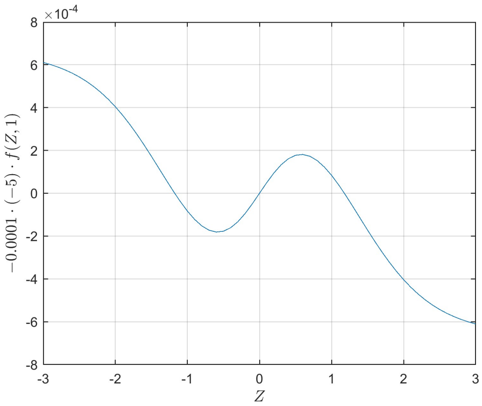
Given the parameters $\mu = 0.0001$; $K = -5$; $c = 1$; $\delta t = 1$, subfigure 1.4a elucidates that the intervention of the "feedback function" can manifest as positive, negative, or zero, contingent on the sampled value $Z_t \sim N(0, 1)$ which is taken within the range of -3 to 3. subfigure 1.4c delineates a series of simulated values, illustrating that $-\mu K f(Z_t)$ predominantly aligns around the mean, thereby exerting a mild influence on the GBM in terms of magnitude as indicated by Equation (1.6).

1.7.1.1 Parameter K influence

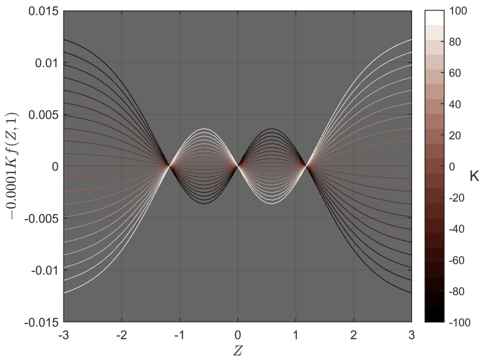
Upon the modification of the parameter K within the range of -100 to 100, while other variables remain constant, subfigure 1.4b illustrates a symmetrical transition in the corrections applied to the returns distribution, along with a variation in the magnitude of these corrections. It's observed that the mechanism's corrective action transitions from a mild influence (central area) to a pronounced one at the points where Z_s values satisfy $-\mu K f(Z_t)\delta t = 0$, which denote the roots of $-\mu K f(Z_t)\delta t$. These roots represent points of equilibrium where the mechanism's corrective action on the GBM is neutralized.

Subfigure 1.4d delves into the distribution variations with K, derived from a simulation of 1000 steps for each K in the range $K = [-100, \ldots, 0, \ldots, 100]$. This subfigure showcases a clustering of values around the mean and at the Z_s points where $-\mu K f(Z_t) = 0$, reflecting the mechanism's self-correcting nature against the GBM's generated values. Essentially, the mechanism

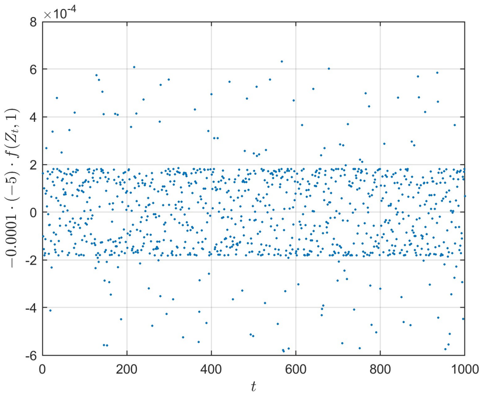
$-\mu K f(Z_t)\delta t$ operates to centralize the returns around the mean, thereby moderating the price movements relative to those generated by the GBM, especially within the "feedback function's" roots where $-\mu K f(Z_t)\delta t = 0$. This behaviour is indicative of a self-correcting system striving to maintain a balanced state in the face of the GBM's inherent volatility. Conversely, beyond these roots, the corrective restraint by the mechanism diminishes, and the price movements become more pronounced in absolute terms compared to those generated by the GBM, highlighting the mechanism's bounded influence.



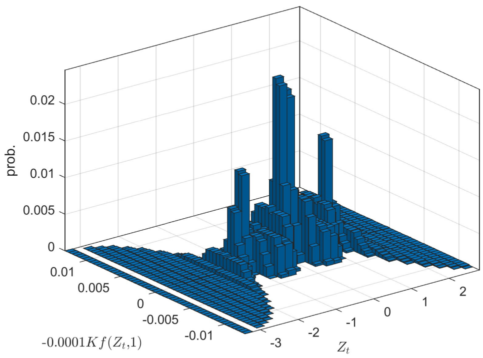
(a) The "feedback function".



(b) The "feedback function" when K varies.



(c) The "feedback function" simulated 1000 times.



(d) The histogram of the realization obtained with $-\mu K f(Z_t, c)$ when K varies.

Figure 1.4: Depiction of the corrective mechanism $-\mu K f(Z, c)$ and $-\mu K f(Z_t, c)$ under various conditions, with $\mu = 0.0001$; $K = -5$; $c = 1$. subfigure 1.4a displays $-\mu K f(Z, c)$ where Z is considered as a subset of the domain of the standard normal density function, restricted to [-3,3], representing a time-independent scenario. subfigure 1.4c showcases 1000 simulations of $-\mu K f(Z_t, c)$, illustrating the mechanism's behaviour over time. subfigure 1.4d extends the analysis by varying K in the range of [-100,100] for each simulation, highlighting the mechanism's response to different K values. Lastly, subfigure 1.4b presents $-\mu K f(Z, c)$ as K varies within the same range, providing a complementary view of the mechanism's sensitivity to K variations.

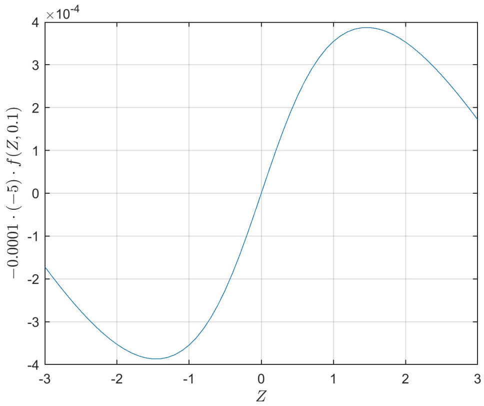
1.7.1.2 Parameter c influence

The parameter c plays a pivotal role in modulating the feedback function's corrective influence on the Geometric Brownian Motion (GBM). By varying c, we can observe how the feedback function adapts and how these adjustments manifest in the returns distribution generated by the IFBM. The referenced subfigures elucidate these dynamics under different settings of c.

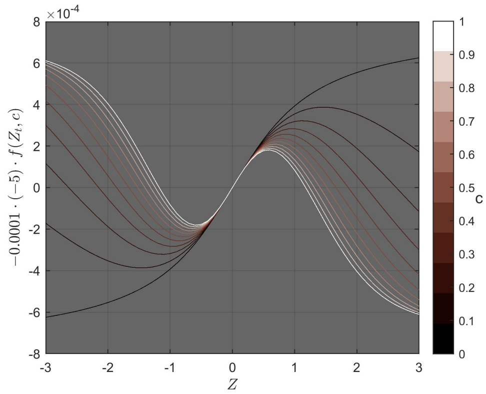
Figure 1.5 presents an analysis analogous to that of Figure 1.4, albeit focusing on the variation of the parameter c. subfigure 1.5a illustrates that with $c = 0.1$ (a relatively small value), the roots of the feedback function are distanced further apart, indicating that the tails of the IFBM-generated returns distribution commence farther from the mean.

Subfigure 1.4c displays a series of simulated values from the feedback function, showcasing a more homogeneous spread of values, with a noticeable concentration towards the margins of the plot compared to the scenario in subfigure 1.4b. This behaviour is accentuated as c approaches one, where the area of low feedback function values diminishes, albeit encompassing the majority of the values. Conversely, with a smaller c, the feedback function exhibits a broader area of influence. These dynamics are further elucidated in subfigure 1.5d, portraying the variations in the feedback function's corrective behaviour as c varies.
The parameter K serves as a pivotal controller in modulating the direction and magnitude of the interventions exerted by the "feedback function" on the mechanics of the Geometric Brownian Motion (GBM). It orchestrates how the feedback function interacts with the underlying GBM, thereby playing a critical role in the subsequent dynamics.

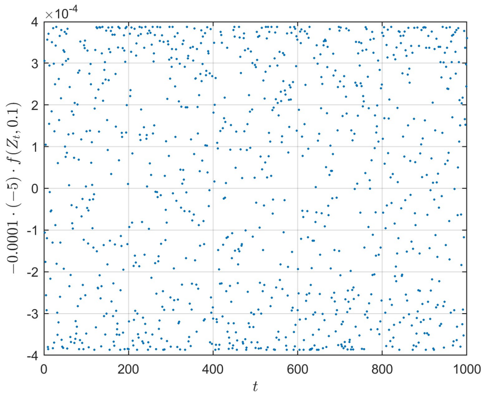
On the flip side, the parameter c specifically impacts the points where $-\mu K f(Z_t)\delta t = 0$, designated as the Z_s roots. These roots are crucial as they delineate the Z_t levels at which the feedback function begins to significantly amend the tails of the GBM generative process. As c enlarges, the resulting distribution of returns exhibits a more platykurtic nature, expanding the tails and moderating the peak, which is indicative of a broader dispersion of returns. This interplay between K and c offers a nuanced control over the feedback mechanism, enabling a tailored adjustment to the behaviour of the GBM, and, thus, the resultant returns distribution.



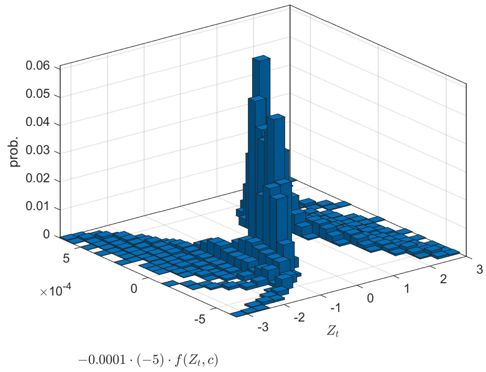
(a) The "feedback function".



(b) The "feedback function" when c varies.



(c) The "feedback function" simulated 1000 times.

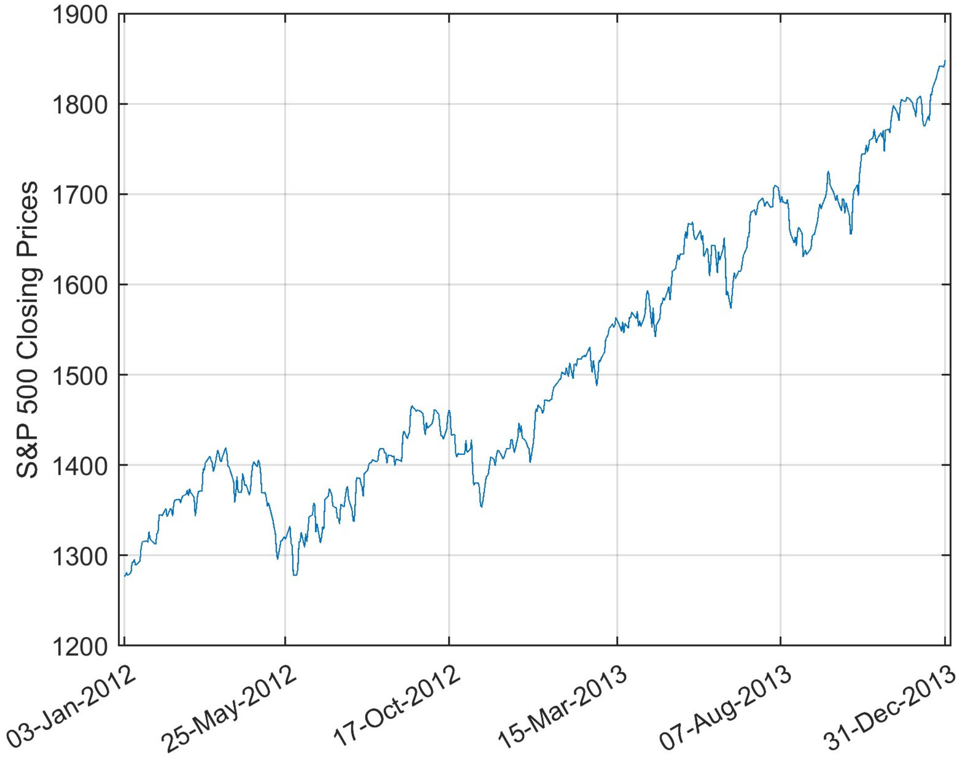


(d) The histogram of the realization obtained with $Kf(Z_t, c)$ when c varies.

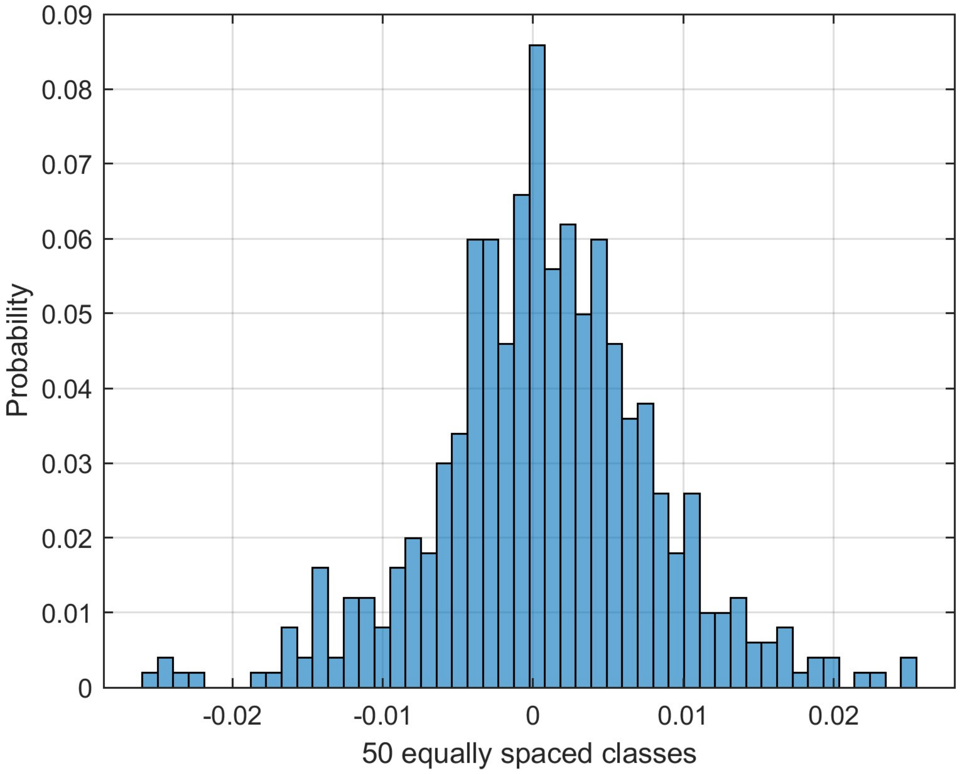
Figure 1.5: The set of plots collectively illustrates the behaviour of the feedback function $-\mu Kf(Z,c)$ and $-\mu Kf(Z_t,c)$ under different settings, with parameters set at $\mu = 0.0001, K = -5$, and $c = 0.1$. subfigure 1.5a displays $-\mu Kf(Z,c)$, where Z represents a sub-set of the domain of the standard normal density function, specifically within the range of [-3,3], thus depicting a time-independent scenario. On the other hand, subfigure 1.5c showcases 1000 simulations of $-\mu Kf(Z_t,c)$, elucidating the temporal evolution of the feedback function. In Subplot 1.5d, simulations are conducted for varying c values within the range [0,1], offering a multidimensional view of the feedback function's behaviour. Lastly, subfigure 1.5b encapsulates the variation of $-\mu Kf(Z,c)$ as c transitions within the range [0,1], portraying the sensitivity of the feedback function to c alterations.

1.7.2 Methodology - IFBM formalised procedure and stability testing

To ascertain the values of parameters K and c, pivotal in encapsulating market behavioural facets, a methodological approach is elucidated in Dhesi et al. (2016), leveraging the chi-square test. This section succinctly delineates the core tenets of this procedure, alongside the nuanced amendments undertaken—discussed in a comparative discourse—to faithfully replicate the findings presented in Dhesi et al. (2019). The discourse navigates through this analytical expedition using the empirical realization of the S&P 500 index during the years 2012-2013, as illustrated in Figure 1.6a, serving as a pragmatic backdrop for this exploratory endeavour.



(a) Daily S&P 500 closing price time series.



(b) Tabulation of the returns in 50 equally spaced classes

Figure 1.6: Data from daily observations of the S&P500 during the years 2012-2013

1 Initially, the mean (μ) and standard deviation (σ) of the returns during the specified period are documented. For the interval spanning 2012 to 2013, these are denoted as $\mu_{12-13} = 0.0007$ and $\sigma_{12-13} = 0.006$, respectively.

2 The core of this procedure rests on the utilization of the chi-square test statistics to facilitate a comparative analysis of empirical distributions. The employment of the chi-square test necessitates ensuring that the frequencies of the tabulated returns do not plummet below a predefined threshold, thereby circumventing unreliable outcomes from the chi-square test. Consequently, any realizations that reside outside the boundaries of $\mu \pm 3\sigma$ are expunged from the series of returns, as advocated in Dhesi et al. (2016). This prudent step aids in the eradication of outliers that could potentially engender complications.
Moreover, the edges of the bins, delineating the categories for data tabulation, are meticulously selected and amalgamated to ascertain that each bin encompasses at least 1.5% of the series' observations (for the case of 2012-13, this equates to 1.5% of 501). While Dhesi et al. (2016) recommends a threshold of 1%, this adjustment aims for a more conservative data representation.

3 Additionally, an optimal number of bins is determined to prevent the reduction of frequency to a critical level. In this instance, data is initially tabulated into 50 bins; subsequent calibration of bin widths and merging of low-frequency bins on the tails are undertaken to satisfy the minimum observation threshold. This procedure culminates in the data representation depicted in Figure 1.7 for the years 2012-13. The ensuing frequencies are designated as Customised Observed Frequencies, foc_h; $h = 1, \ldots, 50$ (with 50 representing the number of bins).
The theoretical framework delineated in Eq. (1.6) is employed, utilizing μ_{12-13}, σ_{12-13}, $\bar{K} = \{-100, \ldots, 0, \ldots, 100\}$ and $\bar{c} = \{0.01, 0.02, \ldots, 1\}$ to simulate the series of returns. Substituting Eq. 1.7 into Eq. (1.6), we obtain the expression:

$$P_{t+\delta t} = P_t e^{\mu_{12-13}\delta t + \sigma_{12-13} Z_t \sqrt{\delta t} - \mu \bar{K}\left[2\exp\left(-\bar{c}\frac{Z_t^2}{2}\right) - 1\right]\arctan(Z_t)\delta t} \tag{1.8}$$

4 A simulation of price series is conducted for each pair derived from the cross-combination of $\bar{K}$ and $\bar{c}$, with $\ell(\bar{K}) = 201$ and $\ell(\bar{c}) = 101$, resulting in a KC matrix housing $20,100$ pairs (K_m, c_m). This leads to a total of $20,100 \times 1000 = 20,100,000$ simulations. For each simulation, the returns are tabulated based on the bin edges defined in Step 2. The frequencies acquired from these tabulations are then averaged across the 1000 simulations for each (K_m, c_m) pair, denoted as Customized Expected Frequencies, $fec_{m,h}$; $m = \{1, \ldots, 20,100\}$; $h = \{1, \ldots, 50\}$ (with 50 representing the number of bins). Consequently, a $50 \times 20,100$ matrix F is generated, encapsulating the average frequencies per bin per pair.

5 In aligning with the chi-square test statistics formulation, the objective is to identify the pair (K_m, c_m) that minimizes the expression:

$$\min \left(\sum_{h=1}^{50} \frac{(foc_h - fec_{m,h})^2}{fec_{m,h}} \right) \forall\, m = \{1, \ldots, 20100\}$$

The optimal pair (K_m, c_m), denoted as $(\hat{K}_{\chi^2;m}, \hat{c}_{\chi^2;m})$, minimizes the discrepancy between the observed and simulated empirical distributions. For the period 2012-13, the optimal pair is $(\hat{K}_{\chi^2} = 2, \hat{c}_{\chi^2} = 0.38)$, as tabulated in Table 1.4 and consistent with the findings in Dhesi et al. (2019)'s Table 1[1]. Figure 1.8 exhibits the superior fit achieved.

It is pertinent to note that the aforementioned procedure was implemented using data from the S&P 500 index. The procedure, as outlined, yielded the estimations of parameters $\hat{K}$ and $\hat{c}$ for the years 2012-2013, which are documented in Table 1.5. The table delineates the outcomes of the procedure under varying simulation counts—specifically at 10, 100, 1000, and $10,000$ simulations. The derived parameters, annotated with a subscript reflecting the simulation count, are tabulated alongside.

Furthermore, a visual representation of these outcomes is provided in Figure 1.9, which underscores the sensitivity of the parameter estimations to the simulation count. Particularly for parameter $\hat{c}$, as depicted in subfigure 1.9b, the influence of simulation count is prominently highlighted, thus laying a substantial groundwork for the subsequent phases of this research.

[1] Minor variations may arise due to the number of simulations.

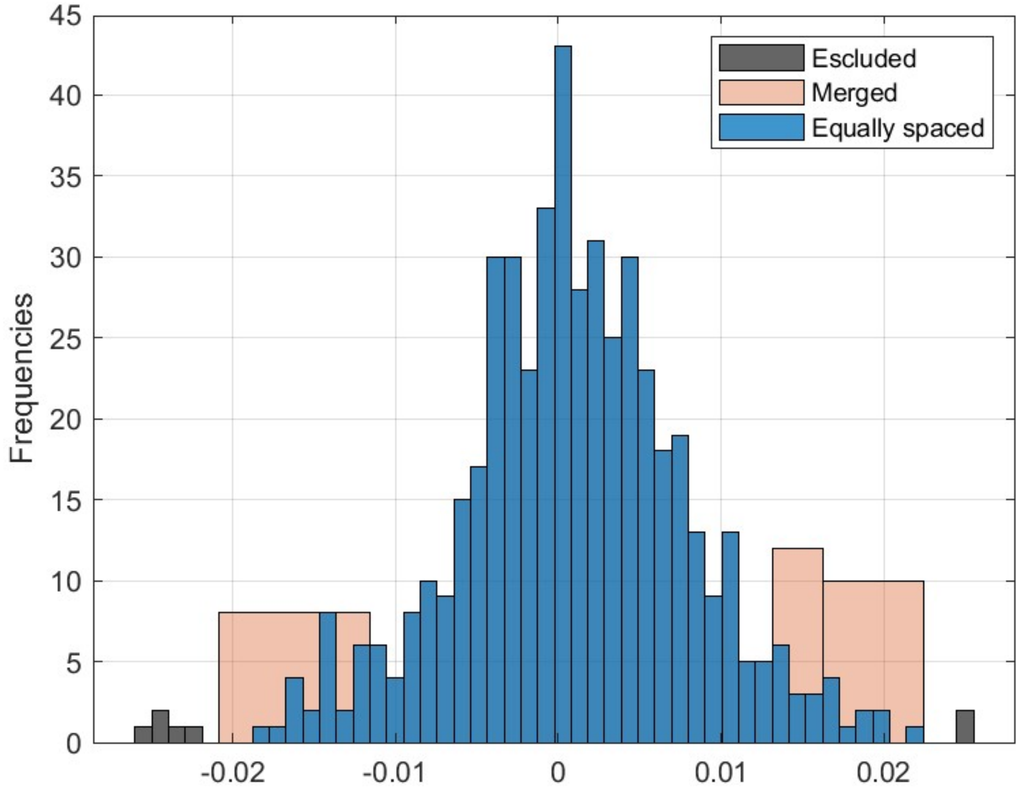


Figure 1.7: S&P 500 - The histogram showcases three distinct categories of data manipulation to prepare for the chi-square test. The black bars represent the data points that have been excised to eliminate outliers, ensuring a more reliable chi-square analysis. The blue bars depict the frequency count post-outlier removal, while the orange bars illustrate the strategic merging of bins, each now encompassing a minimum of 1.5% of the total observations, thus adhering to the requisite threshold for a robust statistical analysis.

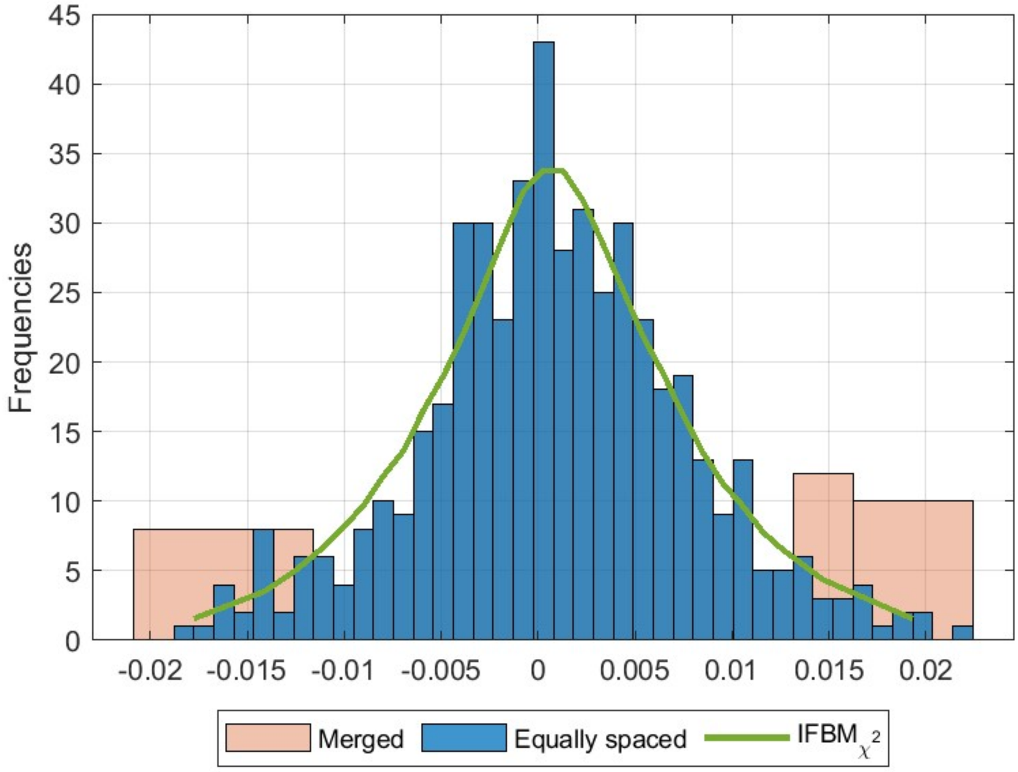


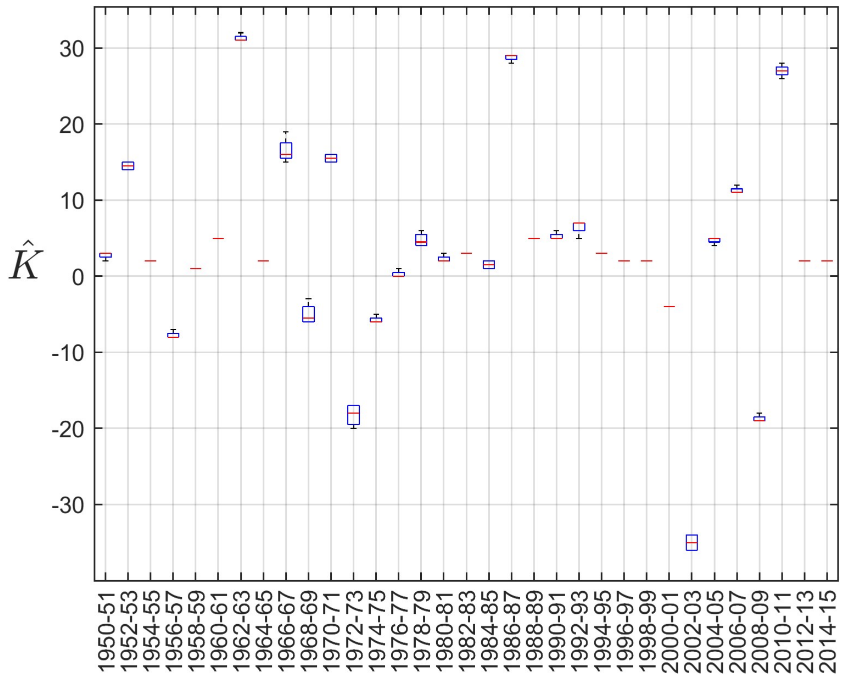
Figure 1.8: S&P 500 - Frequency Levels Represented by $IFBM_I$ Using Estimated Parameters ($\hat{K} = 2, \hat{c} = 0.38$) for the Years 2012-2013.

years	μ	$\hat{K}$	$\hat{c}$
1950-51	0.00071512	3	0.06
1952-53	8.3122e-05	15	0.12
1954-55	0.0011936	2	0.06
1956-57	-0.0002422	-8	0.29
1958-59	0.00078455	1	0.37
1960-61	0.0003544	5	0.24
1962-63	0.00011083	31	0.08
1964-65	0.00040327	2	0.22
1966-67	9.0615e-05	16	0.31
1968-69	-9.0638e-05	-6	0.32
1970-71	0.0001843	15	0.42
1972-73	-8.2405e-05	-20	0.34
1974-75	-0.00015798	-6	0.2
1976-77	8.9621e-05	0	0
1978-79	0.00027817	4	0.33
1980-81	0.00029178	2	0.23
1982-83	0.00058506	3	0.29
1984-85	0.00050213	1	0.21
1986-87	0.00032586	29	0.06
1988-89	0.00064019	5	0.26
1990-91	0.00029319	5	0.2
1992-93	0.00022024	7	0.67
1994-95	0.00055696	3	0.65
1996-97	0.00088309	2	0.09
1998-99	0.00081517	2	0.14
2000-01	-0.00047508	-4	0.29
2002-03	-7.5003e-05	-34	0.62
2004-05	0.00023615	5	1
2006-07	0.00029157	11	0.36
2008-09	-0.00051718	-19	0.32
2010-11	0.00020745	27	0.45
2012-13	0.000738	2	0.38
2014-15	0.0005089	2	0.11

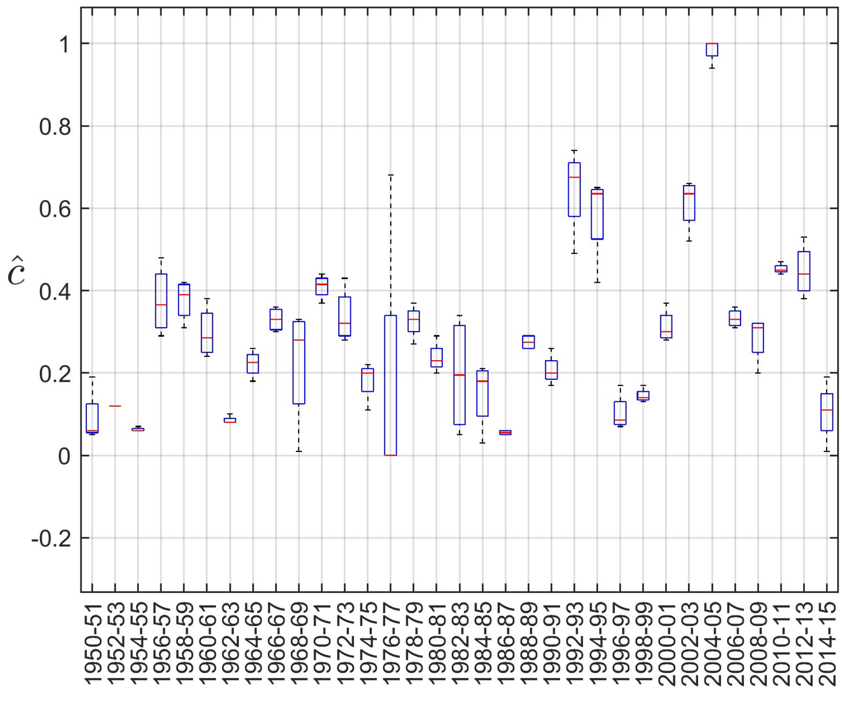
Table 1.4: S&P 500 - Tabulated Results from the Exercise Conducted in Dhesi et al. (2019), Employing 1000 Simulations

years	μ	$\hat{K}_{10}$	$\hat{K}_{100}$	$\hat{K}_{1000}$	$\hat{K}_{10000}$	$\hat{c}_{10}$	$\hat{c}_{100}$	$\hat{c}_{1000}$	$\hat{c}_{10000}$
1950-51	0.00071512	3	2	3	3	0.19	0.05	0.06	0.06
1952-53	8.3122e-05	14	14	15	15	0.12	0.12	0.12	0.12
1954-55	0.0011936	2	2	2	2	0.07	0.06	0.06	0.06
1956-57	-0.0002422	-7	-8	-8	-8	0.48	0.4	0.29	0.33
1958-59	0.00078455	1	1	1	1	0.31	0.42	0.37	0.41
1960-61	0.0003544	5	5	5	5	0.38	0.31	0.24	0.26
1962-63	0.00011083	32	31	31	31	0.1	0.08	0.08	0.08
1964-65	0.00040327	2	2	2	2	0.26	0.23	0.22	0.18
1966-67	9.0615e-05	19	16	16	15	0.35	0.3	0.31	0.36
1968-69	-9.0638e-05	-3	-5	-6	-6	0.01	0.33	0.32	0.24
1970-71	0.0001843	15	16	15	16	0.37	0.41	0.42	0.44
1972-73	-8.2405e-05	-17	-17	-20	-19	0.43	0.28	0.34	0.3
1974-75	-0.00015798	-5	-6	-6	-6	0.2	0.22	0.2	0.11
1976-77	8.9621e-05	0	1	0	0	0	0.68	0	0
1978-79	0.00027817	6	5	4	4	0.27	0.33	0.33	0.37
1980-81	0.00029178	3	2	2	2	0.29	0.2	0.23	0.23
1982-83	0.00058506	3	3	3	3	0.1	0.05	0.29	0.34
1984-85	0.00050213	2	2	1	1	0.03	0.2	0.21	0.16
1986-87	0.00032586	28	29	29	29	0.05	0.05	0.06	0.06
1988-89	0.00064019	5	5	5	5	0.29	0.29	0.26	0.26
1990-91	0.00029319	6	5	5	5	0.17	0.26	0.2	0.2
1992-93	0.00022024	5	7	7	7	0.49	0.74	0.67	0.68
1994-95	0.00055696	3	3	3	3	0.42	0.63	0.65	0.64
1996-97	0.00088309	2	2	2	2	0.17	0.08	0.09	0.07
1998-99	0.00081517	2	2	2	2	0.17	0.13	0.14	0.14
2000-01	-0.00047508	-4	-4	-4	-4	0.37	0.31	0.29	0.28
2002-03	-7.5003e-05	-36	-36	-34	-34	0.66	0.65	0.62	0.52
2004-05	0.00023615	4	5	5	5	0.94	1	1	1
2006-07	0.00029157	12	11	11	11	0.31	0.32	0.36	0.34
2008-09	-0.00051718	-18	-19	-19	-19	0.2	0.32	0.32	0.3
2010-11	0.00020745	26	27	27	28	0.47	0.45	0.45	0.44
2012-13	0.000738	2	2	2	2	0.53	0.46	0.38	0.42
2014-15	0.0005089	2	2	2	2	0.01	0.19	0.11	0.11

Table 1.5: S&P 500 - Tabulated Results of K and c Estimations Across Varying Simulation Counts Utilizing the Chi-Square Based Procedure



(a) S&P 500 - Box Plot Representation of Estimated K Values Across Different Year Pairs



(b) S&P 500 - Box Plot Representation of Estimated c Values Across Different Year Pairs

Figure 1.9: S&P 500 - These box plots depict the variability in the estimations of parameters K and c from Eqs (1.6) and (1.7), across different numbers of simulations employed in the estimation procedure (refer to Table 1.5 for a tabular presentation of these results). Each box plot illustrates the interquartile range (IQR) of the estimations: the bottom and top edges of the box represent the 25th and 75th percentiles, respectively, while the red markers indicate the medians. The whiskers extend to the most extreme data points, not classified as outliers. A key observation from these plots is the absence of convergence towards specific values as the number of simulations escalates from 10 to 10,000; however, the estimations for $\hat{K}$ exhibit a relatively more stable pattern compared to those for $\hat{c}$.

1.7.3 Pivoting Towards Enhanced Analysis: Shannon's Entropy and FTSE 100 Examination

Having successfully replicated the procedure and executed stability tests, the journey of exploration into the dynamics of the Extended IFBM does not halt. The endeavour now veers towards a more comprehensive understanding, prompted by the desire to delve deeper into the model's behaviour under varying conditions. Two significant avenues beckon: the utilization of Shannon's Entropy for nuanced parameter estimation and the application of the IFBM to FTSE 100 data, a market that has also weathered crises such as COVID-19. These pursuits are not merely academic exercises but are tethered to the real-world exigencies of understanding market dynamics amidst different scenarios.

The methodology (described in 1.7.2) initially employed entailed representing the returns of historical, GBM, and IFBM-based prices in distinct histograms, with each bin's frequency meticulously recorded. The Chi-square criterion served as a pivotal tool for comparing the observed frequencies with the expected ones derived from simulations. A prudent step within this procedure entailed the exclusion of return realizations beyond the range of $\mu \pm 3\sigma$, owing to the Chi-square formula's sensitivity to outliers. To enhance the interpretability of the Chi-square results, the bins across the histograms were tailored to ensure each encompassed a frequency of at least 1.5% of the total data set values.

The IFBM model, harbouring two parameters (K and c) requiring estimation, propelled us to simulate these parameters across each biennial period, thereby obtaining a corresponding distribution of returns. Subsequent steps involved computing the average bin heights across the histograms to ascertain the respective frequencies for each two-year time frame. The crux of this procedure hinged on selecting optimal K and c values that, adhering to the Chi-square criterion, minimized the divergence between the historical and simulated frequency distributions.

However, a notable concern emerged regarding the lack of convergence when escalating the number of simulations from 10 to 10,000. This observation hinted at potential limitations within the Chi-square-based approach, thereby sparking curiosity towards exploring alternative methods. Shannon's entropy method emerged as a promising alternative, veering towards quantifying the variability inherent in given probability distributions. Moreover, to further scrutinize the robustness and applicability of the IFBM model, it was deemed imperative to test it under different market conditions. This line of thought naturally led to the application of the IFBM model to FTSE 100 data, with a keen eye on observing the stability of parameters K and c therein. The ensuing sections delve into the intricacies of this method and the findings from the FTSE 100 application, aiming to potentially overcome the convergence challenge and further refine the estimation of parameters K and c.

1.7.4 Methodology - IFBM under Different Market Dynamics: The FTSE 100 Scenario

Transitioning to a different market scenario, the same procedural framework (detailed in 1.7.2) was applied to data from the FTSE 100 index, aiming to comprehend the behaviour of the IFBM under divergent market conditions. This endeavour also included the application of a stability test to the derived estimations. The contrasting market conditions and the subsequent analysis further enrich the comprehension of the IFBM's behaviour, providing a more robust foundation for the overarching investigation.

The results of this application are tabulated and presented in Table 1.6 and Table 1.7. Table 1.6 showcases the outcomes from the exercise conducted as per Dhesi et al. (2019), employing 1000 simulations, while Table 1.7 enumerates the estimations of parameters K and c across varying simulation counts, utilizing the Chi-Square based procedure. The distinct estimations of K and c across different year pairs are further visually represented through box plots in Figure 1.10. subfigures 1.9a and 1.10b distinctly illustrate the variability in the estimations of parameters K and c respectively.

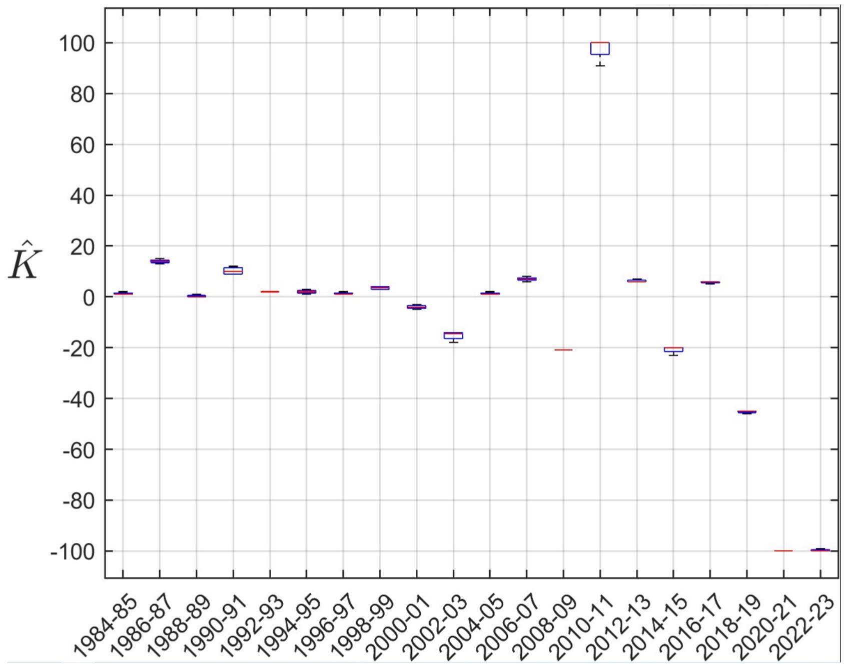
These representations collectively provide a comprehensive insight into the behaviour of the IFBM under the diverse market conditions presented by the FTSE 100 index, thereby contributing to a more nuanced understanding of the model's performance and stability across different market scenarios.
A detailed discussion on the implications and analysis of these results can be found in Section 1.9.

years	μ	$\hat{K}$	$\hat{c}$
1984-85	0.00068898	1	0.72
1986-87	0.00037117	14	0.01
1988-89	0.00064821	0	NaN
1990-91	4.7425e-05	12	0.01
1992-93	0.00062404	2	0.01
1994-95	0.00015738	2	0.95
1996-97	0.00065439	2	0.15
1998-99	0.00057352	3	0.75
2000-01	-0.00048612	-4	0.75
2002-03	-0.00030405	-14	0.28
2004-05	0.0004352	1	0.05
2006-07	0.00025384	7	0.2
2008-09	-0.00033621	-21	0.4
2010-11	2.5834e-05	100	0.46
2012-13	0.00033523	6	0.66
2014-15	-0.0001454	-20	0.39
2016-17	0.00046115	6	0.15
2018-19	-2.7548e-05	-45	0.25
2020-21	-5.7955e-05	-100	0.37
2022-23	-2.8682e-05	-100	0.37

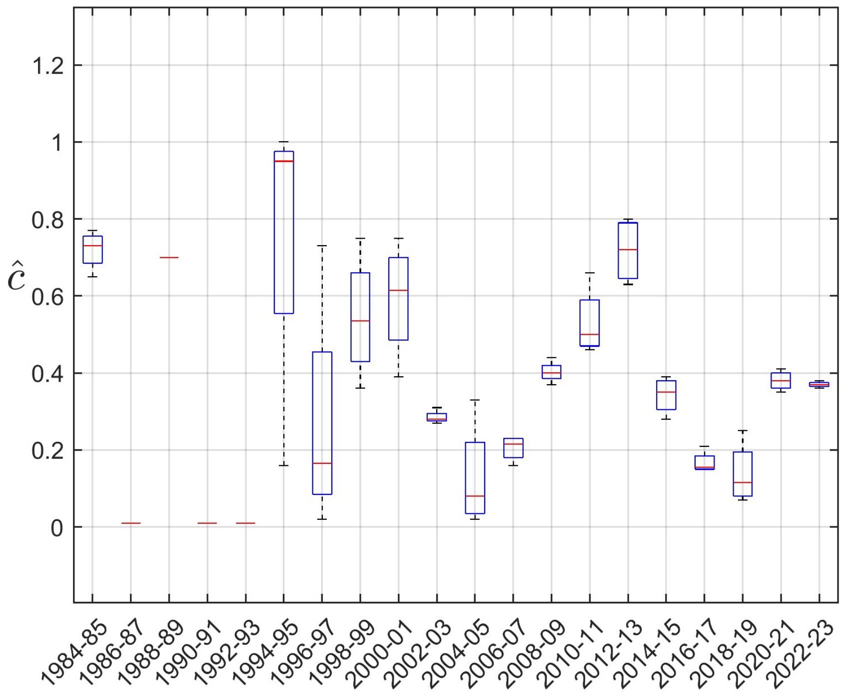
Table 1.6: FTSE 100 – Tabulated Results from the Exercise Conducted in Dhesi et al. (2019), Employing 1000 Simulations

years	$\hat{K}^{\chi}_{10}$	$\hat{K}^{\chi}_{100}$	$\hat{K}^{\chi}_{1000}$	$\hat{K}^{\chi}_{10000}$	$\hat{c}^{\chi}_{10}$	$\hat{c}^{\chi}_{100}$	$\hat{c}^{\chi}_{1000}$	$\hat{c}^{\chi}_{10000}$	$\hat{K}^{I}_{10}$	$\hat{K}^{I}_{100}$	$\hat{K}^{I}_{1000}$	$\hat{K}^{I}_{10000}$	$\hat{c}^{I}_{10}$	$\hat{c}^{I}_{100}$	$\hat{c}^{I}_{1000}$	$\hat{c}^{I}_{10000}$
1984-85	1	2	1	1	0.77	0.65	0.72	0.74	62	62	62	62	0.01	0.01	0.01	0.01
1986-87	15	13	14	14	0.01	0.01	0.01	0.01	10	8	8	8	1	1	1	1
1988-89	1	0	0	0	0.7	NaN	NaN	NaN	60	60	60	60	0.02	0.02	0.02	0.02
1990-91	9	9	12	11	0.01	0.01	0.01	0.01	40	35	40	39	1	1	1	1
1992-93	2	2	2	2	0.01	0.01	0.01	0.01	89	89	89	89	0.01	0.01	0.01	0.01
1994-95	3	2	2	1	0.16	0.95	0.95	1	2	2	2	3	1	1	1	1
1996-97	1	1	2	1	0.73	0.18	0.15	0.02	66	66	66	66	0.01	0.01	0.01	0.01
1998-99	4	4	3	3	0.36	0.5	0.75	0.57	100	100	100	100	0.01	0.01	0.01	0.01
2000-01	-5	-3	-4	-4	0.39	0.58	0.75	0.65	-5	-5	-4	-4	1	1	1	1
2002-03	-18	-14	-14	-15	0.28	0.31	0.28	0.27	-8	-9	-9	-9	1	1	1	1
2004-05	1	2	1	1	0.33	0.11	0.05	0.02	67	67	67	67	0.01	0.01	0.01	0.01
2006-07	6	7	7	8	0.23	0.16	0.2	0.23	6	8	7	7	1	1	1	1
2008-09	-21	-21	-21	-21	0.4	0.44	0.4	0.37	-11	-13	-13	-13	1	1	1	1
2010-11	91	100	100	100	0.66	0.52	0.46	0.48	62	95	94	93	1	1	1	1
2012-13	7	6	6	6	0.78	0.8	0.66	0.63	98	98	98	98	0.01	0.01	0.01	0.01
2014-15	-23	-20	-20	-20	0.28	0.33	0.39	0.37	-9	-14	-14	-13	1	1	1	1
2016-17	5	6	6	6	0.21	0.16	0.15	0.15	100	100	100	100	0.01	0.01	0.01	0.01
2018-19	-45	-46	-45	-45	0.07	0.09	0.25	0.14	-42	-36	-35	-35	1	1	1	1
2020-21	-100	-100	-100	-100	0.41	0.39	0.37	0.35	-57	-56	-56	-57	1	1	1	1
2022-23	-99	-100	-100	-100	0.38	0.37	0.37	0.36	-98	-63	-63	-68	1	1	1	1

Table 1.7: FTSE 100 – Tabulated Results of K and c Estimations Across Varying Simulation Counts.



(a) FTSE 100 – Box Plot Representation of Estimated K Values Across Different Year Pairs



(b) FTSE 100 – Box Plot Representation of Estimated c Values Across Different Year Pairs

Figure 1.10: FTSE 100 – These box plots depict the variability in the estimations of parameters K and c from Eqs (1.6) and (1.7), across different numbers of simulations employed in the estimation procedure (refer to Table 1.7 for a tabular presentation of these results).

1.8 Methodology - IFBM procedure & Shannon's Entropy

In response to the limitations observed in the Chi-square method, an exploration into alternative methodologies was initiated. A pivotal concept from the realm of information theory, which quantifies the extent of information inherent in entities such as events, random variables, and distributions, was employed as a linchpin for this exploration. Within the purview of this research, Shannon's Entropy (Shannon, 1948) is leveraged to quantify the variability, or the amount of information, embedded in a given probability distribution. The endeavour is to compute the entropy's for both observed and simulated returns distributions and to minimize the entropy-based distance between observed and simulated data, thus providing a nuanced understanding of the distributions.
The procedure employing Shannon's Entropy unfolds as follows:

1. Record the mean (μ) and standard deviation (σ) of returns for the specified period, utilizing the previously defined μ_{12-13} and σ_{12-13} for this instance.

2. Tabulate daily returns from the 2012-13 series, subsequently delineating probabilities of returns occurrences across 50 equal-length segments, as depicted in Figure 1.6b. The respective frequencies of each bin in the histogram are noted, from which the proportions of returns falling in each bin are derived. Denote these probabilities as po_h; $h = \{1, \ldots, 50\}$ (with 50 representing the number of bins), and identify the entire empirical distribution as $\bar{PO}$.

3. Similar to the preceding section, employ μ_{12-13}, σ_{12-13}, $\bar{K} = \{-100, \ldots, 0, \ldots, 100\}$ and $\bar{c} = \{0.01, 0.02, \ldots, 1\}$ to simulate returns series using Eq. (1.8). Perform simulations for each pair (K_m, c_m) generated from the linear combination of $\bar{K}$ and $\bar{c}$, resulting in $20,100$ pairs and a total of $20,100,000$ simulations. Each simulation yields returns tabulated as per Step 2, with the resulting probabilities averaged over 1000 simulations for each pair, denoted as $ps_{m,h}$; $m = \{1, \ldots, 20100\}$; $h = \{1, \ldots, 50\}$ (50 being the number of bins), and the entire empirical distribution defined as $\bar{PS}$. This step culminates in a $50 \times 20,100$ matrix M encapsulating the average probabilities per bin per pair.

4. Utilize Shannon's Entropy formulation Shannon (1948) to calculate the entropy for both observed (historical) and simulated data:

$$I(X) = -\sum_{h=1}^{n} p(x_h) \log_2 p(x_h) \tag{1.9}$$

wherein X denotes the distribution of a random variable across a set of states $\mathcal{X}$, and p represents the probability of a state. In this context, $n = 50$ and X signify the empirical distribution of both observed and simulated data. The term $p(x_h)$ encapsulates the probability of observing or simulating a value within the boundaries of bin h. The objective is to identify the pair (K_m, c_m) that minimizes the following expression:

$$(K_m, c_m) : \min \left(I(\bar{PO}_m) - I(\bar{PS}_m) \right) \ \forall\, m = \{1, \ldots, 20, 100\} \tag{1.10}$$

The optimal pair (K_m, c_m) minimizes the empirical distributions' distance per Shannon's Entropy criterion, denoted as $(\hat{K}_{I;m}, \hat{c}_{I;m})$. For the 2012-13 period, the pair minimizing the entropy difference is $(\hat{K}_{I;m} = 49, \hat{c}_{I;m} = 0.01)$.

The stability of the estimated parameters via the Entropy criterion is further evaluated by varying the simulation count. Specifically, price series are generated 10, 100, 1000, and $10,000$ times for each pair (K_m, c_m), examining the robustness of the estimations under different simulation scenarios. This entropy-based procedure presents a novel avenue for parameter estimation, aiming at a more coherent understanding of the underlying distributions and the behaviour of the IFBM under specified market conditions.

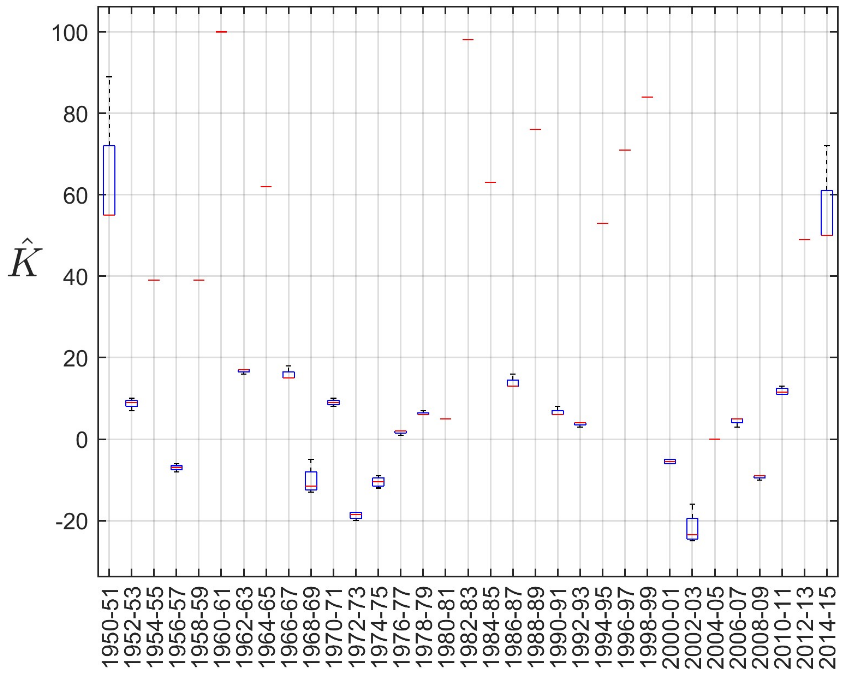
The outcomes of employing Shannon's Entropy procedure are meticulously tabulated and exhibited in Table 1.8 and Table 1.9. Table 1.9 enumerates the estimations of parameters K and c across a spectrum of simulation counts, utilizing the entropy-based procedure. The distinctive estimations of K and c across different year pairs are further illustrated visually through box plots in Figure 1.11. subfigures 1.11a and 1.11b distinctly delineate the variability in the estimations of parameters K and c, respectively. A more in-depth discussion on the implications and analysis of these results is nestled in Section 1.9.

years	μ	$\hat{K}$	$\hat{k}$	$\hat{c}$	$\hat{k}/\hat{c}$	$kurt_{obs}$	$kurt_{sim}$	$I_{obs;50}$	$I_{sim;50}$
1950-51	0.00071512	55	0.039331	0.01	3.9331	8.9124	2.9259	4.1791	4.8696
1952-53	8.3122e-05	9	0.0007481	1	0.0007481	6.2984	3.3986	4.4153	4.5616
1954-55	0.0011936	39	0.046552	0.01	4.6552	15.6885	2.83	3.7255	4.9509
1956-57	-0.0002422	-7	0.0016954	1	0.0016954	5.5254	2.6898	4.3664	4.4995
1958-59	0.00078455	39	0.030598	0.01	3.0598	3.8557	2.5208	4.6911	5.4575
1960-61	0.0003544	100	0.03544	0.01	3.544	6.7828	2.8968	4.0835	5.0723
1962-63	0.00011083	17	0.0018842	1	0.0018842	16.248	2.739	3.5866	3.9271
1964-65	0.00040327	62	0.025003	0.01	2.5003	6.7526	2.8358	4.1632	5.5368
1966-67	9.0615e-05	15	0.0013592	1	0.0013592	4.7018	3.0304	4.527	4.6604
1968-69	-9.0638e-05	-12	0.0010877	1	0.0010877	3.8759	3.1776	4.6972	4.7907
1970-71	0.0001843	9	0.0016587	1	0.0016587	6.4175	2.9228	4.2579	4.4335
1972-73	-8.2405e-05	-20	0.0016481	1	0.0016481	4.4705	3.1088	4.6354	4.7491
1974-75	-0.00015798	-9	0.0014218	1	0.0014218	3.6962	2.7908	4.8081	4.892
1976-77	8.9621e-05	2	0.00017924	1	0.00017924	2.8908	3.3784	5.0889	5.1147
1978-79	0.00027817	6	0.001669	1	0.001669	4.9089	2.7959	4.3504	4.4656
1980-81	0.00029178	5	0.0014589	1	0.0014589	3.4958	3.0376	4.7984	4.8698
1982-83	0.00058506	98	0.057335	0.01	5.7335	4.6627	3.1006	4.4355	5.473
1984-85	0.00050213	63	0.031634	0.02	1.5817	4.0292	2.9854	4.8905	5.2121
1986-87	0.00032586	13	0.0042362	1	0.0042362	80.6733	2.6758	2.8301	3.413
1988-89	0.00064019	76	0.048655	0.01	4.8655	13.4975	3.0237	3.9816	4.8731
1990-91	0.00029319	6	0.0017591	1	0.0017591	4.2283	2.8067	4.6344	4.7284
1992-93	0.00022024	4	0.00088096	1	0.00088096	4.1181	2.9386	4.662	4.7656
1994-95	0.00055696	53	0.029519	0.01	2.9519	4.4817	3.3487	4.572	5.4268
1996-97	0.00088309	71	0.062699	0.01	6.2699	10.2059	3.1585	3.8749	5.2474
1998-99	0.00081517	84	0.068474	0.01	6.8474	5.9224	2.5956	4.2557	5.238
2000-01	-0.00047508	-5	0.0023754	1	0.0023754	4.4182	3.1689	4.5663	4.6969
2002-03	-7.5003e-05	-23	0.0017251	1	0.0017251	4.1436	3.0062	4.7631	4.8485
2004-05	0.00023615	0	0	0	NaN	2.8668	2.9146	5.1041	5.1448
2006-07	0.00029157	5	0.0014578	1	0.0014578	5.3257	3.0708	4.5483	4.7464
2008-09	-0.00051718	-10	0.0051718	1	0.0051718	7.2996	3.1718	4.1922	4.484
2010-11	0.00020745	11	0.0022819	1	0.0022819	6.0113	3.1066	4.3415	4.5555
2012-13	0.000738	49	0.036162	0.01	3.6162	4.1087	3.1577	4.8022	5.4256
2014-15	0.0005089	50	0.025445	0.01	2.5445	4.5575	2.7866	4.8473	5.1029

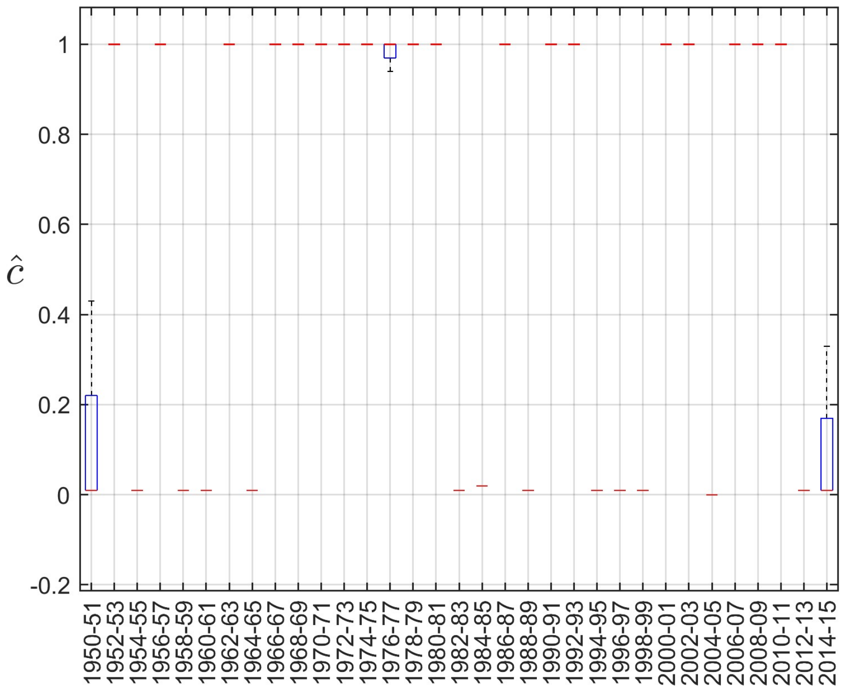
Table 1.8: Results derived from the methodology for estimating $(\hat{K}, \hat{c})$ utilizing Shannon's Entropy, as delineated in Section 1.8, with a set of 1000 simulations. Notably, $kurt_{obs}, I_{obs}, kurt_{sim;50}, I_{sim;50}$ represent the kurtosis and entropy, respectively, computed on the observed returns for the period and on the returns simulated with the optimal $(\hat{K}, \hat{c})$. The subscript 50 signifies that the n in Eq. (1.9) is 50, corresponding to the number of bins utilized in Step 2a of the specified procedure.

years	μ	$\hat{K}_{10}$	$\hat{K}_{100}$	$\hat{K}_{1000}$	$\hat{K}_{10000}$	$\hat{c}_{10}$	$\hat{c}_{100}$	$\hat{c}_{1000}$	$\hat{c}_{10000}$
1950-51	0.00071512	89	55	55	55	0.43	0.01	0.01	0.01
1952-53	8.3122e-05	10	7	9	9	1	1	1	1
1954-55	0.0011936	39	39	39	39	0.01	0.01	0.01	0.01
1956-57	-0.0002422	-8	-6	-7	-7	1	1	1	1
1958-59	0.00078455	39	39	39	39	0.01	0.01	0.01	0.01
1960-61	0.0003544	100	100	100	100	0.01	0.01	0.01	0.01
1962-63	0.00011083	16	17	17	17	1	1	1	1
1964-65	0.00040327	62	62	62	62	0.01	0.01	0.01	0.01
1966-67	9.0615e-05	18	15	15	15	1	1	1	1
1968-69	-9.0638e-05	-5	-13	-12	-11	1	1	1	1
1970-71	0.0001843	10	9	9	8	1	1	1	1
1972-73	-8.2405e-05	-18	-18	-20	-19	1	1	1	1
1974-75	-0.00015798	-11	-12	-9	-10	1	1	1	1
1976-77	8.9621e-05	2	2	2	1	0.94	1	1	1
1978-79	0.00027817	7	6	6	6	1	1	1	1
1980-81	0.00029178	5	5	5	5	1	1	1	1
1982-83	0.00058506	98	98	98	98	0.01	0.01	0.01	0.01
1984-85	0.00050213	63	63	63	63	0.02	0.02	0.02	0.02
1986-87	0.00032586	16	13	13	13	1	1	1	1
1988-89	0.00064019	76	76	76	76	0.01	0.01	0.01	0.01
1990-91	0.00029319	8	6	6	6	1	1	1	1
1992-93	0.00022024	3	4	4	4	1	1	1	1
1994-95	0.00055696	53	53	53	53	0.01	0.01	0.01	0.01
1996-97	0.00088309	71	71	71	71	0.01	0.01	0.01	0.01
1998-99	0.00081517	84	84	84	84	0.01	0.01	0.01	0.01
2000-01	-0.00047508	-6	-6	-5	-5	1	1	1	1
2002-03	-7.5003e-05	-16	-24	-23	-25	1	1	1	1
2004-05	0.00023615	0	0	0	0	0	0	0	0
2006-07	0.00029157	3	5	5	5	1	1	1	1
2008-09	-0.00051718	-9	-9	-10	-9	1	1	1	1
2010-11	0.00020745	13	11	11	12	1	1	1	1
2012-13	0.000738	49	49	49	49	0.01	0.01	0.01	0.01
2014-15	0.0005089	72	50	50	50	0.33	0.01	0.01	0.01

Table 1.9: Tabulated results of $(\hat{K}, \hat{c})$ estimations under varying simulation counts, employing the procedure encapsulated in Shannon's entropy as detailed in Section 1.8.



(a) Box plot illustrating the variability in the estimations of K_s across different year pairs.



(b) Box plot illustrating the variability in the estimations of c_s across different year pairs.

Figure 1.11: These box plots delineate the varied estimations of K and c from Eqs (1.6) and (1.7), contingent on the number of simulations in the estimation procedure employing Shannon's entropy (refer to Table 1.9 for a tabular exposition). The red markers signify the medians, while the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points and are not categorized as outliers. The visualizations suggest a lack of convergence to specific values as the number of simulations escalates from 10 to $10,000$; notably, the estimations of $\hat{K}$ exhibit more promise compared to $\hat{c}$.

1.9 Results - S&P500 Data

1.9.1 Parameter Stability & method agreement

The procedure outlined in Section 1.8 yields the estimated parameters $(\hat{K}, \hat{c})$ for each pair of years considered in this study. Table 1.9 delineates these estimations under Shannon's entropy methodology, while Table 1.5 presents the results under the Chi-Square methodology. A joint examination of Figure 1.11 and Table 1.9 sheds light on the stability of these parameter estimations across different simulation counts—specifically 10, 100, 1000, and $10,000$ simulations.

1.9.1.1 Stability Analysis under Shannon's Entropy

Observing the estimations under Shannon's entropy, the $\hat{K}$ parameter predominantly showcases stability across different simulation counts. For a majority of the years, the values remain consistent, indicating a robustness in the estimation process. Notably, the estimations for years such as 1954-55, 1958-59, and 1982-83 are exemplary of this stability, with their values remaining unchanged across varied simulation counts. There are a few years that experience minor deviations; however, these changes are not drastic enough to challenge the overall consistency observed for $\hat{K}$.

On the other hand, the $\hat{c}$ parameter presents a mixed bag. While a significant portion of the years exhibit consistent values across simulations, suggesting stability, there are specific years that display more pronounced variation. A standout example is 1950-51, where $\hat{c}$ alters notably from 0.43 at 10 simulations to a mere 0.01 at 100 simulations, and it maintains this value for higher simulation counts. Such instances underscore the relative variability in $\hat{c}$ estimations compared to $\hat{K}$.

In conclusion, while the $\hat{K}$ estimations under Shannon's entropy are largely stable across varying simulation counts, the $\hat{c}$ estimations demonstrate occasional variability, necessitating a more cautious interpretation for certain years.

1.9.1.2 Stability Analysis under Chi-Square Methodology

Observing the estimations using the Chi-Square-based procedure, it's evident that the $\hat{K}$ parameter displays a considerable degree of stability across varying simulation counts. For instance, in the years 1952-53 and 1986-87, the $\hat{K}$ estimations remain relatively stable, with values of [14, 14, 15, 15] and [28, 29, 29, 29] across the simulation counts respectively. However, in years like 2002-03, there's a slight yet noticeable change in $\hat{K}$ values from -36 to -34 as simulations increase, although this variation is not drastic.

Conversely, the $\hat{c}$ estimations present a more varied picture. While many years indicate stability, certain years such as 1992-93 and 2004-05 highlight the inherent variability. Specifically, for 1992-93, $\hat{c}$ values fluctuate from 0.49 at 10 simulations to 0.74 at 100 simulations, then slightly decrease to stabilize around 0.67-0.68 for higher counts. For 2004-05, the $\hat{c}$ estimation increases

from 0.94 at 10 simulations to stabilize at 1 for subsequent higher simulation counts.
In summation, while the $\hat{K}$ estimations predominantly exude stability, the $\hat{c}$ estimations, though generally stable, have certain years where more pronounced variations are observed, necessitating a nuanced interpretation.

1.9.1.3 Methods agreement - Shannon's entropy vs Chi-Square

A visual comparison of the results from both procedures is portrayed in Figure 1.12. Comparing the estimations from Shannon's entropy method with those from the Chi-Square procedure, we observe the following:

For $\hat{K}$ estimations:

- The mean difference in $\hat{K}$ values for 10 simulations is approximately 29.55.
- For 100 simulations, the mean difference is about 28.12.
- For 1000 simulations, the mean difference is around 27.82.
- For 10000 simulations, the mean difference stands at about 27.79.

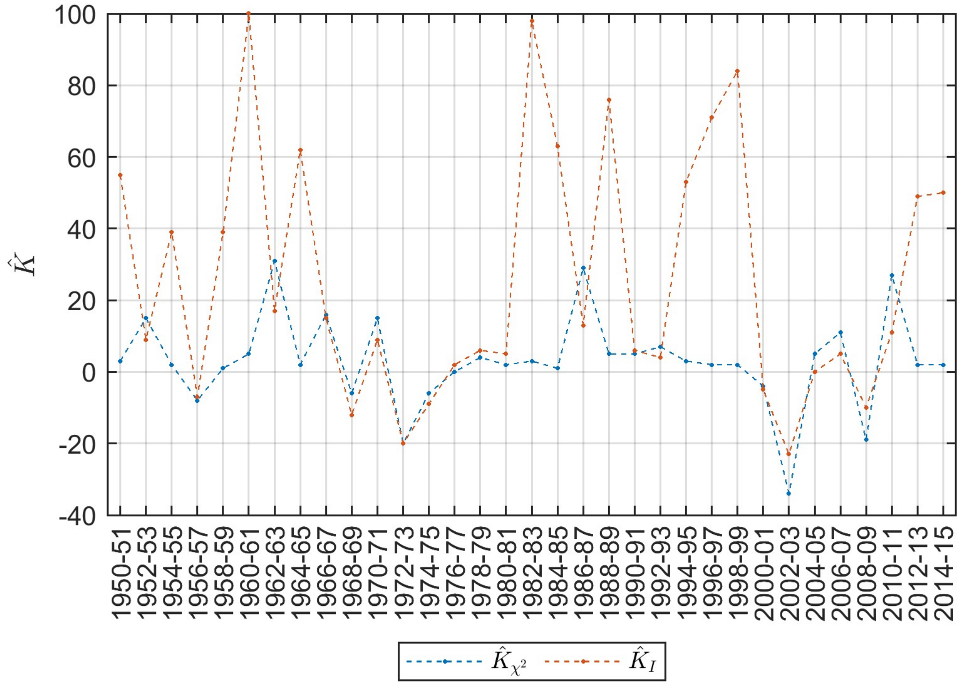
Both methods exhibit a degree of difference in $\hat{K}$ estimations across different simulation counts. Although differences exist, they seem to narrow slightly as the simulation count increases. Notably, Shannon's entropy method generally provides values that either exceed or fall short of those from the Chi-Square method by a consistent margin across different simulation counts.

For $\hat{c}$ estimations:

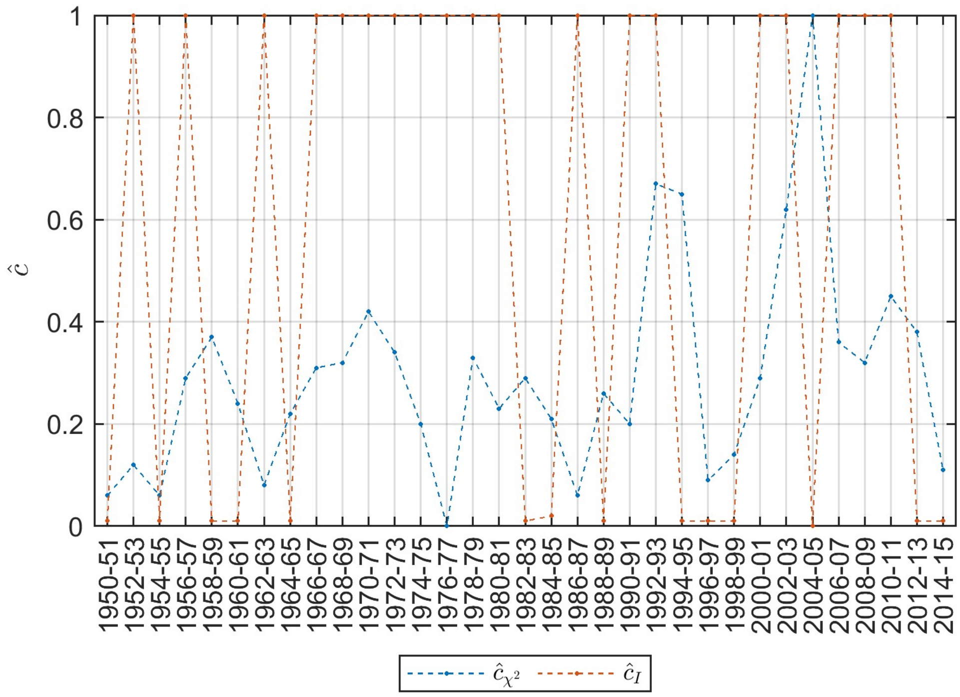
- The mean difference in $\hat{c}$ values for 10 simulations is approximately 0.537.
- For 100 simulations, the mean difference is about 0.500.
- For 1000 simulations, the mean difference is around 0.525.
- For 10000 simulations, the mean difference is approximately 0.532.

The differences in $\hat{c}$ estimations between the two methods are relatively smaller than those for $\hat{K}$. The Shannon's entropy method appears to provide marginally more consistent $\hat{c}$ values compared to the Chi-Square method, especially as simulation counts rise, although the disparities are not significant.

In conclusion, while both methodologies yield relatively stable estimations, Shannon's entropy method is slightly more consistent, particularly for $\hat{c}$ estimations. The consistent parameter estimations across varying simulation counts hint at the model's stability, especially in Shannon's entropy approach. Such stability underscores the model's robustness and reliability in parameter estimation, which is vital for accurately depicting real-world financial dynamics.



(a) $\hat{K}s$ estimated with both the procedures and compared.



(b) $\hat{c}s$ estimated with both the procedures and compared.

Figure 1.12: Comparison of $(\hat{K}_\star, \hat{c}_\star)$ when estimated via $\star = \{\chi^2, I\}$.

1.10 Results - FTSE100 Data

In this section, we delve into the simulation results for the FTSE 100 data, as presented in Table 1.7. We aim to provide insights into the stability and agreement of the $\hat{K}$ and $\hat{c}$ estimations across different simulation counts and compare the findings with those of the S&P 500 data.

1.10.1 Stability of $\hat{K}$ and $\hat{c}$ Estimations

1. **For the Chi-Square method**:
   - $\hat{K}$: For most years, $\hat{K}$ estimations are relatively consistent across varying simulation counts. For instance, in years like "1984-85", "1992-93", and "2016-17", $\hat{K}$ values remain relatively stable across different simulation counts.
   - $\hat{c}$: The $\hat{c}$ values exhibit more fluctuation across the simulation counts. For instance, in "1984-85", $\hat{c}$ changes from 0.77 at 10 simulations to 0.65 at 100 simulations and fluctuates further as the count increases.
2. **For the Shannon's entropy method**:
   - $\hat{K}$: The values of $\hat{K}$ are generally stable across different simulation counts. For example, in years like "1988-89" and "2004-05", the $\hat{K}$ estimations remain consistent irrespective of the simulation count.
   - $\hat{c}$: The $\hat{c}$ estimations remain highly consistent across simulation counts for almost all years.

1.10.2 Agreement Between Chi-Square and Shannon's Entropy Methods

1. **For $\hat{K}$ estimations**: The values estimated using Shannon's entropy method are generally higher than those obtained from the Chi-Square method. For instance, in the year "1986-87", while the Chi-Square method estimates values around 13-15, the Shannon's entropy method gives values around 8-10.
2. **For $\hat{c}$ estimations**: There are notable differences in the $\hat{c}$ estimations between the two methods. In many years, Shannon's entropy method tends to estimate $\hat{c}$ values close to the extremes (either near 0 or 1), while the Chi-Square method provides more varied values.

1.10.3 Comparative Analysis with S&P 500

1. **Stability**: The FTSE 100 data shows a similar trend in stability as the S&P 500 data, with $\hat{K}$ estimations being more stable across simulation counts compared to $\hat{c}$ estimations. However, the Shannon's entropy method's estimations for the FTSE 100 seem to be even more stable compared to the S&P 500 data, especially for $\hat{c}$.

2. **Agreement**: The discrepancies between the two methods' estimations are more pronounced for the FTSE 100 data than the S&P 500 data. Particularly for $\hat{c}$ values, Shannon's entropy method's tendency to estimate values close to 0 or 1 is more evident in the FTSE 100 dataset.

In conclusion, both the FTSE 100 and S&P 500 datasets exhibit relative stability in the $\hat{K}$ and $\hat{c}$ estimations across varying simulation counts, with Shannon's entropy method generally being more stable. The discrepancies between the two methods are more pronounced for the FTSE 100 data.

1.11 IFBM behaviour interpretation

Table 1.8 encapsulates the estimation results achieved with 1000 simulations of the series. A visual comparison of the results from both procedures is portrayed in Figure 1.12.
The $\hat{K}$ estimations exhibit a degree of agreement, particularly between 1966 and 1981, and 2000 to 2011, which are characterized by relatively stable market conditions. Conversely, $\hat{c}_I$ exhibits a dichotomous behaviour with values starkly polarized.
For instance, when $\hat{c}_I = 1$, $\hat{K}_I$ fluctuates between 17 and -23. The extreme cases are as follows:

- The parameter pair ($\hat{K}_I = 17, \hat{c}_I = 1$) occurs during 1962-1963, with a returns' average of 0.0001 for the period. For the same period, the chi-square criterion yields ($\hat{K}_{\chi^2} = 31$, $\hat{c}_{\chi^2} = 0.08$).
- The parameter pair ($\hat{K}_I = -23, \hat{c}_I = 1$) occurs during 2002-2003, with a returns' average of -0.000007 for the period[2]. The parameters for the same period via the chi-square criterion are ($\hat{K}_{\chi^2} = -34$, $\hat{c}_{\chi^2} = 0.62$).

To interpret the above, we resort to Figure 1.13, and reference Dhesi and Ausloos (2016), where the "feedback function" is plotted for the relevant cases discussed herein.
In Figure 1.13, a comparative analysis between the Chi-square and entropy-based methods in terms of their corrective influence on the Geometric Brownian Motion (GBM) is elucidated through the depiction of blue and dashed yellow lines.

- **Chi-Square Method (Dashed Yellow Line):** The chi-square-based estimations exhibit a dominant effect on the GBM base function, as the "feedback function" operates more intensively, either adding or deducting more, to adjust the distribution or fatten the tails accordingly with the value Z_t taken from the standard normal. This method aims to align the shape of the distributions, specifically targeting the kurtosis to ensure a closer match between the observed and simulated distributions. It's noteworthy that the level of kurtosis for 1962-63 is 16.248, necessitating a much higher corrective effort compared to the 2002-03 case. This is mirrored in the oscillations of the "feedback function," signifying a more aggressive corrective action to align the simulated kurtosis with the observed kurtosis.
- **Entropy-Based Method (Blue Line):** In contrast to the Chi-Square method, the entropy-based method exhibits a more gentle corrective action on the GBM. The conditions observed in terms of entropy are much closer to the observed data's entropies, leading to the "feedback function" interventions being closer to each other. This suggests a finer calibration in the entropy-based method as it seeks to match the information content between the observed and simulated distributions. The 1962-63 case highlights

[2]Note that in Figure 1.13's legend the number has been rounded.

an interesting observation where the entropy in the observed data is lower, hinting at a tendency of the GBM function to overestimate the returns' distribution entropy. Hence, the "feedback function" in this method tends to correct more in such instances to achieve a closer alignment with the observed data's entropy.

The distinct regimes of the IFBM, as depicted through the Chi-Square and Shannon's entropy methods, underscore the differential emphasis each method places on aligning either the distribution shape or the information content with the empirical data. This comparative analysis enriches the understanding of the corrective dynamics exerted by each method on the GBM, showcasing the relative merits and considerations inherent in each approach.

The "feedback function", contingent on the value Z_t, either augments or diminishes the GBM component of Eq. (1.6), albeit with varying magnitudes of effect.

The chi-square-based estimations exert a more dominant influence on the GBM base function compared to the entropy-based estimations. The level of kurtosis for 1962-63 is 16.248, indicating a heightened corrective effort on the GBM, as reflected in the oscillations of the "feedback function".

The conditions observed in terms of entropy are more congruent with the observed data's entropies, leading to closer "feedback function" interventions. The entropy estimated closely mirrors the observed one, more so than the kurtosis.

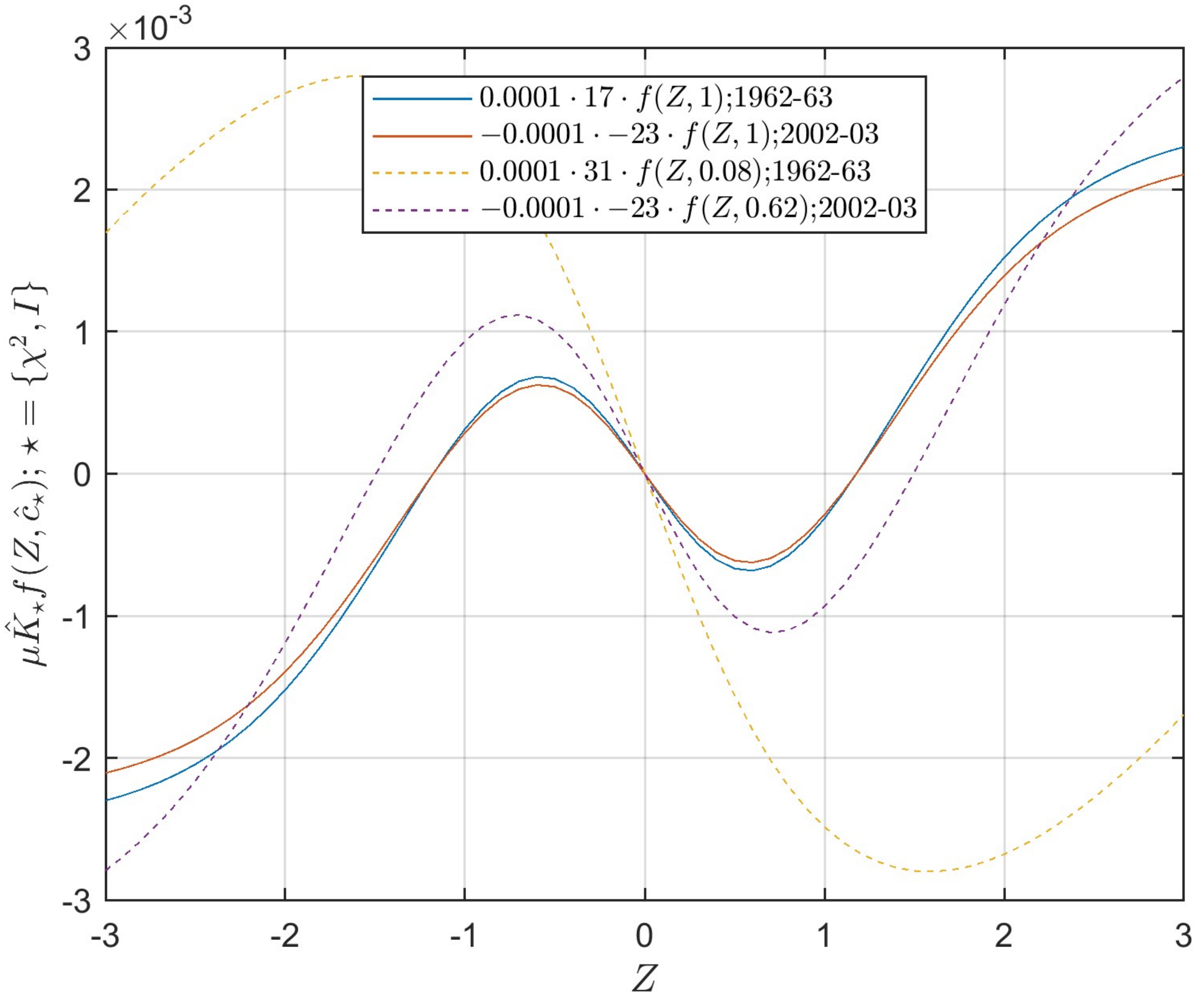


Figure 1.13: This figure contains the function $\mu \cdot \hat{K} \cdot f(Z, \hat{c})$ when it is fed by:

years	μ	$\hat{K}_{\chi^2}$	$\hat{K}_I$	$\hat{c}_{\chi^2}$	$\hat{c}_I$	$kurt_{obs}$	$kurt_{sim}$	$I_{obs;50}$	$I_{sim;50}$
1962-63	0.00011083	31	17	0.08	1	16.248	2.739	3.5866	3.9271
2002-03	-7.5003e-05	-34	-23	0.62	1	4.1436	3.0062	4.7631	4.8485

By employing $\hat{c}_I$ to discriminate the results, the opposite case to $\hat{c}_I = 1$ is $\hat{c}_I \approx 0.01$ (except for 1984-85 when it is estimated at 0.02). Thus, for $\hat{c}_I \neq 1$, $\hat{K}_I$ ranges between 0 and 100. The extreme cases are:

- The parameter pair $(\hat{K}_I = 100, \hat{c}_I = 0.01)$ occurs during 1960-1961, with the returns' average for the period being 0.0003. For the identical period, ($\hat{K}_{\chi^2} = 5$, $\hat{c}_{\chi^2} = 0.24$) is obtained.
- The parameter pair $(\hat{K}_I = 0, \hat{c}_I = \text{`}NaN')$ occurs during 2004-2005, with the returns' average for the period being 0.0002 (Note that when $\hat{K} = 0$, Eq. (1.6) reverts to the GBM, rendering the estimation of $\hat{c}$ redundant). Conversely, the parameters for the same period estimated via the chi-square criterion are ($\hat{K}_{\chi^2} = 5$, $\hat{c}_{\chi^2} = 1$).

Figure 1.14 contains the "feedback" function when the parameters discussed above are used. The entropy-based estimations determine $\hat{K}_I = 0$ for the years 2004-05, essentially proving that the GBM estimate better than the IFBM the empirical distribution in such years when it comes to considering distributions' entropy. In fact, when $\hat{K} = 0$, Eq. (1.6) represents the GBM, the "feedback function" does not act. Indeed, the estimated kurtosis is very close to the observed one; a very similar situation can be appreciated in terms of entropy. It is worth noting that the level of kurtosis for 1960-61 is 6.78, so the effort necessary to correct the GBM is much bigger than in the 2004-05 case. So, also, in this case, the oscillations of the "feedback function" prove it.

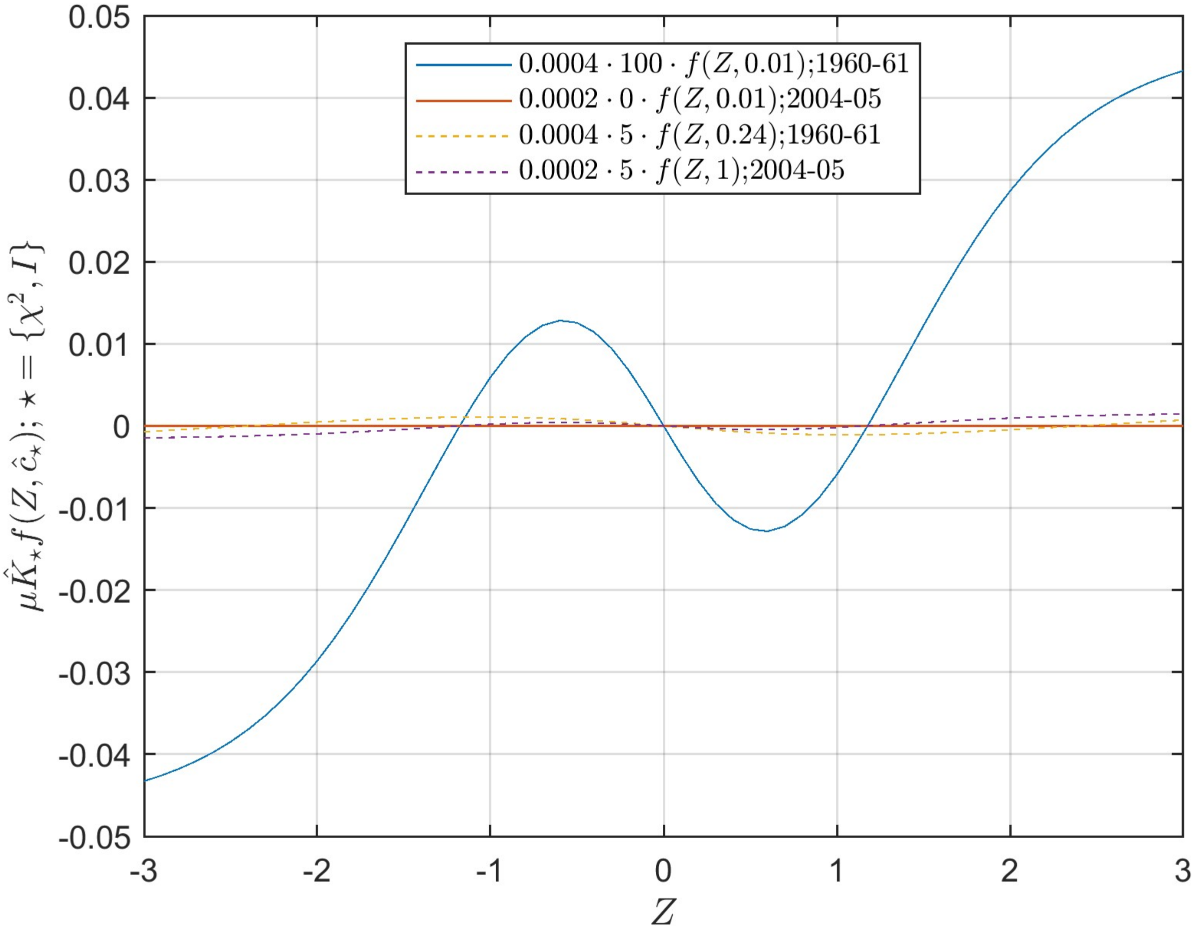


Figure 1.14: This figure contains the function $\mu \cdot \hat{K} \cdot f(Z, \hat{c})$ when it is fed by:

years	μ	$\hat{K}_{\chi^2}$	$\hat{K}_I$	$\hat{c}_{\chi^2}$	$\hat{c}_I$	$kurt_{obs}$	$kurt_{sim}$	$I_{obs;50}$	$I_{sim;50}$
1960-61	0.0003544	5	100	0.24	0.01	6.7828	2.8968	4.0835	5.0723
2004-05	0.00023615	5	0	1	0	2.8668	2.9146	5.1041	5.1448

1.12 Analysis of FTSE 100 Data: Understanding the IFBM Model's Response to Market Dynamics

The FTSE 100 data, especially during the years 2020-23 affected by the COVID-19 pandemic, offers an insightful perspective on the performance and adaptability of the Irrational Fractional Brownian Motion (IFBM) model.

1.12.1 Revisiting the IFBM Model

To appreciate the nuanced responses of the IFBM model to market dynamics, it's essential to break down its mathematical foundation:

$$P_{t+\delta t} = P_t \exp\left(\mu\delta t + \sigma Z_t\sqrt{\delta t} - \mu\hat{K}f(Z_t)\delta t\right)$$

with the feedback function defined as:

$$f(Z_t) = \left[2\exp\left(-\hat{c}\frac{Z_t^2}{2}\right) - 1\right]\arctan(Z_t)$$

This equation highlights the pivotal roles of $\hat{K}$ and $\hat{c}$ in modulating the model based on market feedback.

Impact of Varying $\hat{K}$ Values

- $\hat{K} = 0$: When $\hat{K}$ is zero, the feedback term $-\mu\hat{K}f(Z_t)\delta t$ becomes null, making the IFBM equation revert to the standard GBM model. In this scenario, the model doesn't account for any market feedback and solely relies on the GBM's predictions.
- Positive $\hat{K}$: A positive $\hat{K}$ amplifies the GBM component, suggesting the feedback mechanism is working in tandem with the GBM's predictions. This can indicate that the market is behaving in a manner consistent with the GBM's expectations.
- Extremely Negative (or positive) $\hat{K}$: As observed during the COVID-19 pandemic, an extremely negative $\hat{K}$ value intensifies the feedback mechanism's effect, causing the model to heavily counteract the GBM's predictions. This is the model's way of introducing significant corrections to adapt to adverse market conditions.

Impact of Varying $\hat{c}$ Values

- Low $\hat{c}$: A low $\hat{c}$ value means the feedback mechanism is less sensitive to market movements. The feedback function $f(Z_t)$ would decay slowly, making the model less reactive to market changes.

- High $\hat{c}$: A high $\hat{c}$ value, as seen during the COVID-19 pandemic, implies the model's feedback mechanism is highly sensitive to market movements. This ensures the model remains reactive to minor market shifts but scales down its feedback for larger shocks, preventing overcorrections and ensuring stability.

Balancing $\hat{K}$ and $\hat{c}$

The interplay between $\hat{K}$ and $\hat{c}$ is crucial. While $\hat{K}$ determines the magnitude of the feedback, $\hat{c}$ scales it. Their combined effect ensures that the IFBM model remains adaptive, capturing the intricacies of real-world market dynamics while ensuring stability in its predictions.

1.12.2 Deciphering the Impact of COVID-19 on $\hat{K}$ and $\hat{c}$

During the pandemic years, the notably negative values of $\hat{K}$, especially the extreme of -100, indicate a strong negative feedback in the market returns. This negative feedback acts to counteract positive market trends, suggesting that the IFBM model was adjusting for adverse market conditions. Essentially, the feedback mechanism was working overtime to correct and stabilize the model's predictions, highlighting its adaptability during such tumultuous periods.

The consistent $\hat{c}$ values, especially from Shannon's entropy method, signify the stable scaling factor applied to this feedback across various market scenarios. A high $\hat{c}$ value implies that the model remained sensitive to smaller market movements while ensuring that the feedback mechanism didn't overreact to larger market shocks.

1.12.3 The Feedback Mechanism and Market Dynamics

An extremely negative $\hat{K}$ value, like -100, intensifies the feedback mechanism's effect, causing the model to heavily counteract the GBM's predictions. This could be seen as the IFBM's attempt to bring its predictions closer to the real-world market dynamics during significant disruptions, like the COVID-19 pandemic.

Conversely, the high $\hat{c}$ values show the model's intent to be reactive to minor market movements but prevent overcorrection during major market upheavals. This behaviour showcases the IFBM model's potential to balance sensitivity and stability, even in challenging market conditions.

1.12.4 Efficacy of the IFBM Model During the COVID-19 Pandemic

The pronounced negative $\hat{K}$ values during the COVID-19 period emphasize the IFBM model's ability to recognize and adjust for adverse market conditions. The model's negative feedback response aimed to stabilize its predictions during the pandemic, ensuring it remained adaptive. The high $\hat{c}$ values further reinforced this adaptability, ensuring a consistent scaling of feedback across various market scenarios.

Through its parameters $\hat{K}$ and $\hat{c}$, the IFBM model demonstrated its potential to adapt to real-world financial dynamics, especially during the unprecedented challenges posed by the COVID-19 pandemic. The model's feedback mechanism, governed by these parameters, ensured that it remained both sensitive to market shifts and stable in its predictions, underscoring its relevance and reliability in capturing market dynamics.

1.13 Implications

1.13.1 Insights from Parameter Stability and Method Agreement

1. **Robustness of Estimation Procedures**: The stability observed in the $\hat{K}$ parameter across different simulation counts under both Shannon's entropy and Chi-Square methodologies underscores the robustness of these estimation procedures. This stability indicates that the model's response to market feedback, as governed by $\hat{K}$, can be reliably estimated using either method.

2. **Variability in Feedback Scaling**: The $\hat{c}$ parameter, which scales the feedback mechanism, displays more variability, especially under Shannon's entropy methodology. This suggests that the intensity of feedback might be more challenging to pin down precisely and could vary more significantly based on the estimation method and the number of simulations.

3. **Methodological Agreement and Discrepancies**: While both Shannon's entropy and Chi-Square methodologies offer valuable insights, their differences in $\hat{K}$ and $\hat{c}$ estimations, especially as simulation counts rise, emphasize the nuances in their approaches. Shannon's entropy method appears to be slightly more consistent, especially in $\hat{c}$ estimations. This implies that, while both methods aim to align the IFBM model with empirical data, they might prioritize different aspects or features of the data.

1.13.2 Understanding IFBM behaviour

1. **Consistent Feedback Mechanism**: The agreement observed in $\hat{K}$ estimations, especially during periods of relative market stability, suggests that the model's feedback mechanism remains consistent across these times. This implies that the model can reliably adjust its predictions based on market feedback during such periods.

2. **Dichotomous behaviour in Feedback Scaling**: The polarization observed in $\hat{c}_I$ values suggests that the model might operate in distinct regimes or modes. These regimes could reflect different market conditions or dynamics, with the model either being highly sensitive to market movements or largely insensitive.

3. **Comparative Dynamics of Chi-Square vs. Entropy-Based Methods**: The oscillations of the "feedback function" in the Chi-Square method reflect a more aggressive corrective action to align the simulated kurtosis with the observed kurtosis. In contrast, the entropy-based method seeks a finer calibration, aiming to match the information content between the observed and simulated distributions. This highlights the differential emphasis each method places on aligning either the distribution shape (kurtosis) or the information content (entropy) with empirical data.

1.13.3 IFBM's Response to COVID-19 and Market Dynamics

1. **Adaptive Feedback during Adverse Conditions**: The notably negative $\hat{K}$ values during the COVID-19 pandemic period emphasize the model's adaptability. The strong negative feedback suggests that the model was actively adjusting its predictions to account for adverse market conditions.

2. **Balancing Sensitivity with Stability**: The consistent high $\hat{c}$ values during this period indicate that, while the model was sensitive to market movements, it was also ensuring that the feedback mechanism did not overreact. This behaviour is crucial during major market upheavals, where overreactions could lead to erratic or unstable model predictions.

3. **Relevance and Reliability**: The IFBM model's performance during the challenging times of the COVID-19 pandemic underscores its potential as a reliable tool for capturing market dynamics. Its adaptive feedback mechanism, as evidenced by the pronounced negative $\hat{K}$ values and the consistently high $\hat{c}$ values, demonstrates its relevance in accurately representing real-world financial dynamics.

The insights derived from the stability and agreement analyses, combined with the deep dive into the IFBM behaviour and its response to the COVID-19 pandemic, paint a comprehensive picture of the model's capabilities. The IFBM model emerges as a robust and adaptive tool, capable of capturing intricate market dynamics and offering valuable insights into financial markets' behaviour.

1.14 Limitations, Future Work, & Concluding remarks

The analysis utilizing the Irrational Fractional Brownian Motion (IFBM) model on both S&P 500 and FTSE 100 data provides a richer understanding of asset price dynamics across different markets. However, certain limitations are inherent in the analysis, and opportunities for further exploration abound.

1.14.1 Limitations

1. **Model Assumptions:** The IFBM model, akin to many financial models, is constructed upon certain foundational assumptions which might not always mirror the intricate realities of financial markets (Cont, 2001). For instance, the potential dichotomous behaviour of the feedback scaling parameter $\hat{c}_I$ indicates that the model might not entirely encapsulate some intricate market behaviours. Furthermore, the reliance on parameters K and c to represent market dynamics can sometimes be an oversimplification, especially when considering the nuances of market participants' reactions to fresh information, as explored by (Black, 1976) during turbulent market situations or unparalleled financial events.

2. **Parameter Estimation:** While the Chi-Square and Entropy methods remain valuable tools for parameter estimation, they come with their inherent biases and sensitivities (Tsay, 2005). This is particularly evident in the discrepancies observed in parameter estimations across these techniques and between different markets (S&P 500 and FTSE 100). Notably, the $\hat{c}$ parameter displayed greater stability when analyzed under Shannon's entropy methodology than the chi-square approach. This not only emphasizes the relative sturdiness of the entropy method but also underscores the importance of methodological selection in financial modeling.

3. **Historical Data Dependency:** A prevalent challenge in financial modeling is the heavy reliance on historical data (Engle, 1982). While these data provide invaluable insights into past market behaviours, they may not always be indicative of future market movements. This becomes especially poignant in light of unprecedented market events or significant structural changes in the financial markets, as explored by (Lo, 2001). The challenges associated with extrapolating from historical data accentuate the need for integrating adaptive mechanisms within financial models.

1.14.2 Future Work

1. **Model Extension:** Consider incorporating macroeconomic indicators as exogenous variables to improve the IFBM model's forecasting performance. Indicators like interest rates, inflation rates, or GDP growth have been known to influence asset prices (Stock

and Watson, 1989). Furthermore, incorporating elements from other stochastic models such as Jump Diffusion (Merton, 1976) can address sudden market jumps.

2. **Robustness Checks:** A wider application of the IFBM model across different asset classes, including commodities and currencies, would help test its versatility. Previous research by (Engle and Ng, 1993) on various asset classes could provide a benchmark for comparison.

3. **Comparative Analysis:** A thorough comparison with other volatility models, such as the GARCH family of models (Bollerslev, 1986), can highlight the relative strengths of the IFBM model.

4. **Alternative Parameter Estimation Methodologies:** Exploring advanced estimation techniques, such as the Generalized Method of Moments (Hansen, 1982), could provide more robust parameter estimates, especially in the presence of potential model misspecifications.

5. **High-frequency Data Analysis:** With the increasing availability of tick-by-tick data, applying the IFBM model to high-frequency data can uncover intricate market dynamics that daily data might miss. Research by (Andersen and Bollerslev, 1998) on high-frequency data can serve as a guideline for such explorations.

6. **Market Resilience Analysis:** Investigate how the IFBM model's parameters, especially the feedback mechanism, change during major market events. Such an analysis can echo studies like (Forbes and Rigobon, 2002) that delve into market reactions during crises.

The IFBM model, through the analysis of both S&P 500 and FTSE 100, presents a promising framework for understanding asset price dynamics. The insights gleaned from this analysis emphasize the model's adaptability and potential reliability, especially in the face of significant market upheavals like the COVID-19 pandemic. These insights propel further exploration towards a more accurate and comprehensive understanding of financial markets.

1.14.3 Concluding remarks

The study embarked on a multi-faceted exploration of the Irrational Fractional Brownian Motion (IFBM) model, seeking to unveil its merits in capturing the intricacies of financial market dynamics, particularly when juxtaposed with the conventional Geometric Brownian Motion (GBM) model. The voyage was navigated through a methodological lens, an empirical spectrum, and a practical pathway, each contributing a unique shade to the picture we painted of the IFBM model's potential and performance.

The IFBM model emerged as a promising alternative to the GBM model, offering a richer representation of financial markets by embracing investor irrationality and leptokurtosis. Its acknowledgement of fatter tails and heightened peaks in asset price distributions unveiled a canvas where large price swings aren't anomalies but part of the market's unpredictable essence. This leap from the GBM model, which assumes rationality and overlooks leptokurtosis, establishes the IFBM model as a more realistic portrayal of market dynamics, echoing the empirical realities documented in seminal works ((Mandelbrot, 1963), (Taleb, 2007)).

The empirical expedition into the S&P 500's historical data rendered a testament to the IFBM model's stability and robustness. The stability in the k and c parameters across different simulation counts, especially when estimated using Shannon's entropy, portrays a model grounded in reality yet not swayed easily by market whims. Though the data didn't extend to the pandemic period, the stability observed underscores the model's potential reliability in varying market conditions.

The expedition further led to the formulation of a robust and repeatable algorithm, unlocking the gates for the IFBM model's broader application in financial market analyses. This methodological advancement, bridging the gap between complex theory and practical application, holds the promise of making the IFBM model a versatile tool for both researchers and practitioners.

The algorithm's foray into the tumultuous waters of the FTSE 100 during the COVID-19 crisis rendered visible the IFBM model's adaptability. The pronounced negative feedback mechanism captured by the $\hat{K}$ parameter during this period resonated with the market's adverse conditions, showcasing the model's potential to mirror market dynamics even in extreme scenarios.

Lastly, the employment of Shannon entropy as a lens to scrutinize the information content encapsulated by the IFBM model in portraying market dynamics unveiled a richer understanding of market behaviour and risks. This venture into the realm of information theory accentuated the model's potential in offering a nuanced understanding of the complex interplays inherent in financial markets, especially during turbulent times.

A particular highlight from the analysis is the observed stability in the parameter $\hat{c}$ across different market conditions and datasets. The entropy method, in particular, provided notably more stable estimates for $\hat{c}$ compared to the chi-square method. This consistent estimation, particularly with the entropy method, underscores a level of robustness in the IFBM model. This stability in parameter estimation is crucial as it suggests a reliable scaling factor for the feedback mechanism, aligning the model closer to empirical market dynamics.

In summary, the exploration of the IFBM model, through the complex dynamics of financial markets and the precise realm of methodological formulation, unveiled a model with potential as vast as the markets it seeks to represent. The empirical stability, methodological robustness, and practical applicability observed not only shed light on the model's merits but also pave the way for further explorations into its boundless potential in understanding the enigmatic behaviour of financial markets.

With a more robust understanding of the IFBM model established in Chapter 1, we transition to Chapter 2, delving into the realm of investor sentiment. The study of investor sentiment is particularly relevant as it offers insights into the behavioural aspects that could influence the parameters of the IFBM model. Here, the focus shifts to understanding whether investor sentiment, a distinct yet related aspect of market behaviour, propagates through a community of traders and impacts market dynamics, especially during crises. The juxtaposition of investor rationality/irrationality and sentiment forms a comprehensive narrative, exploring the different, yet interconnected, facets of behavioural finance. Both chapters aim to unravel the complexities of human behaviours that significantly influence market phenomena, emphasizing their paramount role in shaping market trends, especially during times of crises, enriching our understanding of the intricate dance between investor behaviour and market dynamics.

Chapter 2

Financial Markets in the COVID-19 Era

2.1 Introduction

The COVID-19 pandemic, recognized as a global health emergency by the World Health Organization (WHO) (WHO, 2020), has reverberated through economies worldwide, marking its place in the annals of events that reshaped the global financial landscape (Nicola et al., 2020). Historically, pandemics have shown a knack for disrupting economic equilibria, and financial markets, in particular, have often been at the epicentre of these disturbances (Barro et al., 2020).

Financial markets, often seen as the heartbeat of economies, have always been sensitive to global events. Their forward-looking nature means that they constantly adjust to new information, providing real-time feedback on the collective sentiments of the investment community (Fama, 1970). The advent of the COVID-19 pandemic amplified this inherent characteristic of financial markets.

The rapid spread of COVID-19 created significant uncertainty in financial markets, prompting concerns about the potential economic impacts of the pandemic (Goldstein et al., 2020). According to (Goodell and Huynh, 2020; Al-Awadhi et al., 2020; Elnahas et al., 2018), market returns respond negatively to any political, economic, or ecological crisis. Stock market crashes, and pandemics tend to go hand in hand over the long haul. However, the reaction to COVID-19 in 2020 is considered unprecedented (Baker et al., 2020a). Furthermore, Zeren (Zeren and Hizarci, 2020) argues that despite the initially modest number of illnesses and fatalities, the quick decrease in the stock markets is due to the markets' capacity to respond swiftly to incoming information. As (Fernandes, 2020) notes, the pandemic generated "unprecedented levels of uncertainty" for investors and policymakers, who were forced to grapple with a rapidly evolving situation and a range of potential outcomes. (Sansa, 2020) analysed the emotional impact of investing concerns and argued that "fears of a broader outbreak and its economic impact spread to financial markets."

The deluge of daily news, ranging from infection rates to vaccine developments, meant that markets were in a constant state of flux, trying to process and adapt to this new reality (Baker et al., 2020a). In 2020, the prominence of the COVID-19 pandemic in global media reached levels not seen since World War II, as illustrated by The Economist (Economist, 2020). This media focus led to a marked uptick in news consumption globally, especially in the UK (Statista, 2021). The UK government's regular updates on the pandemic became essential viewing for many, serving as a trusted information source during a time fraught with misinformation concerns (Nielsen et al., 2020). With millions tuning in to these briefings, they represented a significant news event during the pandemic.
In an era marked by the ubiquity of information and where the line between factual and misleading news often gets blurred, official communication channels have become the beacon of clarity. Governments and health organizations around the world stepped into this role, with their updates becoming critical for individuals and institutions alike. Economic literature has emphasized the symbiotic relationship between macroeconomic news and stock prices. Government announcements, given their authoritative nature and policy implications, often exert significant influence on market movements (MacKinlay, 1997). Within the tumultuous environment of the pandemic, these updates gained even more prominence. They not only reported on the current state of the health crisis but also indicated the direction of future economic policies, potential lockdown measures, and, most critically, vaccine-related developments—all of which are crucial components in the information landscape that investors navigate. Such information, alongside other global events and economic indicators, directly feeds into the collective psyche of the financial community.

Investor sentiment, which gauges the emotions and attitudes of individual or institutional investors, plays a pivotal role in determining market movements (Baker et al., 2016). This sentiment is a barometer, reflecting investors' collective reactions to various news items, economic indicators, and global events. During the pandemic, this sentiment experienced significant fluctuations. Notable events, such as the announcement of national lockdowns or breakthroughs in vaccine trials, acted as potent triggers, driving pronounced shifts in stock market behaviours (Zhang et al., 2020). Recognizing the profound influence of investor sentiment, especially during turbulent times like the COVID-19 pandemic, this study seeks to unravel its intricacies, examining it through the lens of government briefings.

Positioning the UK as the lens through which we examine these phenomena, this research aims to comprehensively understand the nexus between the COVID-19 narrative, government briefings, and stock market dynamics. By analysing the UK government's daily updates, we unravel their nuanced influences on stock market sentiments. This endeavour not only enhances our understanding of the multifaceted relationship between public health crises, official communication, and financial markets but also illuminates investors' decision-making processes during unprecedented times.

In the forthcoming sections, we will embark on a journey that traces the trajectory of COVID-19 and its multifaceted impact on global stock markets. We will then delve into the UK's governmental response, elucidating its pivotal role as a primary news source during this crisis. A thorough exploration of investor sentiment, its underlying significance, and its profound influence on stock markets will pave the way for our core investigation: we aim to illuminate the broader implications of governmental responses during pandemics on financial markets and the consequent investor decision-making processes.

2.1.1 COVID-19 Pandemic Timeline in the UK

Although the first cases of COVID-19 were recorded in 2019, the Western world at large became aware of the disease in January 2020. On January 29, 2020, the first two infections in the UK were confirmed. Six people died of the sickness in the United Kingdom on March 10th. Upon this, the FTSE, the UK's major index, fell more than 10% on March 12, 2020, on its worst day since 1987 (Zhang et al., 2020), and continued to fall to its lowest level on March 23rd, when the UK Prime Minister declared a national lockdown. The influence of this series of events on the share price of all businesses listed on the London Stock Exchange, the FTSE All-Share index, is shown in Figure 2.1. Between January 2 and March 23, 2020, the FTSE All-Share index dropped by 35%.

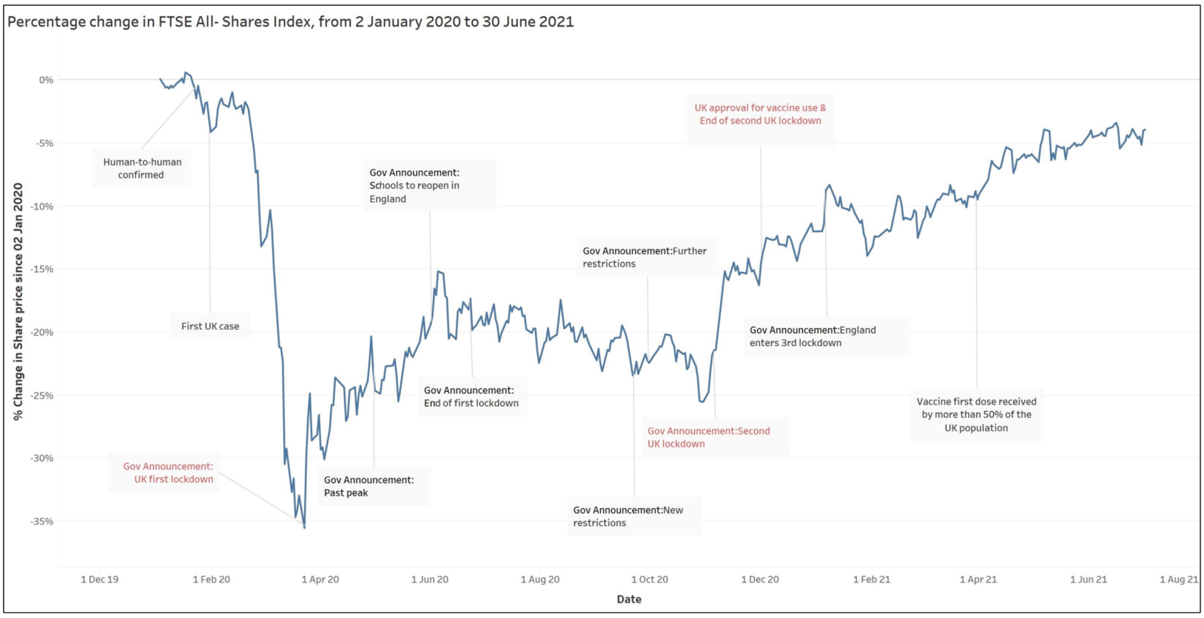


Figure 2.1: Percentage change in FTSE All-Share Index, from January 2, 2020, to June 30, 2021

The UK stock market started to rebound from its lowest point in March 2020 in April. Between April and June, restrictions such as the closure of schools were lifted, while UK Prime Minister

Boris Johnson claimed that the UK was ”past the peak” of COVID-19. Furthermore, the UK human COVID-19 vaccine trials started on April 22, 2020, offering hope for a viable therapy and a way out of the COVID-19 crisis. Between July and October, the share price dropped steadily again. On July 1, 2020, UK firms terminated 11,000 jobs in two days. Meanwhile, owing to an increase in new cases, the UK government reinstated fresh COVID-19 restrictions. The increase in cases might be ascribed to a variety of factors, including the loosening of mixing regulations in the absence of a cure. Furthermore, the £500 million “Eat Out to Help Out” scheme launched by the UK government in August was found to have increased new COVID-19 infections by 8% to 17% (Fetzer, 2022). The scheme was intended to support and create jobs in the hospitality industry to offset the economic impacts of the COVID-19 pandemic. However, this scheme involved indoor public mixing, which increased the spread of the disease.

On September 21, 2020, concerns of a resurgence of coronavirus outbreaks stripped off more than £50 billion in UK stock market value. Unsurprisingly, the Prime Minister of the United Kingdom announced a second four-week national lockdown on October 31, 2020, with non-essential shopping, bars, and restaurants shuttered (Group, 2022).

Former UK Health Minister Matt Hancock announced in November 2020 that a vaccine might be available the following month (for Government, 2022). This statement corresponded with the FTSE All-Share’s recovery and its continued steady ascent with the official approval of a COVID-19 vaccination in the UK. The FTSE All-Share has not dropped below -15% since the official clearance of the vaccine. By May 9, 2021, one-third of people in the UK had been fully vaccinated against COVID-19 (Government, 2021), providing the public and investors with optimism that the worst was behind them.

The vaccination campaign has been successful in the UK, and the government’s decision to prioritize vaccinating the elderly and vulnerable has likely contributed to the decrease in the number of COVID-19 cases and deaths in the country. As a result, the UK government has lifted most COVID-19 restrictions, allowing the economy to reopen and businesses to operate as usual. This has had a positive effect on the UK stock market, with the FTSE All-Share index reaching pre-pandemic levels by mid-2021 (Thomas, 2021).

2.1.2 Governments’ response in COVID-19

Amid economic uncertainties, governments play a pivotal role in shaping investor sentiment and allaying their fears (Eachempati et al., 2021). The extent and nature of government interventions not only stabilize the immediate market conditions but also provide valuable cues about potential future market volatility to those managing equity portfolios (Mirza et al., 2020). In the face of the unprecedented challenges posed by the COVID-19 pandemic, global governments and central banks acted swiftly. They implemented a myriad of stimulus measures to counterbalance the economic disruptions brought about by the virus, with the goal to instil renewed trust in the financial markets. Taking the United Kingdom as a representative example, on March 17, 2020, several initiatives were announced to aid businesses and employees affected by the pandemic's containment measures. These included a year-long business rate relief for entities in the retail, leisure, and hospitality sectors, as well as the Coronavirus Job Retention Scheme (CJRS). Through the CJRS, furloughed employees were assured of receiving 80% of their salaries, with a ceiling of £2,500 monthly per worker. Such measures were conceived to sustain businesses that would have been viable under normal circumstances (Griffith et al., 2020). Furthermore, the UK government prioritized transparency, holding daily press briefings to keep the nation abreast of the pandemic's developments. These updates, closely monitored by the British public—including stock market investors—were invaluable. As (Liu et al., 2020) suggests, navigating a crisis of the magnitude of COVID-19 requires the government to offer clear, timely information, helping the public understand the steps being taken without introducing ambiguity. This approach emphasized the importance of timely government communication in assisting investors during such tumultuous times. Recognizing this, several governments worldwide stepped up their communication efforts. One notable example is the United Kingdom. During the height of the coronavirus pandemic, the UK government initiated regular briefings from 10 Downing Street, the Prime Minister's official residence. These commenced on March 3, 2020, in response to public demands for clarity on government strategies against COVID-19, and while routine press briefings ceased on June 23, sporadic news conferences persisted (BBC). Typically, these briefings were chaired by prominent government figures, including the then-Prime Minister Boris Johnson, and frequently featured expert insights from the medical and scientific community.

This investigation operates under two primary assumptions. Firstly, these briefings acted as an informational bridge, linking technical pandemic experts with the general public. Their purpose was to relay current developments and delineate forthcoming strategies to manage the pandemic within the UK. Secondly, it is postulated that these conferences were keenly observed by investors, influencing their perceptions, investment decisions, and understanding of the pandemic's trajectory.

2.1.3 Government Briefings as the Unit of News in COVID-19

Given the context and the media landscape during the pandemic, this study designates the UK government's daily briefings as the primary "unit of news". This decision is rooted in their widespread coverage, high viewership, and the trust the UK public placed in them as reliable sources of information about COVID-19. In essence, when referring to "Government Briefings as the Unit of News in COVID-19", we emphasize the unique position these briefings held in the information dissemination chain, acting as a bridge between official pandemic responses and public perception and behaviour. In 2020, the COVID-19 pandemic occupied an unprecedented position in global media coverage. As per a report by The Economist (Economist, 2020), the pandemic has surpassed all other subjects in terms of news coverage since the era of the Second

World War. The periodical underscored that during the peak of the pandemic in late March, a notable 80% of their articles contained mentions of the terms "COVID" or "coronavirus".
This increased coverage coincided with a significant increase in media consumption habits worldwide, particularly in the UK, where people reported consuming news more frequently than before the pandemic (Statista, 2021). During the pandemic, the UK government's daily briefings were broadcast on major news outlets. TV and online media were both widely used as sources of information during the pandemic, with 70% of respondents using TV and 79% using online sources (Nielsen et al., 2020). In this article, the authors found that news organisations are the single most widely identified source of news and information about coronavirus.
On the other hand, the concern about misinformation regarding COVID-19 was widespread during the pandemic (Statista, 2020). In the UK, where the government briefings were broadcasted, the government briefings were seen as a reliable source of information during a time when there was a lot of concern about the spread of misinformation. According to a survey conducted by (Nielsen et al., 2020), both news organizations and the national government were predominantly deemed trustworthy sources for obtaining Coronavirus news.
The popularity of the government briefings is evident in the viewership numbers. Boris Johnson's COVID-19 press briefings were watched by millions of people in the UK, with his announcement of the first national lockdown being watched by an audience of 27.4 million across six networks (Guardian, 2020). The announcement of a £330bn lifeline for the UK economy during one of the briefings was expected to have a positive impact on the financial markets (News). This kind of announcement could potentially boost investor confidence and lead to an increase in stock prices and other financial indicators (Allen et al., 2011).
The widespread coverage and high viewership of the UK government's daily updates during the pandemic make them an effective unit of news for this study. Additionally, the trust that the UK public placed in the government as a source of information about COVID-19 further emphasizes the importance of government briefings in understanding public sentiment and behaviour during the pandemic.
In conclusion, the COVID-19 pandemic led to a significant increase in media consumption and the widespread coverage of the UK government's daily briefings, which were seen as a reliable source of information during a time of misinformation. The government briefings were popular, and their impact on the financial markets highlighted their importance for investors. The trust that the UK public placed in the government further emphasizes the significance of government briefings in understanding public sentiment and behaviour during the pandemic.

2.1.4 Government Briefings and Investors' Sentiment in COVID-19

Traditional finance theory posits that an asset's value equates to its anticipated future cash flows. In this paradigm, only systematic risk dictates the balance of expected returns, with any deviations or mispricing rectified by arbitrageurs. This conventional understanding suggests that stock returns and volatility patterns should remain unaffected by externalities like investors' sentiment or emotional state. However, research over the past two decades, including findings by (Huang et al., 2019), indicates that classic financial theories fall short of comprehensively accounting for market returns and volatility. This suggests that a blend of foundational factors and non-financial elements, such as investor sentiment, might better explain stock returns.

Investor sentiment holds critical implications for investment timing. It captures irrational anticipations about a stock's risk-return dynamics that aren't anchored in empirical evidence. Insights from behavioural economics affirm that investors' choices often deviate from rationality, swayed instead by emotions and sentiments (Tikkanen, 2021). Key questions arise: Are market fluctuations during the COVID-19 era driven by unwarranted investor anxieties? Did the pandemic trigger a genuine market downturn, or did it amplify irrational pessimism, especially regarding overvalued stocks?

Building on this, (Kaplanski and Levy, 2010) postulate that negative emotions, such as anxiety, can shape investment decisions. Anxious investors might harbour a gloomier outlook on future returns, leading them to adopt more conservative investment stances. This underlying anxiety fosters a negative sentiment, further influencing investment choices and subsequent asset performance.

The COVID-19 pandemic, accentuated by government pronouncements and actions, has undeniably eroded market confidence, precipitating a steep decline in UK stock prices. At its core, stock prices mirror prospective earnings. Given the pandemic's disruptive impact on economic activities, investors' concerns over future revenues have escalated, fuelling this decline (Liu et al., 2020). Investor sentiment, undeniably, is shaped by a multifaceted confluence of factors. Economic indicators, global events, sectoral performance, and media discourse all play a role in influencing investor behaviour. However, the unique context of the COVID-19 pandemic necessitates a deeper examination of governmental briefings. The COVID-19 pandemic's unprecedented nature and its widespread impact on health and the economy have been well-established (McKibbin and Fernando, 2021). In these challenging times, government briefings became indispensable sources of information. They served as the chief channels for conveying policies, public health directives, and updates on the pandemic's evolution (WHO, 2020). Far from being mere health updates, these briefings also outlined economic policies and interventions aimed at mitigating the pandemic's adverse effects (Takes, 2020). Consequently, they held a dual significance, shaping both public health responses and economic decision-making (Atkeson, 2020).

Several academic studies have highlighted the interplay between government announcements and financial market dynamics. (Baker et al., 2016) demonstrated that governmental policy uncertainty can directly influence firm-level decisions, affecting both investment and employment. Similarly, (Su et al., 2002) underscored the considerable influence of government interventions on stock returns and market volatility. In the specific context of the pandemic, (Baker et al., 2020a) noted that governmental responses to COVID-19 have been primary drivers of stock market reactions globally.

Moreover, the daily nature of COVID-19 briefings and their wide coverage in media made them a focal point of public attention. During the COVID-19 pandemic, investors have exhibited heightened sensitivity to daily updates and developments (Baker et al., 2020b). This increased sensitivity can be attributed to the investor sentiment that has been influenced by the pandemic. Investor sentiment, which is often reflected in trading volume, plays a significant role in stock price sensitivity patterns (Ernawati et al., 2022). The fear and uncertainty surrounding the pandemic have impacted both investors and individuals, leading to a dent in investor sentiment (Van Hoang and Syed, 2021).

One might wonder about the mechanics underlying the transmission of sentiment from government briefings to investor behaviour. A prominent explanation lies in the realm of herding behaviour. As posited by Banerjee (1992), investors often follow perceived prevailing trends or sentiments, especially in times of heightened uncertainty. When a significant segment of the investing community reacts to a stimulus, such as a government announcement, others might follow suit, irrespective of their personal beliefs or analyses. Furthermore, the Media Influence Theory suggests that news and media, including government announcements, can directly influence investor sentiment (Tetlock, 2007).

Amidst this, Information Theory comes into play. In an environment characterized by uncertainty and rapid changes, information becomes a premium asset. Government briefings, by providing timely and official updates, serve as crucial information sources. Investors, in their quest to make informed decisions, seek out these pieces of information and in doing so, may inadvertently amplify the sentiment embedded within them. The daily briefings, with their widespread coverage and analysis, become not just a focal point of investor attention but also a vital conduit for information dissemination. This interplay between herding tendencies, media influence, and the thirst for relevant information provides a plausible mechanism through which the sentiment of the government's daily briefings could be transferred to and mirrored by the investor community.

While myriad factors shape investor sentiment, the centrality of governmental briefings during the COVID-19 pandemic makes them a pivotal determinant of market dynamics. This study acknowledges the broader influences on investor sentiment but posits that, in the unique environment of the pandemic, government announcements held amplified significance. Building on this premise, the research delves deeper into the mechanics of sentiment derivation and its ramifications for the stock market.

This research extends the existing body of literature by exploring the potential impact of government announcements on investor behaviour and the consequent shifts in stock market dynamics amid the unique uncertainties of the COVID-19 era. A central focus is the sentiment

derived from the UK government's COVID-19 briefings.
The study is structured into two primary phases. The first phase, dubbed "sentiment extraction," aims to extract sentiment from the government's pandemic-related communications. Considering existing research that underscores the profound negative impacts of the pandemic on global financial markets (Al-Awadhi et al., 2020; Baker et al., 2020a; Zhang et al., 2020; Ramelli and Wagner, 2020a), we posit a cascading effect: the sentiment from governmental announcements might shape and sway the sentiment of UK investors. The second phase leverages statistical methodologies to analyse the relationship between this investor sentiment and the performance trends of principal UK stock indices, thereby shedding light on the causal dynamics. This exploration contributes to the broader academic dialogue on the influence of political news on stock market behaviours.
The advent of advancements in computational linguistics, Natural Language Processing (NLP), Machine Learning, and econometrics, enhanced by greater accessibility to diverse media content and digital discussions, has spurred their application in a range of financial research areas (Chouliaras, 2016).
Aligned with this trend, our approach harnesses NLP tools to distil sentiment from the widely circulated COVID-19 government announcements. We theorize that the sentiment identified from these official briefings reverberates within the sentiment landscape of the UK stock market, subsequently affecting stock prices and returns.
In synthesizing the interactions between the COVID-19 crisis and the UK stock market, this study endeavours to elucidate how investors responded to the UK government's communications during the pandemic. This initiative enriches the burgeoning body of work delving into the nexus between news propagation and stock market reactions.

2.2 Research motivation & contributions

The stock market's response to the COVID-19 pandemic has been unparalleled, surpassing the impacts of previous infectious disease outbreaks, including the Spanish Flu (Baker et al., 2020a). Given this unprecedented influence, a deeper exploration becomes imperative.
Research by (Khanthavit, 2020) underscores that the pronounced market reactions during the COVID-19 era were intimately tied to extensive media coverage and the official declaration of the pandemic. In the United States, initial market reactions in the late February to early March window predominantly mirrored news updates regarding the pandemic's progression (Baker et al., 2020a). Concurrently, studies by (He et al., 2020) revealed COVID-19's transient yet adverse effects on stock markets across several major countries. Reinforcing this, (Khanthavit, 2020) established a profound negative market response to the pandemic, attributing this largely to the heightened media focus rather than emergent events.
In a detailed analysis of 100,000 COVID-19-related news headlines and articles across four nations, (Ghasiya and Okamura, 2021) pinpointed the United Kingdom as harbouring the highest proportion of negative sentiment throughout an 11-month span of the pandemic. Parallelly, the findings of (Maligkris, 2017) suggest that optimistic political discourses bolster stock returns and trading volumes while curbing volatility. Conversely, pessimistic tones in political declarations induce the opposite effect.
The motivation for this research is rooted in the compelling evidence of the interplay between news, especially concerning COVID-19, and financial markets. Furthermore, the pronounced role of public announcements in modulating public sentiment—and by extension, investor sentiment—stands out. Through this investigation, our objective is to delineate the reactions of investors in the UK market to diverse government announcements made during the course of the COVID-19 crisis.
This endeavour presents a spectrum of novel contributions, ranging from broader insights to more nuanced, targeted findings in the fields we will detail subsequently.

2.2.1 Behavioural finance and market sentiment:

This research offers valuable insights into the fields of behavioural finance and market sentiment, particularly in the exploration of how investor sentiment influences the stock market during crisis situations like the COVID-19 pandemic. The current chapter builds upon the methodological foundations laid in the first chapter, where we employed the innovative "irrational fractional Brownian motion model" approach by (Dhesi et al., 2019) to empirically measure investor rationality/irrationality, a central concept in behavioural finance. In the present chapter, we extend this understanding of irrational investor behaviour by linking Dhesi and Ausloos (2016) 'psychological soliton' concept to market movement during times of crises. In their paper, the 'psychological soliton' represents the propagation of investor sentiment, reflecting how collective emotion or sentiment can spread through a population of traders and influence market behaviour.

The authors posit that this propagation of investor sentiment significantly affects financial markets, especially during crises. This connection provides an additional layer of understanding to the study of market sentiment, expanding its role in behavioural finance.

Our research situates itself in the rich tradition of behavioural finance and sentiment analysis, tracing back to seminal works like Keynes's notion of "animal spirits" (Keynes, 1936) and (Kahneman and Tversky, 1979), which both provided initial insight into sentiment-driven and bias-influenced behaviour in financial decision-making.

Moreover, Shiller's assertion (Shiller et al., 1984) that stock prices are driven by popular narratives and perceptions form a cornerstone of sentiment analysis in finance. Fast-forward to the 2000s, (Baker and Wurgler, 2006) in their paper "Investor Sentiment and the Cross-Section of Stock Returns", examined the effect of investor sentiment on stock returns and found evidence suggesting that investor sentiment does indeed play a significant role. And more recently, (Da et al., 2021) in their paper "Extrapolative Beliefs in the Cross-Section: What Can We Learn from the Crowds?", used a novel dataset from a popular online platform for individual investors and found that the crowd's extrapolative beliefs about stock returns predict cross-sectional stock returns. In essence, the study provides empirical evidence that investor sentiment, especially among non-professional investors, can influence stock prices. It confirms the theoretical notion of "extrapolative beliefs" and shows they can be used to predict future stock returns, particularly in the short term.

These findings align with the theories put forward by scholars like Robert Shiller who argue that investor sentiment and popular narratives can impact financial markets. It would also challenge traditional finance theories like the Efficient Market Hypothesis, which holds that markets are perfectly efficient and that all public information is already reflected in current prices.

By applying these foundational theories, this study enhances the understanding of the empirical relationship between investor sentiment and market volatility during crises. Thus, we aim to enrich the knowledge base in behavioural finance and sentiment analysis, contributing to the ongoing conversation about the integration of investor sentiment and market volatility.

2.2.2 Governmental interventions, political news and financial markets

Our research contributes uniquely to the field of governmental interventions and financial markets by examining the sentiment conveyed in UK government briefings during the COVID-19 pandemic and its immediate impact on the FTSE stock market. This approach provides a distinctive perspective by considering the sentiment expressed in these communications, offering a novel understanding of how sentiments during governmental briefings can significantly sway financial markets during pandemics. Our focus on sentiment analysis aligns this study with existing literature in behavioural finance, integrating it with a political economy perspective. Most current research, including the work of (Pástor and Veronesi, 2013), and (Ramelli and Wagner, 2020a), predominantly focuses on broader, long-term impacts of events and government policies on financial markets, albeit through different lenses: one through a theoretical model of political risk, the other through an empirical analysis of a specific historical event. In contrast, our study explores the immediate effects of sentiment expressed during government briefings, augmenting the current understanding of the role governmental communications play in financial market behaviour.

By incorporating sentiment analysis into the investigation of governmental briefings, our research extends the discourse on government interventions and financial markets. It highlights the importance of considering sentiment in government communication, particularly in times of crisis, and its immediate impact on the market. This novel integration enriches the field and invites further research into the role of sentiment in political communications and their economic implications.

In the context of political news, the research by (Tetlock, 2007), which delved into the role of media sentiment in influencing the stock market, serves as a vital point of reference. Tetlock found that negative words in the media could predict downward pressure on market prices. Our research builds upon and refines Tetlock's line of inquiry by focusing specifically on a single, yet significant, type of media – governmental briefings. In the broader context of Tetlock's study, our research provides a micro-level examination of how sentiment expressed in a particular form of political news (i.e., daily governmental briefings) can immediately influence financial market movements. Tetlock's study surveyed the broader landscape of media sentiment and its short-term effects on market prices. In contrast, our study provides a deeper dive into the immediate influence of sentiment expressed in government briefings, specifically in times of crisis. This specificity enriches the existing body of literature by highlighting how real-time political communications can have a unique and immediate impact on market behaviour.

2.2.3 Application of NLP in financial research

Our research represents a pivotal development in leveraging Natural Language Processing (NLP) tools, especially the Bidirectional Encoder Representations from Transformers (BERT) model (Devlin et al., 2018), in the realm of financial research. The integration of sentiment analysis in finance has its roots in seminal works like that of (Antweiler and Frank, 2004), who employed text analysis to decipher the correlation between online stock message board discussions and stock market volatility. Nevertheless, the infusion of advanced NLP techniques such as BERT in finance is a relatively recent evolution, considering BERT was officially introduced in 2018. Central to our study is the seminal works of (Sousa et al., 2019) and (Ghasiya and Okamura, 2021). In their study, Sousa and his team utilized BERT for sentiment analysis of financial news from reputed media outlets and used this sentiment as an indicator of daily fluctuations

in the Dow Jones Industrial Average (DJIA). Despite achieving superior performance with BERT over traditional methods, they identified a challenge – the complexity and noisiness of the data when trying to link sentiment analysis results directly with the DJIA's performance. Their study provided a broader overview of market trends but did not specifically focus on the immediate impact of individual news items on market movements.
In contrast, our research goes a step further by employing BERT to investigate the direct and immediate impact of daily governmental briefings, a unique and relatively unexplored source of data, on the FTSE stock market. Additionally, while (Sousa et al., 2019) study was set in conventional financial contexts, ours delves into the exceptional conditions surrounding the COVID-19 pandemic, hence offering valuable insights into market dynamics during crisis situations. Against the backdrop of these advancements, (Ghasiya and Okamura, 2021) spotlighted the influence of media sentiments during the COVID-19 pandemic, providing an essential foundation for our research. Drawing inspiration from their findings and the NLP advancements, our study aims to refine and expand upon this by deploying BERT to assess the immediate repercussions of daily governmental briefings on the FTSE stock market. Our research introduces a methodological nuance by employing the BERT model. While Ghasiya and Okamura utilized the larger RoBERTa model (Liu et al., 2019), our choice of BERT offers advantages in terms of computational efficiency without compromising on analytical depth (Devlin et al., 2018). Our approach, informed by(Ghasiya and Okamura, 2021) insights, and distinct from (Sousa et al., 2019) methodology leveraged a pre-trained BERT model. Using a pre-established model brings with it the advantage of leveraging vast amounts of data and training that BERT was initially exposed to. Furthermore, unlike (Sousa et al., 2019), who offered a broader perspective on market trends, we zoomed in on the immediate implications of daily governmental briefings on the FTSE stock market amidst the challenges of the COVID-19 pandemic. This nuanced focus, combined with our specialized NLP approach, not only accentuates the versatility of NLP tools in diverse contexts but also underscores the potential of BERT in contemporary financial analyses. Therefore, our study not only extends the current understanding of BERT's application in financial research but also opens new pathways for future investigations into crisis communication, sentiment analysis, and immediate market responses.

2.2.4 Granular contributions

The significant influence of news on financial markets, especially during pivotal global events like the COVID-19 pandemic, has been well-documented. (Khanthavit, 2020) and (He et al., 2020) have demonstrated the intertwined nature of media coverage and market reactions during the COVID-19 timeline. This sentiment aligns with findings by (Baker et al., 2020b), who examined stock price crashes during the onset of the pandemic. Furthermore, (Albulescu, 2021) delved into the effects of official announcements about COVID-19 infection rates and fatality ratios on the financial market volatility in the United States. Albulescu's empirical findings

underscore the fact that the sanitary crisis significantly amplifies the realized volatility of the S&P 500. This indicates that the persistence of the pandemic plays a pivotal role in financial volatility, posing challenges to risk management activities.
However, while Albulescu's research hones in on the correlation between official health announcements and market volatility in the U.S., our study expands this scope by employing advanced NLP techniques, specifically BERT, to analyse the impact of government announcements in the UK on the FTSE stock market. This distinction not only situates our research within a unique geographical and methodological context but also showcases the versatility of NLP in providing nuanced insights into global financial market dynamics during crisis situations.
The UK's scenario is distinct. While (Ghasiya and Okamura, 2021) provided insights into the negative sentiment predominating UK headlines, other research, such as that by (Ramelli and Wagner, 2020b), explored the global economic consequences of the pandemic, highlighting the UK's unique position.
Given the complex dynamics at play during the COVID-19 pandemic, the role of government announcements becomes paramount in shaping economic landscapes (Baldwin and Di Mauro, 2020). Government announcements often serve as pivotal information sources for investors, guiding them in an environment marked by uncertainty. Historically, government communications, especially during crises, have significantly influenced financial markets ((Baker et al., 2016); (Smales, 2015)). For instance, during the 2008 financial crisis, public addresses by key officials were pivotal in shaping investor sentiment and market reactions (Beetsma et al., 2012). However, the COVID-19 pandemic presented a unique challenge. Unlike financial crises, where economic indicators are primary, the pandemic intertwined public health with economics, making government announcements a critical barometer for both health updates and economic policies ((Takes, 2020); (McKibbin and Fernando, 2021)). Studies such as by (Baker et al., 2020b) have shown that the economic implications of the pandemic have been intricately linked to health-related news and policy announcements.
Yet, the specific influence of government announcements related to COVID-19 on investor behaviour, particularly in the UK, remains an under-explored niche. This research seeks to fill this gap, offering a novel perspective in a domain that demands deeper understanding. Investigating the UK context is especially valuable given its dual role as a major financial hub and one of the countries severely affected by the pandemic ((Kalaitzake, 2021), (Jebril, 2020).

2.2.4.1 Novelty in Analysing COVID-19 Government Announcements:

While numerous studies have analysed the influence of political speeches on various societal dimensions, including financial markets, the specific impact of daily governmental briefings, especially during a public health crisis, remains less explored. Historically, political speeches have been known to sway public opinion, influence voter behaviour, and even affect financial markets. For instance, studies such as (Snowberg et al., 2007), (Cinelli et al., 2021), (Herron, 2000) have

investigated how political events and announcements can cause stock market fluctuations.
However, the COVID-19 pandemic introduced an unprecedented scenario: regular daily briefings by governments worldwide. Unlike the periodic nature of traditional political speeches, these daily communications became a consistent and primary source of information for millions, setting them apart in terms of frequency, reach, and potential impact. Works by Khanthavit et al. (Khanthavit, 2020), He et al. (He et al., 2020), and Alfaro et al. (Alfaro et al., 2020) have explored the broader media landscape and its relationship with market dynamics during the pandemic. Still, a deep dive into the specific influence of these daily briefings remains sparse.
These daily briefings, delivered by heads of state or top health officials, didn't just communicate statistics or guidelines; they were instrumental in framing official narratives, setting public sentiment, and potentially influencing financial markets. The study by (He et al., 2020) further underscores the swift and pronounced negative effects of the COVID-19 outbreak on stock market returns, especially in Asian countries. This research suggests that investor sentiment, driven by fears and uncertainties surrounding the pandemic, acts as a critical transmission channel for these market effects. Importantly, it emphasizes the role of announcements by health officials, indicating that their emotional and psychological impact plays a vital role in shaping financial market reactions. While (Williams and Wright, 2022) delved into the linguistic strategies of the UK's daily COVID-19 briefings and their implications for governmental responsibility, the direct correlation between these briefings and financial market reactions is an area that warrants further exploration.
This research endeavours to bridge this gap, probing how daily official communications, with their inherent linguistic subtleties and emotional undertones shaped by the unique nature of the COVID-19 crisis, can influence market behaviour over an extended period. By focusing on these official briefings and intertwining linguistic analysis with financial outcomes, this study aims to unravel the intricate role of governmental communication in either stabilizing or destabilizing financial markets during an extended crisis period.

2.2.4.2 Temporal Depth of Analysis:

In the initial wave of financial research on the COVID-19 pandemic, the temporal depth of analysis might have been overshadowed. While many studies ((Mirza et al., 2020), (Fahlenbrach et al., 2021), (He et al., 2020)) have focused on the immediate market reactions to the event, there's a clear need to also explore the longer-term effects of the pandemic on financial markets. This research rises to the challenge by offering a comprehensive 16-month analysis, moving beyond the immediate aftermath of the pandemic in 2020. This extended timeline is not merely an analytical choice but a necessity. As (Eachempati et al., 2021) highlighted, market news doesn't swiftly absorb and mirror rapid changes in stock markets. Instead, it necessitates a more extended period for the absorption and reflection of evolving sentiment. Their sentiment analysis, leveraging machine learning on Twitter data, underscored the delayed response of

global markets until the WHO's official declaration of the pandemic, suggesting a lag in market sentiment adjustments.
By expanding its analysis to encompass data up to June 2021, this research captures the nuanced oscillations of market behaviour over a broader timeline. This approach provides invaluable insights into market adaptability, resilience, and the extended influences of significant global events. Furthermore, by moving beyond the immediate year of the pandemic's outbreak, this study stands as a testament to the importance of understanding market behaviour in the long run, ensuring that conclusions drawn are both holistic and reflective of the market's true adaptive nature.

2.2.4.3 Sentiment Analysis of Government Announcements:

Building on the work of Tikkanen (Tikkanen, 2021), who studied the FTSE All Shares Index's daily close-to-close price and found that sentiment data could enhance a model's stock market direction forecasting performance, our research delves further into the complex relationship between governmental announcements and financial market reactions during the COVID-19 pandemic. Ghasiya and Okamura (Ghasiya and Okamura, 2021) underscored a prevailing negative sentiment in COVID-19 news headlines, with a particular emphasis on the notably negative media landscape in the UK. Their findings serve as a cornerstone for our specialized exploration, wherein we focus on the sentiment inherent in the UK government's specific COVID-19 announcements. Ghasiya and Okamura embarked on a large-scale exploration, harnessing a dataset of over 100,000 news articles from four distinct nations. Their comprehensive approach aimed to unveil the overarching sentiments and prevailing topics globally during the pandemic. They found that the UK, among the countries studied, exhibited the most negative sentiment in its news coverage, a finding that was congruent with its severe pandemic impact. Their research, while expansive in scope, primarily centred on general news sentiment, offering a macroscopic view of media narratives. In stark contrast, the current research narrows its lens to the UK's governmental announcements pertaining to COVID-19 and their consequential sway on investor behaviour. This provides a granular insight into the immediate and evolving impact of official communications on financial markets. Furthermore, the thematic essence of the two studies diverges. While Ghasiya and Okamura delineate the sentiments across various sectors, including education and sports, the present research is rooted in financial implications, resonating with findings by (Ramelli and Wagner, 2020b) that emphasize the UK's unique economic position during the pandemic. This is further enriched by the study's temporal depth, spanning a holistic 16-month view, a feature not extensively emphasized by Ghasiya and Okamura. Ghasiya and Okamura (2021) mapped out the broader media landscape during the COVID-19 pandemic, drawing attention to the pronounced negative sentiments in UK headlines. Building on this foundational understanding, our research extends into the financial sphere, particularly focusing on how these sentiments influence the financial markets. (Hassan et al., 2020) delved deep into the textual content of listed firms, constructing measures that identify firms' primary concerns linked to the spread of COVID-19 and other epidemic diseases. Their study discerned the firms that perceived gains or losses from the pandemic and through textual decomposition, illustrated how epidemics affected both demand and supply for these firms. Notably, they identified that the effects of COVID-19 manifested as simultaneous shocks to both demand and supply, with these shocks bearing equal weight on firms' market valuations. However, it was the demand-related impacts that played a pivotal role in the observed collapse of firm-level investments during the crisis. Marrying the insights of Ghasiya and Okamura's broad media analysis with Hassan et al.'s nuanced financial textual exploration, our study seeks to provide a comprehensive understanding of the interplay between governmental announcements, media sentiment, and their consequential effects on investor behaviour and market dynamics.

In this research, we aim to understand how specific COVID-19-related government announcements influenced investor behaviour in the UK. This adds to the existing literature on the relationship between public policy, health crises, and financial market movements. Building on this foundation, we use advanced techniques to study the potential effects of the UK government's announcements on the stock market. As noted by (Liu et al., 2020), the language used in public announcements by health officials can have a significant impact on people's emotions and perceptions.

For clarity, the data we use, which includes UK Government communications and FTSE stock indices (detailed in Section 2.5), focuses only on the initial statements of each COVID-19 briefing, excluding the Q&A section. We believe the initial statements are a better representation of the government's official stance, as the Q&A, influenced by journalists' questions, might not always align with the primary message.

We also study the relationship between the stock market and investor sentiment, using correlation and Granger causality analyses across different time frames. As (Mudinas et al., 2019) highlighted, there's a link between sentiment and stock price changes. Furthermore, based on insights from (Huberman and Regev, 2001), we look into the potential delayed effects of government briefings on the UK stock market.

By examining the relationship between COVID-19 and the UK stock market, we hope to provide valuable insights into how investors reacted to UK government announcements during the outbreak. This research contributes to the broader discussion on the influence of news and public statements on stock markets. For further context, Table 2.1 lists related studies in this area. The following sections will delve into the specific research questions of this study.

2.3 Research questions

In the intricate web of market dynamics during a global pandemic, the flow of information and its subsequent influence on investor behaviour assumes paramount importance, such a point is based on the widely accepted principle that information flow is critical in financial markets, especially during times of uncertainty like a global pandemic. This idea is rooted in the Efficient Market Hypothesis (EMH), which posits that stock prices reflect all available information Fama (1970). This study sits at the intersection of financial theory and information dissemination, striving to uncover the ripple effects of governmental announcements on stock market performance during the COVID-19 crisis.

Figure 2.2 provides a schematic of the research questions driving this investigation. At its core, the research seeks to understand the relationship between the sentiment of UK government updates—both positive and negative—during COVID-19 and the subsequent market returns, both positive and negative. The foundational premise is that information conveyed through governmental briefings can resonate deeply with investors, potentially influencing their perceptions and actions or at least relating to these.

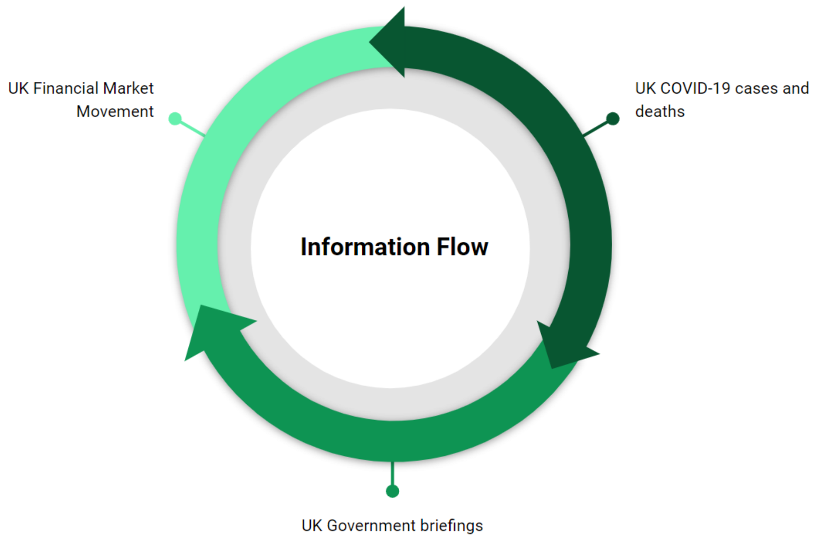


Figure 2.2: Information circulation

The first layer of this inquiry, represented by Research Question 1 (presented below), delves into the correlation between daily COVID-19 cases and fatalities in the UK and the prevailing sentiment of the associated government announcements. This phase aims to juxtapose the emotional tone of the briefings against the stark realities of the pandemic, establishing a baseline for the information flow. The underlying assumption is straightforward: as the pandemic's gravity escalates, it should ostensibly be mirrored in the tenor of governmental communications. With the groundwork laid, Research Question 2 ventures deeper, exploring the potential cascading effects of these government briefings on the stock market. By correlating the sentiment

of these announcements with market returns, this phase seeks to unravel the intricate dance between information dissemination and investor sentiment. It postulates that these briefings, rife with pandemic-related updates, could sway investor moods, subsequently shaping their perceptions about future business prospects and influencing stock market trajectories.
Yet, in the realm of scientific inquiry, it's paramount to isolate variables and ensure that observed relationships are not mere artefacts of confounding factors. This brings us to Research Question 3, designed as a control mechanism. While the first two questions delve into the role of sentiment, this phase investigates the direct correlation between pandemic severity (cases and fatalities) and stock market performance. The objective is clear: to ascertain if the relationships unearthed in the first two questions are genuine, or if they could merely be the by-products of the pandemic's overarching impact on the market.
We will now formally present the research questions in the following format:

- **Question 1 (RQ1)**: How does the severity of the COVID-19 pandemic, as measured by daily cases and fatalities in the UK, impact the sentiment of associated government announcements?
  This question seeks to establish a foundational relationship between the real-world implications of the pandemic and the emotional undertones of governmental communications. It operates on the premise that the gravity of the pandemic situation should be reflected in the sentiment of the briefings, providing the initial layer of information flow.

- **Question 2 (RQ2)**: To what extent does the sentiment of UK government announcements regarding COVID-19 relate to stock market prices and returns?
  Building on the baseline established in RQ1, this question delves into the potential ripple effects of the government's informational broadcasts on investor behaviour. Through correlation and Granger causality analyses, the research aims to uncover the dynamics of information dissemination and its subsequent impact on investor perceptions and market trajectories.

- **Question 3 (RQ3)**: Is there a direct impact of the observable severity of the COVID-19 pandemic (cases and fatalities) on stock market performance, independent of the sentiment from governmental announcements?
  Designed as a control mechanism, RQ3 aims to isolate the direct effects of the pandemic on the stock market from those potentially influenced by government communications. Through both correlation and Granger causality analyses, the research seeks to discern whether the relationships observed in RQ1 and RQ2 are genuine or if they might be overshadowed by the broader implications of the pandemic's progression.

In synthesising the results from these inquiries, the study not only contributes to the rich tapestry of financial literature but also sheds light on the profound power of information in shaping market destinies during global crises.

The remainder of the chapter is structured as follows. In Section 2.4, we examine the relationship between COVID-19 and the stock market, as well as the influence of news and public announcements on investor mood. In Section 2.5, we detail the data sources and collecting methodology, as well as the preparation of the conference briefings' content. The gathering and transformation of market data are described in detail, as are their descriptive statistics. In Section 2.6, we give the approach for correlation analysis, which is supplemented with Granger causality tests. In Section 2.7, we describe the thorough tests conducted to determine if the COVID-19 UK government updates affected the UK stock market. Section 2.8 concludes with closing comments and a discussion of future prospects.

2.4 Related Studies

This section compiles published works and relevant research that have already been conducted on the topic we are investigating. To begin, we review existing literature on how the dissemination of news affects the stock market. This is followed by a lengthy body of material analysing the financial impact of the coronavirus (COVID-19) epidemic over the last two years. Next, we take a look at research that used Machine Learning/Sentiment analysis to understand the economy and the markets. To wrap up, we draw attention to a select number of studies that examined the relationship between sentiment analysis and the economic fallout from the COVID-19 epidemic.

2.4.1 Public news impact on financial markets

The influence of news on financial markets has been extensively studied. First, to analyse the impact of news on stock prices, Niederhoffer's (1971) research relied on New York Times front pages(Niederhoffer, 1971). (Schumaker and Maida, 2018), investigated stock price movements before and after the release of financial news articles using sentiment polarity and correlation analysis for the S&P500 index. This study was able to identify a strong correlation between the period following the release of a financial news article and abnormal price movement. (Chan, 2003), investigates stock response to news and no news by analysing headlines regarding specific firms. The authors discovered that coherence in financial news is substantially connected with and driven by volatility in financial markets, using Drift and reversal following headlines.
(Jazbec et al., 2021), analyses the spread and absorption of large-scale publicly accessible news articles from the Internet to financial markets. The authors calculated information transmission entropy and sentiment extraction by screening and topic modelling from public news items to the US stock market. The authors discovered that, when compared to non-public and commercial news data sources, the sentiment signals from public news have both economic worth and complementing information when utilising a basic sentiment-based trading strategy as an econometric instrument. (Piškorec et al., 2014), proposes a measure of collective behaviour based on financial news on the web – the News Cohesiveness Index (NCI) using NCI to measure the cohesiveness in the news by calculating the average similarity in the financial news and Granger causality to test if NCI-financial is related to the volatility of the market.
(Stankevičienė and Akelaitis, 2014), used an event study analysis to explore the influence of public announcements on stock prices, looking into the relationship between Stock Price Values and Price Changes in the Lithuanian Stock Market. The investigation based on a comparison of mean price movements in various stock price ranges in the Lithuanian stock market revealed that the stocks with the lowest prices had the greatest distortion of results. (Maligkris, 2017), investigates if presidential candidates' political remarks impact stock market results throughout their campaigns. Loughran and McDonald were used to detect the linguistic tone, and regression analysis was utilised to analyse the impact. The authors discovered that economic information

in political speeches increases aggregate market returns and trading volume while decreasing market volatility.

(Atkins et al., 2018), compared predictability of volatility using Financial news and close price in the US stock market. The authors used Machine Learning models of Latent Dirichlet Allocation to represent information from news feeds, and simple naïve Bayes classifiers to predict the direction of movements. The authors concluded that information extracted from news sources is better at predicting the direction of underlying asset volatility movement, or its second-order statistics, rather than its direction of price movement. They conclude that volatility movements are more predictable than asset price movements when using financial news as Machine Learning input, and hence could potentially be exploited in pricing derivatives contracts via quantifying volatility.

(Nam and Seong, 2019), proposes a novel Machine Learning model to forecast stock price movement based on the financial news, considering causality. Using Transfer entropy to find causality and multiple kernel learning is used to combine features of target firms and causal firms in the Korean market.

The authors' experimental findings demonstrate that the suggested technique can forecast the stock price directional movements even when there is no financial news on the target business, but financial news is released on causal companies. Their results suggest that recognising causal links is key in prediction problems.

(Alanyali et al., 2013), quantifies the relationship between decisions taken in financial markets and developments in financial news. Using correlation analysis and a focus on four types of time series for each of the 31 Dow Jones Industrial Average (DJIA) companies: the daily number of mentions of a company's name in the Financial Times, the daily transaction volume of a company's stock, the daily absolute return of a company's stock, and the daily return of a company's stock. The authors discovered a positive relationship between the number of times a firm was mentioned in the Financial Times on a given day and the number of shares traded that day and the day before.

A summary of the aforementioned studies is shown in Table 2.1 along with a summary of the paper and a description of the used approaches.

Authors	Description	Approach
(Schumaker and Maida, 2018)	Investigating stock price movements before and after the release of financial news articles using	Sentiment polarity and correlation analysis using for S&P500
(Chan, 2003)	Examines stock reaction to news and no news using headlines about individual companies	Drift and reversal after headlines
(Jazbec et al., 2021)	Quantifies the propagation and absorption of large-scale publicly available news articles from the World Wide Web to financial markets	Quantify the information transfer from public news articles to the U.S. stock market through Transfer entropy and sentiment Extraction via Screening and Topic Modelling
(Piškorec et al., 2014)	Proposes a measure of collective behaviour based on financial news on the web – the News Cohesiveness Index (NCI)	NCI to measure the cohesiveness in the news by calculating the average similarity in the financial news and Granger causality to test if NCI-financial is related to the volatility of the market
(Stankevičienė and Akelaitis, 2014)	Impact of Public Announcements on Stock Prices: Relation between Values of Stock Prices and the Price Changes in Lithuanian Stock Market	Event study analysis
(Maligkris, 2017)	Examines whether political speeches of presidential candidates influence stock market outcomes during their campaigns	Loughran and McDonald to identify the linguistic tone and regression analysis
(Atkins et al., 2018)	Comparing predictability of volatility using Financial news and close price in the US stock market	Machine Learning models of Latent Dirichlet Allocation to represent information from news feeds, and simple naïve Bayes classifiers to predict the direction of movements
(Nam and Seong, 2019)	Proposes a novel Machine Learning model to forecast stock price movement based on the financial news considering causality in the Korean market	Transfer entropy is used to find causality and multiple kernel learning is used to combine features of target firm and causal firms.
(Alanyali et al., 2013)	Quantifies the relationship between decisions taken in financial markets and developments in financial news	Correlation analysis and a focus on four types of time series for each of the 31 Dow Jones Industrial Average (DJIA) companies.

Table 2.1: News and financial markets related work – List of references about News and financial markets related work. Each row represents a study, starting from the left side; in the first column there are the references; in the second the description of the respective papers; and in the third column is the approach used.

2.4.2 COVID-19 and the financial market

The academic community has been responding to the COVID-19 crisis for the last two years. The effects of the COVID-19 pandemic on the stock market have been the topic of previous research, utilising a number of different theoretical paradigms. Examining the pandemic's effect through the number of cases and fatalities, via financial news and government policies, Google searches, and social media platforms were the frameworks that stood out in the literature. An overview of prior research is provided below, organised by subject themes. Tables 2.2, 2.3, 2.4 summarises the previous literature related to COVID-19 and financial markets.

COVID-19 and the stock market

(Sharif et al., 2020), studied the connectedness between the spread of COVID-19, oil price volatility shock, the stock market, geopolitical risk and economic policy uncertainty in the US within a time-frequency framework. The author(s) used the coherence wavelet method and the wavelet-based Granger causality tests applied to US daily data. The study found that the COVID-19 outbreak has a higher impact on geopolitical risk and economic uncertainty in the United States than on the stock market.

(Shehzad et al., 2020), analysed the non-linear behaviour of financial markets of the US, Germany, Italy, Japan, and China during the COVID-19 and Global Financial Crises period using Asymmetric Power Generalized Autoregressive Conditional Heteroscedasticity model. COVID-19, according to (Shehzad et al., 2020), has a significant negative impact on market returns in the United States and Japan. Furthermore, COVID-19 has had a greater impact on the volatility of the stock markets in the United States, Germany, and Italy than the Global Financial Crisis (GFC).

(So et al., 2021), studied the impacts of the COVID-19 pandemic on the connectedness of the Hong Kong financial market and compared the impacts of the previous financial crises in the past 15 years. The author(s) constructed dynamic financial Networks based on correlations and partial correlations of stock returns. (So et al., 2021) found that in the partial correlation Networks during the COVID-19 epidemic, Network density and clustering are greater than in prior crises when Network density and clustering may be explained by co-movement with market indices as in normal times.

(Şenol and Zeren, 2020), examined the effect of the COVID-19 outbreak on global markets between January 21, 2020 and April 7, 2020 using the Fourier Co-integration test. The study showed that stock markets reacted promptly to the risks posed by COVID-19, and stock market indexes depreciated rapidly.

(Zhang et al., 2020), maps the general patterns of country-specific risks and systemic risks in the global financial markets Correlation analysis and minimum spanning tree. According to the author, global financial market risks have escalated significantly in reaction to the epidemic. Individual stock market movements are obviously connected to the magnitude of each country's

epidemic.
(Mirza et al., 2020), assess the price reaction, performance, and volatility timing of European investment funds during the outbreak of COVID-19 using a GARCH-based event study. The author(s) demonstrated that market liquidity has decreased, which has harmed market returns.
(Al-Awadhi et al., 2020), investigated the COVID-19 pandemic effect on the Chinese stock market using panel data regression. The author(s) found that The medical and pharmaceutical industries had positive abnormal returns, whereas restaurants, hotels, and motels had negative abnormal returns during the COVID-19 pandemic.
(Griffith et al., 2020), described how the impact of COVID-19 has varied across industries, using data on share prices of firms listed on the London Stock Exchange.
(Mazur et al., 2021), investigate the US stock market performance during the crash of March 2020 triggered by COVID-19. In this study, the author(s), analysed single-day extreme events by sector for industry returns and volatility. The study found that stocks in the healthcare, food, natural gas, and software sectors performed exceptionally well during the March 2020 stock market crash, generating high returns, whereas firms in the crude petroleum, real estate, entertainment, and hospitality sectors plummeted significantly, losing more than 70% of their market capitalisation.
(Takyi and Bentum-Ennin, 2021), evaluated and quantified the short-term impact of the coronavirus disease of 2019 on stock market performance in thirteen African countries, using daily time series stock market data and Bayesian structural time series. According to Bayesian posterior estimates, the author(s), demonstrated that stock market performance in Africa has been considerably lowered during and after the occurrence of COVID-19, typically by -2.7% to -21%.
(Samitas et al., 2022), identifies volatility and contagion risk among stock markets during the COVID-19 pandemic using dependence dynamics and Network analysis on a bi-variate basis. The study demonstrates "instant financial contagion" as a consequence of the shutdown and the transmission of the new coronavirus.

COVID-19 – impact of cases and fatalities

(Giangreco, 2020), investigated the correlation between case fatality rate and social behaviour during the Italian COVID-19 outbreak, using Correlation analysis of case fatality rate compared with social habit variables. According to an examination of all Italian areas, the fatalities are confined and have no correlation with any social conduct. (Sansa, 2020), investigates the relationship between the COVID-19 confirmed cases and both China and USA stock markets using descriptive statistics and regression models. The research results demonstrated a positive substantial association between the COVID-19 verified instances and all financial markets (Shanghai Stock Exchange and New York Dow Jones) in China and the United States from March 1st to March 25th, 2020.

(Onali, 2020), investigated the impact of COVID-19 cases and related deaths on the US stock market (Dow Jones and S&P500 indices) using Generalized autoregressive conditional heteroscedasticity (GARCH) and Vector autoregression (VAR) models. The author(s) indicated that variations in the number of cases and fatalities in the United States and six other countries heavily impacted by the COVID-19 problem had little effect on US stock market returns, except the number of reported cases in China. However, there is evidence of a beneficial influence on the conditional heteroscedasticity of the Dow Jones and S&P500 returns for certain nations.
(Albulescu, 2021), empirically investigates the effect of the official announcements regarding the COVID-19 new cases of infection and fatality ratio, on the financial market's volatility in the United States using ordinary Least Squares (OLS) regression to test both global and US reported data. This paper found that the new infection cases reported at the global level and in the US amplify the financial volatility.

COVID-19 – impact of news, social media, and Google searches
(Goodell and Huynh, 2020), assesses the reactions of US industries to sudden COVID-related news announcements, concomitantly with an analysis of levels of investor attention to COVID-19. Using Capital asset pricing (CAPM) model and ordinary least squares (OLS).
(Cepoi, 2020), found a piece of novel empirical evidence on the relationship between COVID-19-related news and stock market returns across the top six most affected countries by the pandemic, using the Panel quantile regression framework. The study demonstrates that stock markets exhibit asymmetric reliance on COVID-19-related information such as fake news, media coverage, or contagion. The outcome implies that suitable communication channels should be used more intensively to reduce COVID-19-related financial turmoil.
(Haroon and Rizvi, 2020), explored whether the media reporting of COVID-19, panic amongst investors, and global sentiment have played a role in the previously unseen volatility in the equity markets using exponential GARCH models. The authors found that the excessive fear caused by news sources has been linked to increased volatility in the financial markets. The authors' findings for specific economic sectors show that panic-laden news contributed more to volatility in sectors believed to be most impacted by coronavirus outbreaks.
(Khanthavit, 2020), investigated how and how early the world and national markets reacted to COVID-19 events and news coverage using an Event study and Abnormal returns analysis. The study concludes that the responses were to COVID-19's enormous media attention and pandemic announcement, not to the real unfolding events and conditions.
(Baker et al., 2020a), examined the role of COVID-19 developments in recent stock market behaviour and drew comparisons to previous infectious disease outbreaks, In their methodology, the authors, examined next-day newspaper explanations for each daily move in the US stock market. The evidence the author(s) have gathered suggests that government restrictions on commercial activity and voluntary social distancing, both of which have powerful effects in a

service-oriented economy, are the primary reasons the US stock market reacted so strongly to COVID-19 than to previous pandemics in 1918-1919, 1957-1958, and 1968.

(Cinelli et al., 2020), provided an in-depth analysis of the social dynamics in a time window where narratives and moods in social media related to COVID-19 have emerged and spread. The research showed that the interaction paradigm imposed by the unique social media or/and the specific interaction patterns of groups of users involved with the issue drives knowledge dissemination.

(Cerqueti and Ficcadenti, 2020), investigates the relationship between the Google search volumes of "coronavirus" and the stock index prices of different markets, using design indicators able to capture the connection between anxiety for the pandemic and expectations on the future outcomes of financial markets. The authors found that anxiety is manifested via the intensity of the searches run on Google related to the COVID-19 virus.

(Topcu and Gulal, 2020), explored the predictive power of Google searches on stock market volatility during the COVID-19 pandemic using regression analysis and Driscoll-Kraay estimator. The author(s) found that the COVID-19 outbreak's effect is considerably lower in developing countries where governments implemented necessary actions on time and announced greater stimulus packages.

(Lyócsa et al., 2020), explores the predictive power of Google searches on stock market volatility during the COVID-19 pandemic using trends analysis and ordinary least squares. The author(s) reveal that Google searches for coronavirus are not just correlated; they forecast variation in the future for every nation studied.

2.4.3 Sentiment Analysis and Financial markets

Investment banks and hedge funds are among the many financial firms focusing on investor sentiment in order to refine their market predictions. According to (Mudinas et al., 2019), two of the most well-known companies in the financial news and data industry, Thomson Reuters and Bloomberg, have lately extended their services to include company sentiment research. The following is a non-exhaustive collection of research into the use of sentiment analysis in the business world. The results of this literature synthesis are summarised in Table 2.5.

(Zaenen and van den Bosch, 2007), explored a computable metric of positive or negative polarity in financial news text, which is consistent with human judgments and can be used in a quantitative analysis of news sentiment impact on financial markets. The author(s) elaborated a Lexical cohesion-based metric of sentiment intensity and polarity in text and evaluated this metric relative to human judgments of polarity in financial news. (Mao et al., 2011), compared a range of different online sources of information (Twitter feeds, news headlines and volumes of Google search queries) using sentiment-tracking methods and compared their values for financial prediction of market indices, such as the DJIA, trading volumes, implied market volatility (VIX) and gold prices.

The authors' findings suggest that typical investor intelligence surveys are lagging indicators of the financial markets. Weekly Google Insight Search volumes on financial search inquiries, on the other hand, have predictive value. The frequency of occurrence of financial keywords on Twitter in the preceding 1-2 days, as well as an indicator of Twitter Investor Sentiment, are proven to be extremely statistically significant predictors of daily market log return.
(Casarin and Squazzoni, 2013), computed the Bad News Index as the weighted average of negative sentiment words in the headlines of three distinct news sources. Their results showed that the press and markets influenced each other in generating market volatility. Their findings support the reflexive character of stock markets. When the situation is ambiguous and unexpected, market behaviour may even reflect qualitative, large picture, and subjective information, such as streamers in a newspaper, the economic and informational worth of which is debatable.
(Ranco et al., 2015), explored the effects of Twitter Sentiment on Stock Price Returns for companies that form the Dow Jones Industrial Average (DJIA) index using an Event study to identify events as Twitter volume peaks, Support Vector Machine (SVM) for sentiment extraction, Correlation analysis, and Granger causality tests.
All things considered, the authors discovered a low Pearson correlation and Granger causality between the two-time series in question. They do, however, discover a robust correlation between Twitter sentiment and abnormal returns across Twitter's busiest times. The authors use an event analysis to demonstrate that the direction of cumulative anomalous returns is mirrored by the polarity of sentiment during Twitter peaks. While the total aberrant returns are small (about 1-2%), the correlation remains statistically significant for many days after the occurrences.
(Kalyani et al., 2016), examined non-quantifiable data such as financial news articles about a company and predicted its future stock trend with news sentiment classification using Support Vector Machine (SVM), Random forests, and Naïve Bayes classification. In their study, the authors found that stock trends can be predicted using news articles and previous price history.
More recently, (Mudinas et al., 2019), investigated the causal relationship between sentiment attitude/emotion signals and stock price movements through financial news using various sentiment signal sources and different time periods. The study used Support vector Machine (SVM) for emotions and sentiment detection, Various stock indices (DJIA, S&P500 and JPM), and Granger causality to test the causal relationship between emotions/sentiment polarity of financial news.
Although the authors found no Granger causality between positive and negative attitudes and stock prices, they did find that positive and negative emotions did have a significant impact on the price of select companies. (Tikkanen, 2021), tested if emotions from a smaller subgroup, i.e., people from the UK, perform better than emotions from the worldwide public in predicting close-to-close price direction of selected investment instruments from the London Stock Exchange.

The study used the sentiment of Twitter users as a proxy for the UK public sentiment. The authors discovered that the emotion "Fear" was found to Granger-cause differences in the close-to-close pricing in the worldwide dataset, while the emotion "Sadness" was shown to do the same in the UK dataset. (Eachempati et al., 2021), analysed the differential impact of COVID-19 through the sentiment of tweets and the connectedness of countries' stock markets using VAR spillover over for the pre-lockdown and post-lockdown phase. The study used 72000 tweets, extracted between January–May 2020 and closing prices of stocks in 4 countries USA, UK, China, and India. The results highlighted that the pre-lockdown period of the pandemic, considers the UK to be second most influential with positive spillover.

(Ghasiya and Okamura, 2021), investigated COVID-19 News Across Four Nations using Topic Modelling and Sentiment Analysis Approaches such as top2vec and RoBERTa for sentiment classification. The study used 100,000 COVID-19 news headlines and articles from four countries for 11 months. The conclusions of the authors were that the worst affected country, i.e., the UK, also had the highest percentage of negative sentiment during the studied period.

(Farimani et al., 2022), conducted an analysis of information gain of market data and mood in specialized financial newsgroups for price prediction using a BERT-based transformer language model fine-tuned for financial domain sentiment analysis. The author's experiments demonstrate the effectiveness of considering the mood of financial news when making market predictions.

Table 2.2: COVID-19 and financial markets related work – Part 1

Authors	Description	Approach
(Cerqueti and Ficcadenti, 2020)	Investigates the relationship between the Google search volumes of "coronavirus" and the stock index prices of different markets	Design indicators able to capture the connection between anxiety for the pandemic and expectations on the future outcomes of financial markets
(Cepoi, 2020)	Novel empirical evidence on the relationship between COVID-19 related news and stock market returns across the top six most affected countries by the pandemic	Panel quantile regression framework
(Albulescu, 2021)	Empirically investigates the effect of the official announcements regarding the COVID-19 new cases of infection and fatality ratio, on the financial markets volatility in the United States	Applied Ordinary Least Squares (OLS) regression to test on both global and US reported data
(Mazur et al., 2021)	This paper investigates the US stock market performance during the crash of March 2020 triggered by COVID-19	Analysing single-day extreme events by sector for industry returns and volatility
(Sharif et al., 2020)	Studies the connectedness between the spread of COVID-19, oil price volatility shock, the stock market, geopolitical risk and economic policy uncertainty in the US within a time-frequency framework	The coherence wavelet method and the wavelet-based Granger causality tests applied to US daily data
(Haroon and Rizvi, 2020)	Explores whether the media reporting of COVID-19, panic amongst investors, and the global sentiment has played a role in the previously unseen volatility in the equity markets	Exponential GARCH models
(Shehzad et al., 2020)	Analyses the non-linear behaviour of financial markets of the US, Germany, Italy, Japan, and China during the COVID-19 and Global Financial Crises period	Asymmetric Power Generalized Autoregressive Conditional Heteroscedasticity model
(So et al., 2021)	Studies the impacts of the COVID-19 pandemic on the connectedness of the Hong Kong financial market and compares the impacts of the previous financial crises in the past 15 years.	Construction of dynamic financial Networks based on correlations and partial correlations of stock returns
(Giangreco, 2020)	Study of the correlation between cases fatality rate and social behaviour of the Italian COVID-19 outbreak	Correlation analysis of cases fatality rate compared with social habit variables
(Şenol and Zeren, 2020)	Investigates the effect of COVID-19 outbreak on global markets between January 21, 2020 and April 7, 2020	Fourier Co-integration test

Table 2.3: COVID-19 and financial markets related work – Part 2.

Authors	Description	Approach
(Onali, 2020)	Investigates the impact of COVID-19 cases and related deaths on the US stock market (Dow Jones and S&P500 indices)	GARCH and VAR models
(Samitas et al., 2022)	Identifies volatility and contagion risk among stock markets during the COVID-19 pandemic	Dependence dynamics and Network analysis on a bi-variate basis
(Al-Awadhi et al., 2020)	Investigates the COVID-19 pandemic effect on the Chinese stock market	Panel data regression
(Goodell and Huynh, 2020)	Assesses the reactions of US industries to sudden COVID-related news announcements, concomitantly with an analysis of levels of investor attention to COVID-19	CAPM model and ordinary least squares (OLS)
(Lyócsa et al., 2020)	Explores the predictive power of Google searches on stock market volatility during the COVID-19 pandemic	Trends analysis and ordinary least squares
(Zhang et al., 2020)	Maps the general patterns of country-specific risks and systemic risks in the global financial markets	Correlation analysis and minimum spanning tree
(Erdem, 2020)	Analyses whether there is a relationship between the freedom of countries and their stock market movements in response to COVID-19 announcements	Panel regression analysis with 75 countries using their stock market index returns and volatilities as dependent variables and their COVID-19 data, their level of freedom, as independent variables.
(Zaremba et al., 2020)	This study is the first attempt to examine the influence of non-pharmaceutical policy responses to the COVID-19 pandemic	Capital Asset Pricing Model, Three and four-factor models
(Mirza et al., 2020)	Assess the price reaction, performance, and volatility timing of European investment funds during the outbreak of COVID-19	GARCH based event study
(Eachempati et al., 2021)	Statistical analysis of the relationship between daily market returns, COVID-19 outbreak and the market liquidity for firms listed in the FTSE 100	Regression models as measures for market liquidity including the trading turnover ratio (TVR), bid-ask spread (BAS) and high-low inter-day price (HLP)
(Rizwan et al., 2020)	Investigating how COVID-19 impacted the systemic risk in the banking sectors of eight of the most COVID-19 affected countries	CATFIN for systemic risk and Kruskal-Wallis one-way analysis-of-variance)

Table 2.4: COVID-19 and financial markets related work – Part 3

Authors	Description	Approach
(Liu et al., 2020)	Evaluates the short-term impact of the coronavirus outbreak on 21 leading stock market indices in major affected countries	Event study investigating the abnormal returns (ARs) and cumulative abnormal returns (CARs)
(Cinelli et al., 2020)	provide an in-depth analysis of the social dynamics in a time window where narratives and moods in social media related to the COVID-19 have emerged and spread	Models the spread of information using epidemic models and provide basic growth parameters for each social media platform
(Griffith et al., 2020)	Describes how the impact of COVID-19 has varied across industries, using data on share prices of firms listed on the London Stock Exchange	Trends and correlation analysis
(Sansa, 2020)	Investigates the relationship between the COVID-19 confirmed cases and both China and USA	Descriptive statistics and regression models
(Baker et al., 2020a)	examine the role of COVID-19 developments in recent stock market behaviour and draw comparisons to previous infectious disease outbreaks	Examines next-day newspaper explanations for each daily move in the US stock market – Text-based analysis and EMV tracker
(Topcu and Gulal, 2020)	Explores the predictive power of Google searches on stock market volatility during the COVID-19 pandemic	regression analysis and Driscoll-Kraay estimator
(Azimli, 2020)	Examines the impact of the novel coronavirus (COVID-19) on the degree and structure of risk-return dependence in the US	Quantile regression analysis
(Li et al., 2020)	Investigates whether the Infectious Disease EMV tracker proposed by (Baker et al., 2020a) has the additional predictive ability for European stock market volatility during the COVID-19 pandemic	Heterogeneous autoregressive models and Model Confidence set (MCS) test
(Takyi and Bentum-Ennin, 2021)	Evaluates and quantifies the short-term impact of the coronavirus disease of 2019 on stock market performance in thirteen African countries, using daily time series stock market data spanning 1st October 2019 to 30th June 2020	Bayesian structural time series
(Khanthavit, 2020)	How and how early the world and national markets reacted to COVID-19 events and news coverage	Event study and Abnormal returns analysis

Table 2.5: Sentiment analysis and financial markets related work

Authors	Description	Approach
(Mudinas et al., 2019)	Investigated the causal relationship between sentiment attitude/emotion signals and stock price movements through financial news using various sentiment signal sources and different time periods	Support vector Machine SVM for emotions and sentiment detection, Various stock indices (DJIA, S&P500 and JPM), Granger causality to test causal relationship between emotions/sentiment polarity of financial news
(Tikkanen, 2021)	Predicting the FTSE All-Share index daily close-to-close price direction using sentiment analysis on tweets from the UK	Support vector Machine and Granger causality test
(Zaenen and van den Bosch, 2007)	Explores a computable metric of positive or negative polarity in financial news text which is consistent with human judgments and can be used in a quantitative analysis of news sentiment impact on financial markets	Lexical cohesion-based metric of sentiment intensity and polarity in text, and an evaluation of this metric relative to human judgments of polarity in financial news.
(Ranco et al., 2015)	Explores the Effects of Twitter Sentiment on Stock Price Returns for companies that form the Dow Jones Industrial Average (DJIA) index	Event study to identify events as Twitter volume peaks, Support Vector Machine for sentiment extraction, Correlation analysis, and Granger causality tests
(Kalyani et al., 2016)	Taking non-quantifiable data such as financial news articles about a company and predicting its future stock trend with news sentiment classification.	Support Vector machine, Random forests, and Naïve Bayes classification
(Eachempati et al., 2021)	Analysing the differential impact of the COVID-19 through sentiment of tweets and the connectedness of countries stock markets using VAR spillover over for pre-lockdown and post-lockdown phase	72000 tweets with extracted between January–May 2020 and closing prices of stocks of 4 countries USA, UK, China, and India
(Ghasiya and Okamura, 2021)	Investigating COVID-19 News Across Four Nations: A Topic Modelling and Sentiment Analysis Approach using top2vec and RoBERTa for sentiment classification	100,000 COVID-19 news headlines and articles from four countries for 11 months
(Farimani et al., 2022)	Analysis of information gain of market data and mood in specialized financial newsgroups for price prediction	BERT-based transformer language model fine-tuned for financial domain sentiment analysis.

2.5 Data collection and description

This section outlines the methodology employed in data collection for both the UK government's COVID-19 briefings and the UK stock market data. We begin with the acquisition and preparation of textual data from the government's briefings. Following this, we provide a comprehensive overview of the data and variables associated with the UK stock market employed in this study.

2.5.1 Data Collection from the COVID-19 UK Government Briefings

In order to analyse the sentiment fluctuations in the UK government's stance on the Coronavirus pandemic, we collected data from their daily COVID-19 briefings. The primary source for this data is the official UK government website, Gov.UK (Wikipedia contributors, 2022). The timeframe for our research commences on the first trading day subsequent to the UK's initial COVID-19 briefing and concludes on the day the government detailed the vaccine roll-out strategy, preceding the lifting of all COVID-19 public restrictions. This translates to a period stretching from the 3rd of March 2020 to the 23rd of June 2021. Over this duration, a total of 160 COVID-19 briefings were held. It's noteworthy that our dataset only comprises the official statements initiated at the onset of each briefing, excluding the Q&A sessions.

2.5.2 Web scraping procedure

The collection of briefings was automated using a web scraping technique, a prevalent data mining method to harvest information from websites. In this case, our target was the UK government's official site. Two primary Python libraries facilitated this process:

- *requests* - Enables making HTTP requests to access web content.
- *BeautifulSoup* - An efficient package for parsing HTML and XML structures, generating a parse tree that aids in data extraction.

Initially, the target website was visually inspected to comprehend the layout and structure of the content. The UK government lists the COVID-19 statements chronologically at: *https://www.gov.uk//collections/slides-and-datasets-to-accompany-coronavirus-press-government-conferences*. Each statement possesses a unique URL. For illustration, the statement for July 31st, 2020 can be accessed at: *https://www.gov.uk/government/speeches/prime-ministers-statement-oncoronavirus-COVID-19-31-july-2020*.
Subsequently, the HTML source code for each URL was analysed to identify the relevant content elements, such as the title, date, and the actual transcript. After pinpointing these elements, the *requests* library fetched the content, looping over all 160 URLs to aggregate the statements. *BeautifulSoup* was then employed to parse and extract the requisite data from the fetched HTML content. The accumulated daily briefings were systematically catalogued in a table.

Finally, each briefing underwent a rigorous text preprocessing regimen to prime it for subsequent sentiment analysis.

2.5.3 COVID-19 UK conference briefings text Pre-processing

Upon collecting the conference briefing data, we subjected the daily textual content to analysis using Natural Language Processing (NLP) algorithms, focusing on the content of each COVID-19 government update. Our approach is rooted in standard text mining practices—a specialized domain within artificial intelligence (AI) that transforms unstructured textual input into structured, analysable data (Feldman and Sanger, 2007). The methodologies and techniques utilized for text processing in our research are illustrated in Figure 2.3. Our preprocessing

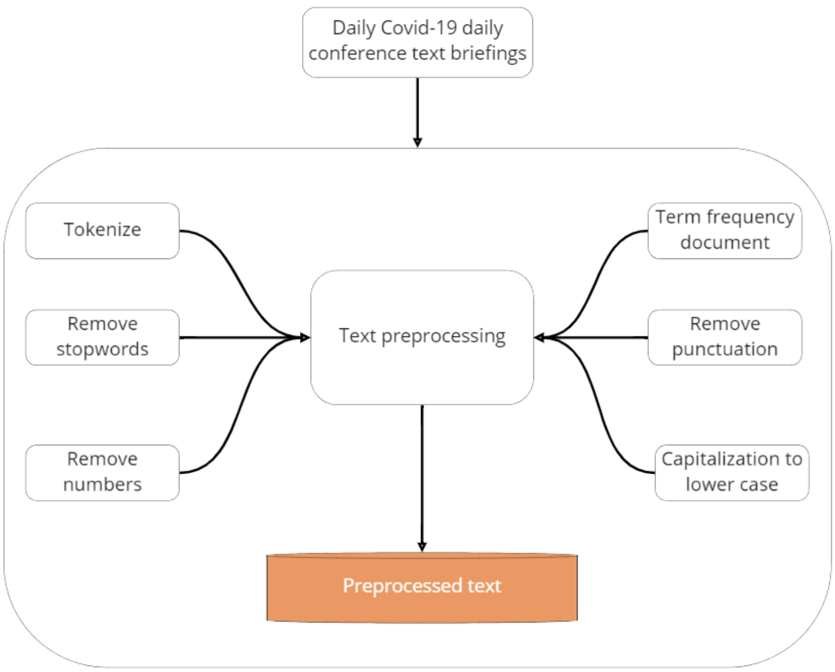


Figure 2.3: Text preprocessing procedure

commenced with the standardization of the corpus. This involved converting all text to lowercase to ensure uniformity. Subsequent steps included the removal of URLs, email addresses, stop words, special characters, and numbers. Following the cleansing, the text was tokenized—splitting the monolithic string into discernible sub-units or tokens, commonly referred to as words (Grefenstette, 1999). With a cleansed corpus at hand, we constructed a Term Document Matrix (TDM). This matrix enumerates terms and their respective frequencies across all texts, laying the foundation for word cloud analysis. The word cloud offers insights into the most recurrent terms within the COVID-19 conference briefings.

As illustrated in Figure 2.4, each cluster signifies a specific subject. For illustrative purposes, the clusters represent UK ministers, categorizing remarks made by each speaker throughout the pandemic briefings. It's interesting to observe that the predominant themes of speeches can be

Figure 2.4: UK ministers Word-cloud throughout the COVID-19 pandemic

discerned merely from the vocabulary chosen by each minister. For instance, in the discourses of Robert Jenrick, the Secretary of State for Housing, Communities, and Local Government, terms like "rough sleepers", "local council", and "domestic violence" frequently surfaced. Similarly, Gavin Williams, the Secretary of State for Education, often gravitated towards words such as "school" and "children", reflecting his office's core concerns.

In addition to the word cloud, we also compiled a daily frequency time series for words. This time series is designed for subsequent in-depth analytical exploration to discern patterns and trends over the duration of the pandemic for specific words.

While the word cloud analysis highlights the terms most commonly associated with the COVID-19 conference briefings, the mere frequency of word usage does not suffice to gauge the stock market's reaction to the pandemic. Therefore, this research extends its ambit to extract sentiment polarity from the COVID-19 UK government updates, delving deeper into this aspect in the methodology Section 2.6.2.

2.5.4 Data collection for the UK stock market

This study is centred on the UK stock market data. We procured hourly data, aligning with the timeframe of the daily COVID-19 conference briefings. Specifically, daily data for the FTSE 100, FTSE 250, FTSE 350, and FTSE All-Shares index was amassed.

The FTSE All-Share Index, maintained by the FTSE Russell—a subsidiary of the London Stock Exchange Group—is a capitalization-weighted index (Exchange, n.d.). Such an index weights its constituents, in this case, firms, based on their shares' market value. It amalgamates the FTSE 100, FTSE 250, and FTSE Small Cap Indexes, encapsulating approximately 600 of the 2000+ businesses listed on the London Stock Exchange. Among these, the top 100 represent the highest capitalisation, with the subsequent 250 being mid-capitalized.

Additionally, daily opening and closing prices of shares from AstraZeneca plc, a FTSE 100 con-

stituent, were obtained. AstraZeneca, a prominent multinational bio-pharmaceutical company, specializes in the discovery, development, and distribution of prescription medicines (Birmingham et al., 2018). Given their pivotal role in developing a COVID-19 vaccine, examining the impact of investor sentiment on this specific stock provides a focused case study, enriching our comprehensive understanding of the pandemic's effects on the market.
It's pivotal to recognize that the majority of the COVID-19 conference briefings were conducted around 5:00 pm UK time, post the closure of the London Stock Exchange market.
During the specified period of the COVID-19 conference briefings, the daily data collection was not necessarily synchronized with the exact timing of the briefings. Rather, data was collected for the same day on which a conference briefing occurred.
The London Stock Exchange is open from Monday to Friday, except for public holidays in England and Wales. The standard trading hours for the London Stock Exchange are from 8:00 a.m. to 4:30 p.m. UK time, except on the Friday before Christmas Day when trading hours are from 8:00 a.m. to 12:30 p.m. Thus, we collected hourly stock prices starting from 8:00 a.m. and continuing each subsequent hour. This data is available as such on the Bloomberg Terminal, where data is only available at the top of the hour, meaning precisely at the start of each hour. The first available hourly data point of a trading day would correspond to the hour starting at 8:00 a.m. The second hourly data point for the trading day would correspond to the hour ending at 9:00 a.m., and so forth. The last available hourly data point would correspond to the hour ending at 4:00 p.m. instead of 4:30 p.m. (market closing time), as hourly data are only available at the top of the hour, and the last hour available is 04:00 p.m. For this reason, in this research, 04:00 p.m. is assumed to be the last trading data point.
The COVID-19 conference briefings were held daily at 5:00 pm UK time after the London Stock Exchange had closed at 4:30 pm. To explore the impact of these briefings on the UK stock market, we analysed stock returns at three different periods to provide a nuanced understanding of market behaviour:

- (1) On the day of the COVID-19 conference briefing – Build up Returns (Rt_b):

  Rt_b was calculated between market opening (08:00 am) and market closing (04:00 pm) of the same day (d) and on a day a conference briefing was about to occur, as this period represents the build-up to the conference and the potential effects of pre-conference news and announcements. The following formula was used to calculate Rt_b:

  $$Rt_b = \frac{price(close)_d - price(open)_d}{price(open)_d} \tag{2.1}$$

- (2) Day post COVID-19 conference briefing – Overnight Returns (Rt_o):

  Rt_o was calculated over the period that covers the last data point recorded at market closing ($d-1$), which is at 4:00 pm, right before the start of the COVID-19 conference briefing at 05:00 pm, until market reopening on the next day (d) at 08:00 am. Overnight returns are particularly interesting to study because the COVID-19 briefings occurred after market closing, which means that investors had to wait until the next day to respond to any news or announcements made during the briefings. Therefore, analysing overnight returns can help capture the market's reaction to the previous day's events, including any developments or news that may have been discussed during the briefing. The following formula was used to calculate Rt_o:

$$Rt_o = \frac{price(open)_d - price(close)_{d-1}}{price(close)_{d-1}} \tag{2.2}$$

- (3) Day post COVID-19 conference briefing – First 2 hours of trading returns (Rt_f):

  Rt_f was calculated for the day following a conference briefing during the first two opening hours between 8:00 ($h = 0$) am and 10:00 am ($h = 2$); overnight returns (Rt_o) can capture any immediate reaction to news or events that occur after market close, while the first 2 hours of the next day after the briefings can provide insight into how that reaction has evolved once the market reopens and additional information becomes available. The following formula was used to calculate Rt_f:

$$Rt_f = \frac{price(open)_{h=2} - price(open)_{h=0}}{price(open)_{h=0}} \tag{2.3}$$

The rationale for employing three return variables is anchored in capturing the nuanced reactions of the stock market to the COVID-19 conference briefings at various times:

Build-up Returns (Rt_b): This return measures market anticipation and early reactions on the day of a conference briefing. The financial market is a forward-looking entity and anticipatory moves often occur before significant events. This notion stems from the Efficient Market Hypothesis (EMH), which suggests that stock prices instantly reflect all available information, and thus always trade at their fair value on exchanges (Fama, 1970). This means that investors, in their collective wisdom, are always looking ahead, trying to anticipate future events and adjust their trading strategies accordingly. By observing the build-up returns, we can gauge how the market is positioning itself in anticipation of the briefing.

Overnight Returns (Rt_o): As the conference briefings occurred after the market close, this metric is essential to capture any immediate reactions to the announcements. The overnight period is when the market digests the information presented during the briefing and any other related news. Analyzing these returns provides insight into the immediate sentiment and forecasts of investors based on the briefing content.

First 2 hours of trading returns (Rt_f): This metric captures how the immediate reactions (from overnight returns) evolve when faced with the reality of a new trading day and additional information. It's a window into the sustained or changed sentiment from the overnight period and offers a more granular look at investor behaviour.

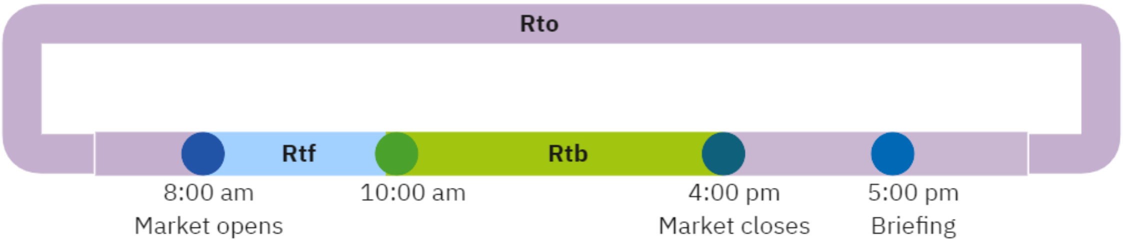


Figure 2.5: Visual representation of the timeline for the calculation of the returns.

The combination of these three return variables offers a comprehensive view of the returns across different time intervals surrounding the conference briefings. While each variable individually provides a snapshot, collectively, they form a more holistic picture of market dynamics.

Table 2.6 provides a summary of statistics of the daily stock returns analysed during the COVID-19 pandemic. It is important to note that after excluding weekends, which are observed on Saturdays and Sundays in the United Kingdom and during which the London Stock Exchange remains closed, a total of 138 data points have been retained. These data points represent instances where the London Stock Exchange was open and coincided with the occurrence of the UK COVID-19 conference briefings.

Statistics	FTSE100	FTSE250	FTSE350	FTSE All-Shares	AstraZeneca PLC
N	138	138	138	138	138
μ	0.04	0.05	0.05	0.05	0.02
σ	0.90	0.86	0.89	0.87	1.52
Min	-3.07	-2.79	-2.88	-2.83	-4.85
25%	-0.41	-0.45	-0.38	-0.38	-0.94
50%	0.01	-0.00	0.00	0.00	-0.05
75%	0.38	0.46	0.36	0.35	1.02
Max	3.98	4.71	4.96	4.94	4.70
$Skew.$	0.76	1.09	1.27	1.26	0.02
$Kurt.$	6.49	8.24	8.73	8.81	3.45

Table 2.6: Summary statistics of UK daily stock returns during the COVID-19 pandemic

Table 2.6 presents a comprehensive summary of the UK's daily stock returns throughout the COVID-19 pandemic. The table not only offers central tendencies, denoted by μ, but also provides an in-depth view of dispersion measures, which include standard deviation (σ), minimum and maximum values, and interquartile ranges. Additionally, it gives insights into the distribution shape via skewness and kurtosis metrics.

A notable observation is the skewness values for the FTSE 250, 350, and All-Shares indices, which exceed one, suggesting a right-skewed distribution. In contrast, the FTSE 100's skewness value, being less than one, indicates only a slight skew. AstraZeneca plc's daily returns display near symmetry, with a skewness of just 0.02.

The skewness in the FTSE 250, 350, and All-Shares indices suggests more significant occasional gains than losses. However, consistent positive skewness can sometimes be a precursor to market corrections if the market has been under-reacting to positive news. The near-zero skewness of AstraZeneca plc's daily returns indicates a relatively symmetrical return distribution.

The pronounced kurtosis values for the FTSE indices, exceeding three, highlight the potential for extreme returns. Such leptokurtic behaviour implies that while these stocks might exhibit expected performance patterns most of the time, there exists a higher risk of significant deviations. This can be attributed to increased market sensitivity, especially given the uncertainties of a pandemic. Investors often view such distributions as indicators of heightened risk due to their potential for extreme positive and negative returns (Taleb, 2007).

The kurtosis values for all FTSE stocks surpass three, signalling a leptokurtic distribution. In the context of stock returns, this implies that these stocks experienced greater volatility and a heightened potential for extreme price changes than a standard normal distribution would suggest (Hair Jr et al., 2021). Such behaviour can be indicative of increased market sensitivity, especially during unpredictable events like a pandemic. Visual representations of the return distributions for each stock are available in Figure A.1 within the Appendix.

Furthermore, this research extends beyond mere stock returns. It delves into both closing and opening stock prices and their respective fluctuations amidst the COVID-19 pandemic. Detailed findings on these aspects are elaborated upon in the Results and Discussion section 2.7.

Complementing the stock data, information on the number of COVID-19 cases and fatalities in the UK was sourced from *https://ourworldindata.org*, ensuring temporal alignment with the UK stock market data collection period.

2.6 Methodology

This section introduces the machine learning and deep learning methodologies employed in the sentiment polarity classification. Subsequently, we delve into the techniques adopted for analysing the correlation between the UK stock market's response and the COVID-19 pandemic. Finally, we provide an in-depth exploration of the Granger causality approach, enabling the assessment of potential Granger causal relationships among the study's focal variables.

2.6.1 System design

Following the system design delineated below, we first extract the sentiment from these briefings. Then, we correlate this sentiment with the UK stock market performance. Finally, we employ the Granger causality approach to discern any potential causative relationships between the extracted sentiment and market movements.

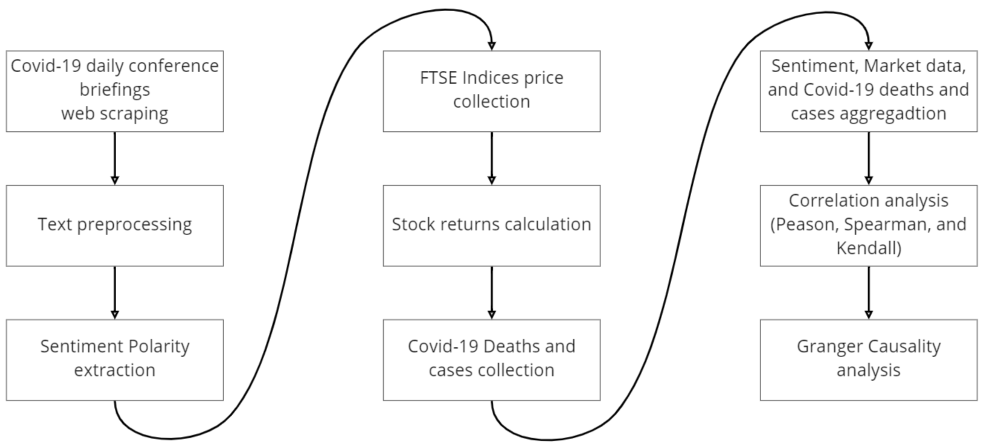


Figure 2.6: System Design

Initially, web scraping is employed to collate the daily COVID-19 conference briefings. This method automates the extraction of text from the UK government's conference briefings published online (further elaborated in the data collection and description section 2.5). Following the removal of punctuation, stop words, and subsequent tokenisation, among other textual preprocessing steps (illustrated in Figure 2.3), the briefings are primed for sentiment classification. Here, both lexicon rule-based and transformers-based techniques are utilised (detailed in subsection 2.6.2) to discern the sentiment polarity of the daily COVID-19 UK updates. Subsequent stages involve gathering the FTSE indices prices and AstraZeneca PLC stock details, from which stock returns are computed. In addition, data on COVID-19-related UK fatalities and cases are amassed to serve as benchmark variables. Ultimately, with the sentiment polarity scores of each UK daily update, alongside the FTSE and AstraZeneca plc prices and stock returns, we proceed to correlation analysis and Granger causality tests.

2.6.2 Sentiment analysis

For years, finance researchers have endeavoured to quantify and assess the impact of information on financial markets. Recent advancements in computational linguistics, Natural Language Processing (NLP), Machine Learning (ML), and econometrics have empowered scholars to apply these methodologies across diverse financial research areas (Chouliaras, 2016). In this research, we harness the capabilities of Natural Language Processing and Machine Learning to delve deeply into the market's response to the COVID-19 crisis. In particular, we explore sentiment classification methods to gauge the daily investor sentiment within the UK's financial market throughout the COVID-19 pandemic.

Sentiment classification primarily serves two core objectives: emotion recognition and polarity detection within textual or transcribed data (Tikkanen, 2021). Emotion recognition strives to identify specific emotions conveyed in texts, such as joy, anger, or sadness. In contrast, polarity detection categorises texts based on their overarching positive or negative sentiment. While this typically results in a binary classification—namely, "positive" or "negative"—it can be further refined to include nuanced categories like "neutral", "very positive", and "very negative" (Cambria et al., 2017).

Sentiments can manifest at various levels: individual words, sentences, or entire documents. Analysis can be tailored accordingly. In our research, we focus on extracting sentiment from individual sentences. By aggregating these, we derive the overall sentiment of a specific COVID-19 update.

Sentiment analysis can be categorised into three primary approaches (Kansal et al., 2020). Firstly, there's the Rule/Lexicon-Based approach. In this method, a lexicon—a collection of words with their respective polarities—is employed. Each word maps to a specific sentiment, and the overall sentiment of a sentence is derived from the combined sentiment of its constituent words. Typically, lexicon sentiment analysis yields a polarity score ranging from -1 to 1, where -1 signifies a strongly negative sentiment, and 1 indicates a robust positive sentiment.

The second technique is the Machine Learning-Based approach. This method hinges on a classifier trained extensively using a vast repository of labelled data, enhancing its efficacy in predicting the sentiment of a given text or document.

Lastly, there's Cross Domain sentiment analysis. This approach is invoked when a model or classifier, initially trained on data from one domain (referred to as the "input domain"), is subsequently applied to data from a distinct domain (the "output domain"). This application holds true irrespective of whether the data in the latter domain is labelled.

For our research, we employ a combination of the Rule/Lexicon-Based and Machine Learning-Based approaches to extract sentiment from the daily COVID-19 government updates.

2.6.3 Rule/Lexicon-Based approach

We commence our exploration with the Rule/Lexicon-Based approach, utilising the Valence Aware Dictionary for Sentiment Reasoner (VADER) model (Hutto and Gilbert, 2014) and Text-Blob (**?**). Both are NLP Lexicon-based sentiment analysers. These techniques draw upon dictionaries that associate lexical features with emotion intensities, termed sentiment scores.

2.6.3.1 Valence Aware Dictionary and Sentiment Reasoner (VADER)

Using the VADER-based approach, the sentiment score of a sentence is determined by aggregating the sentiment scores of each word found within the VADER dictionary that appears in the sentence. The creators of VADER then employ Hutto's (Hutto and Gilbert, 2014) normalisation (refer to Eq. 2.4) to generate the final sentiment scores. These scores span from -1, denoting negative sentiment, to 1, representing positive sentiment.

$$\frac{x}{\sqrt{x^2 + \alpha}} \tag{2.4}$$

where,

- x is the sum of the sentiment scores of the constituent words of the sentence, and
- α is a normalization parameter that was set to 15 in the authors' (Hutto and Gilbert, 2014) experiment.

2.6.3.2 Text-Blob

This method employs the Naïve Bayes model for the classification task. This classifier, a supervised Machine Learning technique, leverages the Bayes theorem to determine sentiment distribution across the data. The model has been trained using the Natural Language Toolkit (NLTK) to identify the valence of aggregated tweets (**?**). The equation, grounded in Bayes' theorem, used to forecast sentiment probability is as cited in (Wagenmakers, 2007; Medhat et al., 2014; Manguri et al., 2020):

$$P(label/features) = \frac{P(label) * P(features/label)}{P(features)} \tag{2.5}$$

where,

- $P(label$ is the prior probability of a label, and
- $P(features/label)$ is the prior probability that a given feature set is being classified as a label and
- $P(features)$ is the prior probability that a given feature set is occurring

2.6.4 Limitations of Lexicon-based Sentiment Analysis

Historically, lexicon-based sentiment analysis has been the favoured approach for predicting financial market movements (Farimani et al., 2022). However, this technique is not without its drawbacks. The primary limitation of rule-based systems is their inability to accurately interpret linguistic nuances and contextualise them. For instance, while lexicon-based sentiment analysis might correctly deem the statement "I love Statistics" as positive, it could misclassify the phrase "I do not love Statistics" as negative (Tikkanen, 2021). Additionally, these systems may struggle with sarcasm, idiomatic expressions, and cultural variations in language use.

This is where Machine Learning (ML) comes into the picture. ML, a subset of artificial intelligence, focuses on developing algorithms that improve through experience. Its application to economic issues has a long-standing history, traceable back to 1974 (Lee and Lee, 1974). To the best of our knowledge, the study by (Wang et al., 1984) was the pioneering effort employing an ML approach explicitly for an economics topic. Although (Gogas and Papadimitriou, 2021) references the term 'AI' in their research, it's plausible that 'ML' would have been the more accurate designation. By 1988, the application of ML in economics had advanced, with (White, 1988) utilising Neural Networks (NN) to predict daily stock returns for IBM. Since then, the adoption of ML within the realm of economics has seen significant growth.

Within the realm of ML, Neural Networks (NN) stand out, especially for sentiment analysis tasks. Comprising interconnected layers of algorithms called neurons, they're designed to recognise and interpret patterns. According to (Otter et al., 2020), Neural Networks are conceived as interconnected nodes or neurons. Each neuron accepts various inputs and subsequently produces an output. The nodes in the output layers determine a weighted sum of the values sourced from the input nodes. Following this, they employ fundamental nonlinear transformation functions on these aggregate sums to generate outputs. When the network exhibits errors or discrepancies at the output nodes, weight adjustments are made to rectify them. In contemporary networks, such rectifications are typically achieved using stochastic gradient descent combined with the derivatives of errors at nodes—a process termed back-propagation.

In the context of Natural Language Processing, Neural Networks undergo training with labelled text data. Each piece of this data is introduced into the network individually for assessment. The primary objective of the Neural Network is to identify and harness combinations of elements with predictive capacities. For instance, when presented with textual data, the network discerns and represents sentiment by amalgamating these elements into coherent representations (Colnerič and Demšar, 2018). Given the evident limitations of lexicon-based systems, it's clear that the depth and adaptability offered by ML, and especially NN, present a promising direction for refining sentiment analysis methodologies.

One standout model in the realm of Neural Networks that has garnered significant attention for its prowess in NLP tasks is BERT (Bidirectional Encoder Representations from Transformers). Developed by researchers at Google, BERT revolutionised the way we understand and process language through deep learning.

2.6.4.1 Transformer Architectures: BERT and its Variants

The introduction of the Attention Mechanism in NLP has marked a significant shift in the way textual data is processed and understood (Bahdanau et al., 2014; Vaswani et al., 2017). This mechanism, which provides context to words in a given sentence, has led to the development of models that outperform many state-of-the-art algorithms, particularly when combined with bidirectional techniques (Sousa et al., 2019).

One such groundbreaking model is the Bidirectional Encoder Representations from Transformers (BERT) (Devlin et al., 2018). BERT revolutionised the NLP landscape by achieving state-of-the-art results on a range of tasks, including sentiment classification. Its bidirectional nature means it can understand the context from both the left and right of a word in a sentence, making it particularly adept at understanding the nuances of human language.

Following BERT's success, there have been several variants aimed at refining and optimising its foundational architecture. One notable variant is RoBERTa (A Robustly Optimized BERT Pretraining Approach) (Liu et al., 2019). While BERT relies on masked language modelling for pretraining, RoBERTa modifies this approach by using more data, larger batch sizes, and removing the next-sentence pretraining objective. As a result, RoBERTa often outperforms BERT in specific NLP tasks.

Recognising the significance of Transformer-based architectures, (Wolf et al., 2019) introduced the transformers' library in 2020. This library is dedicated to supporting these architectures and has made deploying pre-trained models like BERT and RoBERTa more accessible.

In this study, we leverage the capabilities of the transformers' library, namely the BERT model, to classify sentiment in text. Nevertheless, the choice to employ BERT over RoBERTa in this study was made after careful consideration. BERT's pre-training methodology, which involves both masked language model and next-sentence prediction tasks, might be more suited for capturing the semantic relationships present in the government briefings, which often have intricate contextual nuances. Furthermore, BERT's extensive usage and validation across a myriad of NLP tasks provides a well-established foundation for sentiment analysis, particularly in the context of financial data.

Moreover, while RoBERTa has shown impressive results, its optimisations might not necessarily translate to significant improvements in the specific task of sentiment analysis for our data set. Given the computational costs, it was deemed prudent to adopt BERT, a tried and tested model with proven efficacy in similar scenarios.

2.6.5 Illustrative Sentiment Analysis Examples

In this research, we use both VADER and Text-Blob, based on the Lexicon approach, to calculate the sentiment mood time series. To improve our sentiment scores, we also use a BERT-based transformer language model. It's important to note that our main goal isn't to find the best sentiment model but to analyse the sentiment in the UK government's COVID-19 announcements and see its possible effect on the UK stock market. Despite using different methods, the sentiment scores are consistent across all. A clear relationship between the three sentiment score methods can be seen in Figure **??** in the Appendix.
To further cement our understanding of these methods, let's delve into some specific instances of sentiment classification. Figure 2.7 presents a few representative examples, highlighting the capabilities and occasional limitations of our chosen algorithms.

Example 1: *On the 3rd of March 2020:*

Input sentence: {"and let me be absolutely clear that for the overwhelming majority of people who contract the virus, this will be a mild disease from which they will speedily recover as we already seen"}

Output sentiment classification:

BERT based: [{'label'} : 'POSITIVE', 'score' : 0.99}]
VADER based: {'neg' : 0.92 ,'neu' : 0.92, 'pos' : 0.07, 'compound' : 0.43}
TEXTBLOB based: 0.31

Example 2: *On the 23rd of March 2020:*

Input sentence: {"parks will remain open for exercise but gatherings will be dispersed"}

Output sentiment classification:

BERT based: [{'label'} : 'NEGATIVE', 'score' : 0.99}]
VADER based: {'neg' : 0.0 ,'neu' : 1.0, 'pos' : 0.0, 'compound' : 0.0}
TEXTBLOB based: 0.0

Example 3: *On the 3rd of July 2020:*

Input sentence: {"because we are not out of the woods,"}

Output sentiment classification:

BERT based: [{'label'} : 'NEGATIVE', 'score' : 0.98}]
VADER based: {'neg' : 0.0 ,'neu' : 1.0, 'pos' : 0.0, 'compound' : 0.0}
TEXTBLOB based: 0.0

Figure 2.7: Sentiment classification examples according to the methods described in Section 2.6.3, and 2.6.4.1

Looking at the first example in Figure 2.7, the UK Prime Minister addresses the British public about the COVID-19 pandemic. All three sentiment analysis methods have rightly identified the sentiment as positive. Using the BERT-based model, the statement is labelled as "positive" with a high confidence score of 99%. The VADER approach gives a compound sentiment score, which reflects 43% positivity – here, any score above zero is considered positive. Meanwhile, Text-Blob offers a singular polarity score of 31%, marking the statement as positive since its score exceeds zero.
In the second instance, only the BERT-based model correctly identifies the sentiment as "negative", and with a confidence score of 99%. This highlights a limitation of Lexicon/Rule-based methods, which can struggle with statements that combine positive and negative elements. The third example further underscores BERT's accuracy. It effectively identifies the sentiment in the idiomatic phrase *"because we are not out of the woods"* as negative, demonstrating a confidence score of 98%. This showcases BERT's advantage over Lexicon/Rule-based models in interpreting idiomatic expressions.
While these isolated examples provide insights into the individual performance of each sentiment analysis method, a comprehensive evaluation requires a more structured aggregation of these scores. The following section delineates the procedure we employed to cohesively aggregate sentiment scores for entire briefings.

2.6.6 Sentiment Score Aggregation Procedure

To comprehensively evaluate the sentiment of each COVID-19 conference briefing, we incorporated three distinct sentiment analysis algorithms: VADER, Text-Blob, and BERT.

- **VADER and Text-Blob**: These algorithms generate polarity scores within a range of -1 to 1. A score closer to -1 signifies a negative sentiment, whereas a score closer to 1 indicates a positive sentiment.

- **BERT**: BERT provides labels along with confidence intervals. These labels are subsequently converted into numerical scores. Specifically, a negative label is given a value of -1, while a positive label is assigned a value of 1.

Figure 2.8: Sentiment scores aggregation procedure

For a holistic sentiment score of each briefing:

1. The briefing text is segmented into individual sentences.

2. Each sentence is then assigned a sentiment score using the aforementioned algorithms.

3. The individual scores are aggregated, culminating in the final sentiment score for the entire briefing. For instance, if a briefing's sentences have scores of -0.5, 0.5, and 0.99, the aggregate sentiment score would be calculated as $0.99 + 0.5 \quad 0.5 = 0.99$. A higher value implies a more positive briefing sentiment, and conversely, a lower or negative value indicates a more negative sentiment.

Figure 2.7 is an example of the final output table for the first five COVID-19 government updates in the United Kingdom.

Conference briefing	Number of sentences	Sentiment polarity (Text-Blob)	Sentiment polarity (VADER)	Sentiment polarity (BERT)
*Day*1	18	3.24	4.47	4
*Day*2	21	3.36	4.18	5
*Day*3	49	5.48	3.84	1
*Day*4	42	4.45	4.02	-2
*Day*5	31	3.67	3.66	5

Table 2.7: Final table with 5 COVID-19 UK government updates and their sentiment scores

2.6.7 Challenges and Nuances in Briefing Sentiment Analysis

Having outlined our sentiment aggregation procedure and presented preliminary results, it's crucial to delve deeper into the complexities and challenges associated with interpreting sentiment from government briefings on a topic as multifaceted as COVID-19. The inherently subjective nature of sentiment analysis presents a challenging aspect in the context of COVID-19 briefings. For instance, consider the sentence, "The lockdown has successfully flattened the curve." At face value, this statement appears to be positive, as it suggests that the lockdown measures have achieved their intended effect. However, the notion of a 'successful lockdown' could be fraught with negative sentiment for individuals who have experienced economic hardship, social isolation, or mental health struggles due to the restrictions. Therefore, the sentiment associated with this statement is complex and could be interpreted as both positive and negative, depending on the context and the audience. This example underscores the intricate challenges associated with language and sentiment in these COVID-19 briefings, necessitating a nuanced approach to sentiment analysis.
In light of these complexities, we undertook a qualitative comparative assessment involving BERT, VADER, and Text-Blob. In this assessment, the BERT model consistently demonstrated a more accurate interpretation of sentiment from the COVID-19 daily briefings compared to VADER and Text-Blob. We manually inspected a range of predictions across these models to

ensure our selection was optimally attuned to this nuanced context.
BERT's superior performance, as observed in this comparative assessment, aligns with the findings in the literature. BERT is known for its ability to understand the context of a word based on its surroundings, thanks to its sophisticated bidirectional training mechanism. This capacity allows BERT to capture the semantic meanings of sentences more accurately, thereby making it a highly suitable choice for our study (Devlin et al., 2018).
Conversely, VADER and Text-Blob utilize simpler techniques such as lexical approaches and rule-based methods. While these methods have shown to perform well in general scenarios (Sudhir and Suresh, 2021), they may struggle with the complex language structures or specialized terminologies present in our data, as observed in our qualitative assessment. Thus, for the subsequent stages of our research, specifically the correlation and Granger causality analysis, we opted to exclusively utilise the sentiment scores derived from the BERT model. This decision aims to ensure that our analyses are rooted in the most accurate and contextually aware sentiment data available.

In the subsequent phase of our methodology, we will delve into financial modelling, specifically correlations and the Granger causality test, to ascertain the influence of the sentiment polarity derived from the UK's COVID-19 conference updates on the nation's stock market.

2.6.8 Financial modelling via correlations

Being a global health crisis that created strong effects on the real economy (Goldstein et al., 2021), one would expect to see correlations between the COVID-19 pandemic and the stock market. In the process of investigating this hypothesis, we examine how the UK financial market movements are related to the COVID-19 pandemic through the number of COVID-19 fatalities and cases. We also scrutinise scrutinize the UK stock market by looking for correlations between the most relevant UK stock indexes and the public mood, which arguably would translate to the market investors' sentiment. The investors' sentiment is implied through the sentiment polarity of the COVID-19 UK daily government updates. For more details regarding the sentiment analysis calculation, refer to the Sentiment classification section 2.6.2.

Correlation analysis

Correlation is a statistical measure of a monotonic relationship between two variables. A monotonic connection between 2 variables is one in which, as the value of one variable increases, the value of the other variable also increases, or, as the value of one variable increases, the value of the other variable declines. In correlated data, the change in magnitude of one variable is coupled with the change in magnitude of another variable, either in the same or opposite direction. In other words, higher values of one variable are typically associated with either higher (positive correlation) or lower (negative correlation) values of the second variable and vice versa (Schober et al., 2018). In this study, we employed:

- **Pearson correlation coefficient**

  The Pearson correlation coefficient (r_{XY}), named after the English mathematician and bio-statistician Karl Pearson, is a statistical measure of the linear connection between two variables X and Y and is defined as follows (Profillidis and Botzoris, 2018):

  $$r_{XY} = \frac{(cov(X,Y))}{\sigma_X.\sigma_Y} \tag{2.6}$$

  where,

  - $cov(X,Y)$ is the covariance between X and Y, and
  - $\sigma_{\mathbb{X}}, \sigma_{\mathbb{Y}}$ is the standard deviation of X and standard deviation of Y

  The values of r_{XY} range between $[-1\ ,\ 1]$ where $r_{XY} = 0$ indicating that there is no linear relationship between X and Y, and the relationship becomes more substantial as the absolute value of r_{XY} increases and ultimately approaches the coefficient -1 or 1. The Pearson correlation coefficient assumes that both variables should be normally distributed and is very sensitive to outliers. For non-normal distributions (for data with extreme values outliers), the Pearson correlation coefficient should be calculated from the ranks of the data and not from their actual values. The coefficients designed for this purpose are Spearman's and Kendall's Tau coefficients (see the below paragraphs for definitions). These coefficients can be calculated as a measure of linear and nonlinear monotonic (i.e., continuously increasing or decreasing) relationships without any assumptions (Akoglu, 2018).

- **Spearman rank correlation coefficient**
  A Spearman coefficient is similar to a Pearson correlation coefficient. However, the Spearman rank correlation coefficient is computed using rankings rather than actual values for each variable. The Spearman correlation coefficient is not limited to continuous variables. The coefficient quantifies strictly monotonic correlations between two variables using ranks. In addition, this trait renders a Spearman coefficient comparatively resistant against outliers. As for the Pearson coefficient, the Spearman coefficient varies between -1 and 1. It may be regarded as expressing anything from a perfect monotonic connection with a value of 1 to the absence of any correlation where the coefficient is 0, (Caruso and Cliff, 1997).

- **Kendall Tau rank correlation coefficient**
  The Kendall rank correlation coefficient (Abdi, 2007), measures the degree of similarity between two sets of rankings assigned to the same collection of objects. This coefficient is dependent on the number of object pair inversions required to change one rank order into the other. In order to accomplish this, each rank order is represented by the set of all pairs of objects (e.g., $[a, b]$ and $[b, a]$ are the two pairs representing the objects a and b, and a value of 1 or 0 is assigned to each pair based on whether its order matches or does not match the order in which the two objects were placed.

The correlation analysis using the above correlation coefficient is presented in the Results and Discussion section 2.7.

2.6.9 Financial modelling via Granger Causality test

Lastly, we examine possible justifications for the UK stock market response to COVID-19 by performing a Granger causality test. First proposed in 1969 by the British econometrician Sir Clive Granger Granger (1969), the Granger causality method is a statistical hypothesis test that answers the question: Can the previous values of a time series X be used to forecast the current value of a time series Y? If the past values of X help in forecasting the current value of Y, it is said that time series "X Granger causes time series Y" Tikkanen (2021). The null hypothesis H_0 is rejected at any α level provided that the probability value is less than that level. In his seminal paper "Testing for Causality: A Personal Viewpoint",(Granger, 1980). Granger defines causality based on two fundamental axioms:

- **Axiom A.** The past and present may cause the future, but the future cannot cause the past.
- **Axiom B.** I_t contains no redundant information, so if some variable Z_t is functionally related to one or more other variables, in a deterministic fashion, then Z_t should be excluded from I_t.

Based on these axioms, Granger provides a definition of causality as follows:
A variable Y_t is said to cause X_{t+1} if the probability distribution of future values of X_{t+1} given the set of all past information I_t differs from the probability distribution of X_{t+1} given all past information excluding Y_t (denoted as $I_t - Y_t$), for some set A. This can be represented mathematically as:
Let A be an arbitrary set. Y_t is said to Granger-cause X_{t+1} if:

$$\mathbb{P}(X_{t+1} \in A | I_t) \neq \mathbb{P}(X_{t+1} \in A | I_t - Y_t) \tag{2.7}$$

Where:

- $\mathbb{P}$ refers to the probability operator,
- A is an arbitrary non-empty set,
- $I(t)$ denotes the set of all information available up to and including time t,
- $I - X(t)$ is the set of all information available up to and including time t, excluding information about X.

In a Granger causality test, this statement serves as the null hypothesis, asserting that past values of Y_t do not provide useful information in predicting future values of X_{t+1}. If this null hypothesis is rejected, then we can conclude that Y_t Granger-causes X_{t+1}.
To conduct such a test, we commonly begin by modelling X_{t+1} as an autoregressive (AR) process. We then assess whether adding lagged values of Y_t to this AR model improves the

prediction of X_{t+1}. If so, we conclude that Y_t is a Granger cause of X_{t+1}.
Let us assume that Y and X are two variables having stationary time series of data or observations. To test the null hypothesis that X does not Granger-cause Y, we first find the appropriate p lagged values of Y (the order p of the $AR(p)$ process) to include in an AR process of Y (Profillidis and Botzoris, 2018):

$$y_t = c + \Phi_1 . y_{t-1} + \Phi_2 y_{t-2} + ... + \Phi_p y_{t-p} + \epsilon_t \tag{2.8}$$

where,

- y_t is a time series variable measured at time t, and
- y_{t-1} is a time series variable measured at time $t-1$, and
- $\Phi_1, \Phi_2, ..., \Phi_p$ are the parameters of the AR model, and
- c is the intercept of the $AR(p)$ process, and
- ϵ_t is white noise.

Next, the (2.8) equation is augmented by including lagged values of the variable X:

$$y_t = c + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + ... + \Phi_p y_{t-p} + \omega_1 x_{t-1} + \omega_2 . x_{t-2} + ... + \omega_n x_{t-n} + \epsilon_t \tag{2.9}$$

where,

- n is the longest lag length for which the lagged value of the variable X has been proved statistically significant.

The Granger causality test is a test of a joint hypothesis that lagged values of X are not statistically significant. Therefore, the null hypothesis is:

$$H_0 : \omega_1 = \omega_2 = ... = \omega_n = 0 \tag{2.10}$$

While the alternative hypothesis:

$$H_1 : \omega_n \neq 0, \; for \; at \; least \; one \; value \; of \; n \tag{2.11}$$

Thus, to test the null hypothesis, one needs to estimate two models. One is a restricted model which omits historical values of X, represented in Eq. (2.8) while the second model (unrestricted) has the full specification mentioned in Eq. (2.9). Furthermore, to test for Granger causality, one needs to carry out a χ^2 test which compares the restricted model (2.8) with the unrestricted model (2.9).
We retain in Eq. (2.9) all lagged values of the variable X that are statistically significant, provided that jointly all of them contribute to the explanatory ability of Eq. (2.9) according to

the χ^2 test.
In this study, the Granger causality test is applied to investigate the influence of the COVID-19 pandemic on the behaviour of the Financial Times Stock Exchange (FTSE) indices in the United Kingdom. The FTSE indices behaviour, denoted by M_{c19_t}, can be represented by either market returns or market prices at time t. We test the Granger causality through three hypotheses:

1. H_0: The sentiment polarity of COVID-19 conference briefings, denoted by $S_{c19_{t-1}}$, does not Granger cause M_{c19_t}.

   The null hypothesis is tested by estimating a restricted and an unrestricted model.

   The restricted model excludes the historical values of sentiment polarity and is given by:

   $$M_{c19_t r} = c + \Phi_1 M_{c19_{t-1}} + \ldots + \Phi_p M_{c19_{t-p}} + \epsilon_t \tag{2.12}$$

   where:

   - $M_{c19_t r}$ represents the market state at time t in the restricted model,
   - c is a constant,
   - $\Phi_1, ..., \Phi_p$ are parameters to be estimated,
   - ϵ_t is the error term at time t.

   The unrestricted model includes the historical values of sentiment polarity, $S_{c19_{t-1}}$:

   $$M_{c19_t u} = c + \Phi_1 M_{c19_{t-1}} + \ldots + \Phi_p M_{c19_{t-p}} + \omega_1 S_{c19_{t-1}} + \ldots + \omega_n S_{c19_{t-n}} + \epsilon_t \tag{2.13}$$

   where:

   - $M_{c19_t u}$ represents the market state at time t in the unrestricted model,
   - $\omega_1, ..., \omega_n$ are the additional parameters representing the influence of historical values of sentiment polarity.

   We then perform a χ^2 test which compares $M_{c19_t r}$ with $M_{c19_t u}$. If the χ^2 test value exceeds a certain critical value, we reject the null hypothesis that $S_{c19_{t-1}}$ does not Granger cause M_{c19_t}.

2. H_0: The number of COVID-19 cases, denoted by $C_{c19_{t-1}}$, does not Granger cause M_{c19_t}.

   Again, we test this null hypothesis by estimating a restricted model that excludes the historical values of $C_{c19_{t-1}}$ and an unrestricted model that includes these values:

$$M_{c19_t r} = c + \Phi_1 M_{c19_{t-1}} + ... + \Phi_p M_{c19_{t-p}} + \epsilon_t \tag{2.14}$$

$$M_{c19_t u} = c + \Phi_1 M_{c19_{t-1}} + ... + \Phi_p M_{c19_{t-p}} + \omega_1 C_{c19_{t-1}} + ... + \omega_n C_{c19_{t-n}} + \epsilon_t \tag{2.15}$$

A similar χ^2 test is then performed to compare these two models. If the test result is greater than a certain critical value, we reject the null hypothesis that $C_{c19_{t-1}}$ does not Granger cause M_{c19_t}.

3. H_0: The number of COVID-19 related deaths, denoted by $D_{c19_{t-1}}$, does not Granger cause M_{c19_t}.

   The testing of this null hypothesis involves estimating a restricted model excluding the historical values of $D_{c19_{t-1}}$ and an unrestricted model that includes these values:

$$M_{c19_t r} = c + \Phi_1 M_{c19_{t-1}} + ... + \Phi_p M_{c19_{t-p}} + \epsilon_t \tag{2.16}$$

$$M_{c19_t u} = c + \Phi_1 M_{c19_{t-1}} + ... + \Phi_p M_{c19_{t-p}} + \omega_1 D_{c19_{t-1}} + ... + \omega_n D_{c19_{t-n}} + \epsilon_t \tag{2.17}$$

   The χ^2 test is then performed to compare these two models. If the test result exceeds a certain critical value, we reject the null hypothesis that $D_{c19_{t-1}}$ does not Granger cause M_{c19_t}.

In these hypotheses and models, the variables are defined as follows:

- M_{c19_t} represents the state of the market (either market returns or prices) during the COVID-19 pandemic at time t,
- $C_{c19_{t-1}}$ represents the number of COVID-19 cases at time $t-1$,
- $D_{c19_{t-1}}$ represents the number of deaths due to COVID-19 at time $t-1$,
- $S_{c19_{t-1}}$ represents the sentiment polarity of COVID-19 conference briefings at time $t-1$.

Prior to the application of the Granger causality test, it is critical to ensure that the time series under investigation are covariance stationary. This precondition stems from the fundamental assumption of Granger causality, which mandates that the analysed signals or time series must be stationary. A stationary series is characterized by time-invariant properties such as consistent

mean, variance, and autocorrelation structure.
In order to test for stationarity, we perform Unit root tests, which are a characteristic of time series that makes the signal non-stationary. For the purpose of this study, we perform the following unit root tests: Augmented Dickey-Fuller Test (Dickey et al., 1984) and Kwiatkowski test(Shin and Schmidt, 1992).
Following the unit root tests, and subject to the satisfactory condition of stationarity for all the Granger Causality studied time series, we make use of χ^2 a test constructed with a Wald test (Gourieroux et al., 1982) and assert when values of the variable X provide statistically significant information about the evolution of the future values of the variable Y. In cases where the studied time series is non-stationary. We apply a transformation by squaring the given time series, i.e., converting the studied time series from non-stationary to stationary. See non-stationary results in 2.23.
Of course, the Granger causality is not necessarily true causality. In fact, Granger causality merely gives information about predicting abilities; it does not reveal the underlying causal link between two variables (Maziarz, 2015). According to (Granger, 1980), Granger causality, grounded on the axiom that the future cannot cause the past, primarily aids in forecasting. Its relevance is questionable if this axiom is not accepted. While not a measure of absolute causality, Granger causality indicates if one variable contains predictive information about another. The quality of the results is reliant on the sophistication of the analysis. Note that its applicability is limited to data sequences, not unique events or ultimate causes. Hence, Granger causality interpretations should remain within these constraints and not extend to philosophical or theological domains.
The application of Granger causality in this research hinges on its inherent axiom that the future cannot cause the past. It underpins the sequential investigation of this study, which explores whether the sentiment expressed in COVID-19 speeches, Granger causes subsequent stock market movements. Nevertheless, it's imperative to realize that the effectiveness of our Granger causality tests is largely contingent upon the sophistication of our analysis, particularly our ability to properly specify our model. Moreover, while a significant Granger causality may suggest that the sentiment of these speeches contains information that can predict stock market movements, it does not establish that these speeches are the only or even the primary cause of such movements. Indeed, other factors not included in our analysis could also be driving the market. Lastly, given that Granger causality is tailored to sequences of data, it's well-suited to our time-series examination of speeches and stock market data, but we must refrain from drawing conclusions about one-off speeches or unique market events based solely on Granger causality.
The hypothesis grounding this portion of our research is that investors' sentiment in the UK stock market is reflected in the UK government's daily COVID-19 updates. To prove this, we extract sentiment polarity scores from each COVID-19 government update by applying

traditional Machine Learning methods alongside advanced AI techniques on the text of the daily briefings. These scores create a time series, which we then juxtapose with the time series of the UK stock indices through correlation analysis and Granger causality tests. If the sentiment conveyed in a briefing is positive, we postulate that the briefing's impact on the market will be favourable, increasing the likelihood of a rise in index prices. We extend this analysis by comparing the sentiment scores and the stock market time series to the number of COVID-19 cases and deaths, aiming to gain further insight into the information chain.

2.7 Results and discussion

This section presents the findings from the experiments carried out throughout the present thesis to answer the research questions introduced in section 2.3. The data used comprises the computed sentiment scores of the daily COVID-19 briefings, the daily frequencies of the mentions of the word "vaccine" during the COVID-19 briefings, the computed UK market stock returns, the main UK stock indexes prices, and the COVID-19 deaths and cases time series. Section 2.5 contains further information and a description of the data-gathering method used. We compute and report the results of the correlation findings, which are supplemented by Granger causality, using the methodologies described in section 2.6.
First, we investigate the relationship between the number of COVID-19-related deaths and cases and the inferred investors' sentiment scores. This is to understand whether the calculated sentiment scores reflect the severity of the COVID-19 pandemic. Second, we examine the interrelationship between the inferred investors' sentiment scores and the UK stock index prices and returns. Finally, we look into the COVID-19 fatalities and cases and the UK stock indexes variables relationship.
It is important to note that the aforementioned investigations were broken down into three distinct time periods. We expand on this in the below section 2.7.1.

2.7.1 The COVID-19 pandemic in periods

In order to understand the pandemic evolution in relationship to the financial variable over time, we segregate the changes of this epidemic into three different time periods that we mark as periods 1, 2, and 3. Figure 2.9 presents the evolution of the UK's new daily COVID-19 deaths

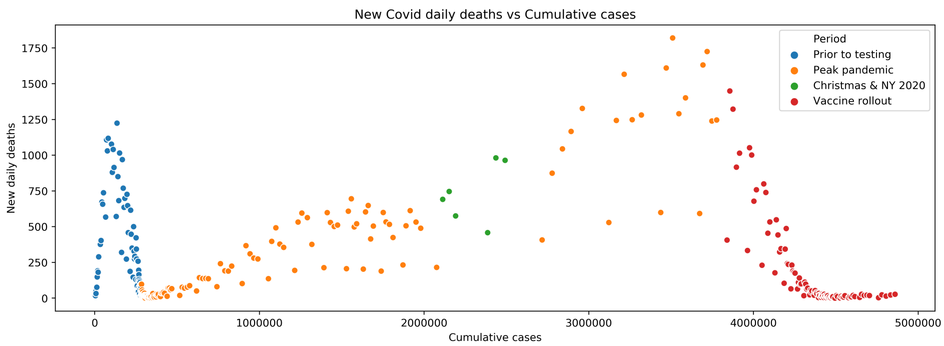


Figure 2.9: Cumulative COVID-19 cases vs New COVID-19 daily deaths

plotted against the cumulative number of new COVID-19 cases between the 3^{rd} of March 2020 and the 23^{rd} of June 2021. The three periods are as follows:

- **Period 1**: Prior to testing [March 2020 – June 2020]
  This interval (in blue) corresponds to the time of the pandemic when self-testing for

the illness was not prevalent owing to the scarcity of COVID-19 tests. In reality, the seven-day average virus tests reported exceeded 100,000 only on 30/06/2020. And, since the COVID-19 fatalities were more accurately measured, we see a substantially greater number of deaths as compared to the less precise number of cases owing to a lack of testing. This time period is distinguished by the inconsistency of the COVID-19-related cases in comparison to the number of new COVID-19 deaths. And is consequently isolated in its own temporal period.

- **Period 2**: Peak pandemic [July 2020 – January 2021]
  This timeframe (in orange) corresponds to the pandemic period when the number of cases was more accurately measured at a large scale owing to the availability of testing facilities such as home self-testing kits. In this situation, as the number of reported cases grows, so does the number of fatalities. The green data points belong to the Christmas and New Year's festive period when testing slowed owing to the holiday season. With the absence of a cure for the COVID-19 virus and with the abundance of testing, we can see a positive relationship between the number of new deaths and the cumulative number of cases. This correlation is statistically significant, especially after 10 days of contamination. See Figure A.4 in the appendix.

- **Period 3**: Vaccine roll-out [February 2021 – June 2021]
  This time period (in red) corresponds to the phase pandemic when the United Kingdom began to experience the benefit effects of the vaccine's efficacy in the form of a steady drop in the number of COVID-19 deaths despite the high number of cases that were a result of less stringent lockdown rules and more mixing. The vaccination campaign started in January 2021, and the assumption here is that the vaccine would take effect a month later, therefore, this time begins in February, a month after the vaccine campaign began in the United Kingdom.

2.7.2 Correlations of Sentiment polarity vs COVID-19 fatalities and cases

This section presents the correlations between the sentiment of the UK government's daily COVID-19 updates and the reported number of fatalities and cases. The focus is on Period 1 (Prior to testing), as described in section 2.7.1. The objective is to address the first research question highlighted in Section 2.3. Throughout the Results and Discussion section 2.7, we emphasize statistically significant findings, omitting non-significant ones for brevity.

2.7.2.1 Period 1 – Cases/Deaths vs Sentiment

Table 2.8 provides insights into these correlations. A salient observation during Period 1 is the positive relationship between the number of new daily COVID-19 cases and the sentiment scores from the UK government's updates. All correlation techniques, namely Pearson, Spearman, and Kendall, confirm this significant relationship. Notably, a stronger association emerges at $lag = 2$, suggesting that the sentiment of a government update is closely tied to the number of cases reported two days prior.

Table 2.8: Correlation of Sentiment Polarity with COVID-19 Deaths and Cases for Period 1. This table illustrates the relationship between the number of reported COVID-19 cases and deaths and the BERT-based sentiment scores of the daily UK government updates during the "Prior to testing" phase. The "lag" denotes a shift in days in the time series; e.g., "New Daily Cases ($lag = 1$)" refers to cases reported a day prior to the government statement.

Correlations	Period	Sentiment Polarity (BERT)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
New daily cases	Period 1	0.24**	0.24***	0.17***
New daily cases (lag=1)	Period 1	0.31***	0.30***	0.21***
New daily cases (lag=2)	Period 1	0.35***	0.33***	0.23***
New daily deaths	Period 1	0.27***	0.27***	0.19***
New daily deaths (lag=1)	Period 1	0.34***	0.34***	0.24***
New daily deaths (lag=2)	Period 1	0.21**	0.25***	0.17***

* $p-value < 0.1$,** $p-value \leqslant 0.05$, *** $p-value < 0.01$

Interestingly, the correlation is more pronounced for the "new daily deaths" metric at $lag = 1$. Specifically, the sentiment of a government announcement is significantly influenced by the number of fatalities reported just a day prior. This suggests that fatalities might have had a more immediate and pronounced impact on the sentiment of official announcements.

One potential interpretation of this trend is the government's intention to maintain a sense of hope and optimism in the face of rising pandemic numbers. Despite the growing number of fatalities and cases during Period 1, the government's updates leaned towards a positive tone, possibly as an effort to instil confidence and resilience among the public during these challenging times.

2.7.2.2 Period 2 & 3 – Cases/Deaths vs Sentiment

An analysis of Table 2.9 for Period 2 (Peak pandemic) indicates a positive correlation between the number of new daily COVID-19 cases and the sentiment scores of the UK government's updates, similar to Period 1. However, a notable difference emerges in this period: only at $lag = 5$ do all correlation methods (Pearson, Spearman, and Kendall) reflect a statistically significant association. This suggests that the sentiment of a UK government update is significantly influenced by the number of cases reported five days earlier, potentially due to a time lag in data processing and strategy formulation before public communication.

Table 2.9: Correlation between Sentiment Polarity and COVID-19 Cases for Period 2. The table showcases the relationship between reported COVID-19 cases and BERT-based sentiment scores from daily UK government updates during the "Peak pandemic" phase. A "lag" denotes a day shift in the time series; for instance, "New Daily Cases ($lag = 1$)" refers to cases reported a day prior to the government announcement.

Correlations	Period	Sentiment Polarity (BERT)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
New daily cases	Period 2	0.15	0.14	0.11
New daily cases (lag=1)	Period 2	0.30*	0.24	0.19*
New daily cases (lag=2)	Period 2	0.32**	0.21	0.17
New daily cases (lag=5)	Period 2	0.33**	0.32**	0.23**
New daily cases (lag=10)	Period 2	0.33**	0.19	0.12

* $p - value < 0.1$,** $p - value \leqslant 0.05$, *** $p - value < 0.01$

Interestingly, in Period 2, there's no statistically significant correlation between the number of new daily deaths and the sentiment of government updates, which is why it's not included in Table 2.9. Moreover, in Period 3 (Vaccine roll-out), neither the number of new daily COVID-19 cases nor the fatalities show a statistically significant correlation with the sentiment scores of the government's announcements. This could reflect a communication shift, focusing more on vaccination progress and less on daily pandemic metrics as a measure of success and control. It's worth highlighting that while some significant correlations emerge between the number of cases and government sentiment in Period 2, only Period 1 consistently showcases a correlation between both the number of COVID-19 cases and fatalities and the sentiment of the UK government updates.

2.7.2.3 Period 1, 2, & 3 – "Vaccine": A Control Variable in Sentiment Validation

In our endeavour to analyse sentiment through textual analysis, it's essential to ensure the robustness and validity of our sentiment scoring. The term "vaccine" served as a pivotal control variable in this validation process. Given its central role in the discourse around COVID-19 and its potential impact on sentiment, the frequency of its mentions in government briefings provided a tangible metric to cross-check and validate our sentiment calculations.

Delving into Table 2.10: In Period 1, the sporadic mentions of "vaccine" aligned with the broader global uncertainty of the pandemic's early days. However, as we transitioned to Periods 2 and 3, the term's increasing frequency in government announcements resonated with the growing optimism and anticipation surrounding vaccine development and distribution. This shift in sentiment, as mirrored by the term "vaccine", not only validates our sentiment analysis methodology but also underscores the intricate relationship between specific terms and the broader sentiment landscape.

Table 2.10: Correlation between the Word Frequency of "Vaccine" and Sentiment Polarity – Across All Periods. This table illustrates the relationship between the occurrence of the term "vaccine" in UK COVID-19 government announcements and the BERT-based sentiment scores of these updates.

Correlations	Period	Sentiment Polarity (BERT)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
Word frequency (keyword=Vaccine)	Period 1	0.08	0.05	0.04
Word frequency (keyword=Vaccine)	Period 2	0.53***	0.26*	0.20*
Word frequency (keyword=Vaccine)	Period 3	0.51***	0.51***	0.35***

* $p - value < 0.1$,** $p - value \leqslant 0.05$, *** $p - value < 0.01$

The term "vaccine" was chosen as a control variable to validate our sentiment scores. Given its pronounced impact on public sentiment during the pandemic, it served as a tangible benchmark against which the accuracy and sensitivity of our sentiment analysis could be gauged.

By tracking the frequency and sentiment associated with "vaccine" across different periods of the pandemic, we could assess how well our sentiment analysis captured prevailing public sentiments and narratives. If our sentiment scores aligned well with the expected sentiment around "vaccine" (e.g., more positive as vaccine developments progressed), it would validate our sentiment scoring methodology.

Using "vaccine" as a control variable helps ensure that our sentiment analysis is both robust and reflective of real-world narratives and sentiments. It acts as a quality check, ensuring our sentiment scores aren't merely artefacts of our analysis method but resonate with actual public sentiment.

2.7.2.4 Summary Insights—Part 1

The findings presented provide insights that directly address Research Question 1, which seeks to understand the impact of the rising number of COVID-19 deaths and cases on the sentiment of the government's daily COVID-19 statements.

Government Sentiment vs. COVID-19 Numbers:

- In **Period 1** (Prior to Testing), there's a significant positive correlation between the daily number of new COVID-19 cases and the sentiment scores of UK government updates. Interestingly, the sentiment seems to be more positive as the number of new daily cases increases. This positive correlation suggests that, in the early days of the pandemic when testing was limited, the government may have adopted a more optimistic tone in their updates despite the rise in cases.
- In **Period 2** (Peak Pandemic), this correlation shifts. Only at a lag of 5 days do we see a consistent positive correlation across all three correlation methods. This might indicate that, during the peak of the pandemic, the government's sentiment reaction to case numbers was delayed, possibly due to the evolving nature of the situation and the time taken to collate and process information.
- In **Period 3** (Vaccine roll-out), no significant correlation exists between COVID-19 numbers (cases or deaths) and government sentiment. This suggests that, as vaccines were being rolled out and a solution to the pandemic was in sight, the daily numbers might have had a diminished impact on government sentiment.

Absence of Death Correlation in Period 2:

In **Period 2**, there's no statistically significant correlation between the number of new daily deaths and the sentiment of the government updates. This could suggest that, during the pandemic's peak, the sentiment of government announcements might have been more influenced by other factors, such as economic indicators, lockdown measures, or vaccine developments, rather than the immediate toll of the pandemic.

Significance of the Term "Vaccine":

The term "vaccine" served as a pivotal control variable, and its frequency and associated sentiment serve as a temporal gauge of the pandemic's progression and the government's shifting stance. The minimal mention in **Period 1** mirrors the early days' uncertainty, whereas its increasing prominence in **Periods 2** and **3** indicates growing optimism around pandemic

control and economic revival. The positive sentiment associated with "vaccine" mentions further validates the sentiment analysis methodology and suggests a direct relationship between the term's prominence and the broader sentiment landscape.
In conclusion, the findings unequivocally address Research Question 1, illustrating that the rising number of COVID-19 cases and deaths did impact the sentiment of the government's daily statements. Specifically, in Period 1, the government appeared to maintain a positive sentiment despite the rise in cases, possibly as an approach to managing public sentiment. By Period 2, the sentiment's correlation with case numbers became delayed, suggesting a more measured governmental response to the peak of the pandemic. By Period 3, as vaccines were rolled out, daily numbers seemed to have a diminished impact on government sentiment, indicating a shift in focus towards recovery and solutions. This dynamic interplay between case numbers, deaths, and sentiment underscores the nuanced approach the government took in managing public communications during the pandemic.

2.7.3 Correlations of Sentiment polarity vs UK stock market movement

In this section, we elucidate the correlation between sentiment scores and the opening and closing prices of the FTSE stock indices across the three time periods specified in section 2.7.1. The aim is to discern the influence of the sentiment scores, derived from the government's daily COVID-19 briefings, on the UK stock market during the COVID-19 pandemic, thereby addressing research question two as presented in 2.3.

We initially directed our focus towards returns, aligning with conventional financial research methodologies. Returns, as a measure of percentage change in stock prices, are often employed for their normalization of stock performance **?**. However, our empirical findings led us to pivot our focus. Specifically, while returns did not yield significant correlations, absolute stock prices did. Such prices, untouched by normalization, can sometimes capture insights that returns might overlook Grossman and Stiglitz (1980). Thus, the decision to concentrate on stock prices was data-driven, reflecting an adaptability in our research approach.

2.7.3.1 Period – Sentiment vs Stock prices

The most salient findings are from Period 2 (Peak of the Pandemic). As observed in Table 2.11, during this period, a statistically significant positive relationship emerges between sentiment scores of government updates and the subsequent day's FTSE stock indices, both for opening and closing prices. This relationship holds across all correlation measures. An intriguing observation is the lack of a statistically significant correlation for FTSE250, hinting at potential unique factors influencing this specific index which warrants a deeper investigation.

Table 2.11: Correlations of Sentiment polarity vs FTSE prices: The table showcases the relationship between the BERT-based sentiment scores of the COVID-19 daily UK government updates during Period 2 (Peak pandemic) and the FTSE stock prices' opening and closing values on the subsequent day after a COVID-19 announcement. For example, for the FTSE100 index, the table elucidates the influence of sentiment scores from the government's COVID-19 announcements on the following day's FTSE100 opening and closing prices.

Correlations	Period	Sentiment Polarity (BERT)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
FTSE100 (Next day **Open** price)	Period 2	0.33**	0.31**	0.22**
FTSE100 (Next day **Close** price)	Period 2	0.34**	0.30*	0.21*
FTSE250 (Next day **Open** price)	Period 2	0.20	0.18	0.12
FTSE250 (Next day **Close** price)	Period 2	0.22	0.20	0.15
FTSE350 (Next day **Open** price)	Period 2	0.31*	0.30*	0.20*
FTSE350 (Next day **Close** price)	Period 2	0.32*	0.29*	0.21*
FTSE All Shares (Next day **Open** price)	Period 2	0.30*	0.30*	0.20*
FTSE All Shares (Next day **Close** price)	Period 2	0.31*	0.29*	0.20*

* $p-value < 0.1$,** $p-value \leqslant 0.05$, *** $p-value < 0.01$

2.7.3.2 Impact of 'Vaccine' Mentions on Stock Market Dynamics

In the context of validating our sentiment analysis, the term "vaccine" was earlier identified as a pivotal control variable, given its substantial influence on public sentiment. However, beyond sentiment validation, the frequency of this term also bears potential insights into market dynamics during the pandemic. The global anticipation surrounding vaccine development and deployment has had significant socio-economic implications. It was not just a medical solution, but a potential economic catalyst — a key to reopening economies, reviving businesses, and restoring investor confidence.

Given this backdrop, it becomes imperative to understand the direct impact of the term "vaccine" on stock market movements. While sentiment scores capture the broader tone of government communications, the sheer mention frequency of "vaccine" could act as a more immediate and tangible indicator for investors. Frequent mentions might signify rapid developments, imminent approvals, or roll-out strategies — all of which could have direct ramifications on the stock market. By correlating the frequency of "vaccine" mentions with stock market movements, we aim to dissect this relationship and quantify the extent to which this single term influenced investor behaviour and market dynamics during the pandemic.

The analysis of the term "vaccine" in relation to stock market movements ties directly into Research Question 2. By examining the frequency of the term "vaccine" and correlating it with stock market movements, we are essentially exploring a specific facet of this broader question. It helps discern the direct influence of a particular term within government statements on stock market dynamics, which is a subset of understanding the overall sentiment influence.

Table 2.12 reveals a noteworthy observation: there exists a robust positive correlation between the frequency of the term "vaccine" in the UK government's daily COVID-19 announcements and the subsequent day's performance of all FTSE indices, both for opening and closing prices. This association is consistently significant across the three correlation metrics, namely, Pearson, Spearman, and Kendall.

Table 2.12: Correlation between "Vaccine" Mentions and FTSE Performance. The table showcases the relationship between the frequency of "vaccine" mentions in the UK government's daily COVID-19 briefings and the next day's opening and closing prices of various FTSE indices during Period 2 (Peak pandemic). To illustrate, the row corresponding to the FTSE100 index examines how the frequency of "vaccine" mentions in a government briefing impacts the FTSE100's opening and closing prices the following day.

Correlations	Period	Word frequency (keyword=Vaccine)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
FTSE100 (Next day **Open** price)	Period 2	0.67***	0.76***	0.56***
FTSE100 (Next day **Close** price)	Period 2	0.66***	0.78***	0.57***
FTSE250 (Next day **Open** price)	Period 2	0.64***	0.78***	0.59***
FTSE250 (Next day **Close** price)	Period 2	0.63***	0.80***	0.61***
FTSE350 (Next day **Open** price)	Period 2	0.67***	0.77***	0.58***
FTSE350 (Next day **Close** price)	Period 2	0.66***	0.80***	0.58***
FTSE All Shares (Next day **Open** price)	Period 2	0.67***	0.77***	0.57***
FTSE All Shares (Next day **Close** price)	Period 2	0.66***	0.80***	0.58***

* $p-value < 0.1$,** $p-value \leqslant 0.05$, *** $p-value < 0.01$

Given the global anticipation and the potential impact of a COVID-19 vaccine, it's not surprising that its mention in government briefings would have a profound effect on the stock market. As the vaccine was heralded as the crucial instrument to curtail the pandemic, restore normalcy, and rejuvenate economies, any government communication hinting at its progress or distribution would likely be perceived as a beacon of hope. This leads to the hypothesis: the more frequent the mentions of the vaccine in government updates, the more bullish the stock market response, reflecting heightened investor optimism and confidence in economic recovery.

2.7.3.3 AstraZeneca Stock Dynamics in Response to 'Vaccine' Mentions and Its Unique Position in the Analysis

In the pursuit to answer our second research question—how does the sentiment of the government's daily COVID-19 briefings influence stock market movements—we delve into specific examples that shed light on unique market dynamics. AstraZeneca, as a prime player in the COVID-19 landscape, offers an insightful case study.
Table 2.13 showcases intriguing dynamics specific to the AstraZeneca stock in relation to the frequency of the term "vaccine" in the UK government's COVID-19 daily updates during Period 2 (Peak pandemic).

Table 2.13: Correlation of the word frequency "Vaccine" with AstraZeneca plc's Prices and Volume. This table presents correlation outcomes between the frequency of "vaccine" mentions in the UK COVID-19 daily government communications and the subsequent day's opening, closing prices, and trading volume for AstraZeneca plc during Period 2 (Peak pandemic).

Correlations	Period	Word frequency (keyword=Vaccine)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
AstraZeneca plc (Next day **Open** price)	Period 2	-0.47***	-0.57***	-0.41***
AstraZeneca plc (Next day **Close** price)	Period 2	-0.46***	-0.57***	-0.40***
AstraZeneca plc (Next day **Trading volume**)	Period 2	0.42***	0.36***	0.26***

* $p-value < 0.1$,** $p-value \leqslant 0.05$, *** $p-value < 0.01$

Being a constituent of the FTSE 100 and a direct player in COVID-19 vaccine development and distribution, AstraZeneca occupies a unique position compared to other companies. Its share price during this period wasn't solely influenced by the general market sentiment, but also by a myriad of factors tied to its direct involvement in the pandemic response. From vaccine trial outcomes and governmental approval timelines to its production and delivery capacities, each of these elements added layers of complexity to the stock's performance.
The robust negative correlation observed between the mentions of 'vaccine' and both the opening and closing prices of AstraZeneca's stock hints at the potential challenges the company faced, possibly from reported side effects Vogel and Kupferschmidt (2021). On the flip side, the positive correlation with trading volume indicates a heightened market activity, potentially driven by investors reacting to evolving news about the vaccine.
While this study primarily aims to gauge the broader FTSE market reaction to the UK government's COVID-19 briefings, AstraZeneca's inclusion was essential. It epitomizes the diverse ways the pandemic has shaped the trajectories of FTSE 100 companies. Recognizing the multi-faceted influences on AstraZeneca's stock price, this analysis does not delve deep into individual company dynamics, but instead seeks to understand market reactions from a more holistic viewpoint. This exploration into AstraZeneca's stock behaviour in relation to government briefings, offers a detailed lens into the multifaceted ways sentiment can influence

stock market movements, thereby enriching our understanding of the second research question.

2.7.3.4 Isolated Influence of Sentiment on AstraZeneca Stock Returns

Table 2.14 reveals a nuanced relationship between the sentiment scores of the government's COVID-19 updates, as gauged by the BERT model, and AstraZeneca's stock returns. Notably, the overnight returns (Rt_o) of AstraZeneca demonstrate a weak positive correlation across all correlation measures. This is particularly significant given that this is the only instance where we observed a statistically significant correlation with stock returns in our analysis. Conversely, the immediate stock returns following the briefing (Rt_f) and returns on the day of (Rt_b) the briefing did not present any notable correlation.

Table 2.14: Sentiment Polarity and AstraZeneca's Stock Performance. This table showcases correlations between the BERT-derived sentiment scores from the UK's daily COVID-19 updates and AstraZeneca's stock performance metrics during Period 2 (Peak pandemic). Specifically, it contrasts sentiment with the overnight stock returns (Rt_o), the returns in the two hours post-briefing (Rt_f), and the entire day's returns on the briefing day (Rt_b). Notably, the overnight returns (Rt_o) represent the sole significant correlation with stock returns observed in our study.

Correlations	Period	Sentiment Polarity (BERT)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
AstraZeneca plc (Rt_o)	Period 2	0.18***	0.14*	0.09*
AstraZeneca plc (Rt_f)	Period 2	-0.08	-0.06	-0.04
AstraZeneca plc (Rt_{be})	Period 2	-0.10	-0.13	-0.09*

* $p-value < 0.1$,** $p-value \leqslant 0.05$, *** $p-value < 0.01$

Several interpretations emerge from these findings. For example, the market, or a subset of investors, might require time to assimilate and respond to the sentiment expressed in government briefings. Additionally, AstraZeneca's stock is potentially influenced by a myriad of factors beyond just government sentiment. Elements like clinical trial outcomes, international distribution agreements, and vaccine production rates are just a few of the multifaceted influences.

In summary, while sentiment scores from government updates offer a unique lens into potential stock behaviour, the intricate dynamics of AstraZeneca reiterate the complex nature of stock market reactions during unprecedented events like a pandemic. This correlation serves as a reminder that while sentiment provides insights, a holistic approach that encompasses multiple influences is crucial to decipher stock behaviour, especially for stocks like AstraZeneca that are at the heart of the pandemic discourse. The analysis was strategically broken down into three periods, capturing different phases of the pandemic. The most pronounced findings emerged from Period 2, which represented the peak of the pandemic.

2.7.3.5 Summary Insights—Part 2

The overarching goal of this section was to uncover the relationship between sentiment scores sourced from the government's daily COVID-19 briefings and the FTSE stock indices across three distinct periods. This exploration directly seeks to address Research Question 2: How does the sentiment of the government's daily COVID-19 briefings influence stock market movements?. The most pronounced findings emerged from Period 2, which represented the peak of the pandemic.

- **Stock Prices vs. Returns:** Traditional financial research often gravitates towards returns. However, this study's empirical observations emphasized absolute stock prices, revealing that they might capture nuances potentially missed by normalized returns.
- **Significant Findings during Peak Pandemic:** In Period 2, a statistically significant positive relationship emerged between sentiment scores and both the opening and closing prices of various FTSE indices. Curiously, FTSE250 remained an outlier, lacking a significant correlation and suggesting unique dynamics affecting this index. A potential explanation is provided in the conclusion section 2.8.
- **The 'Vaccine' Factor:** Beyond being a pivotal control variable in sentiment analysis, the frequency of the term "vaccine" in government briefings was found to profoundly influence market dynamics. Its regular mention possibly signalled rapid vaccine-related developments, acting as an indicator for investors.
- **Nuanced Influences on AstraZeneca Stock Returns:** Sentiment scores had a weak but positive correlation with AstraZeneca's overnight returns. This stood out as a unique observation, being the only significant correlation with stock returns throughout the study.
- **Interpretative Highlights for AstraZeneca:** AstraZeneca's stock movements can be attributed to a multitude of factors, including but not limited to government sentiment.

While sentiment scores provide a valuable lens into the potential behaviour of stocks, the intricate dynamics seen, especially with stocks like AstraZeneca, emphasize the importance of multifaceted analysis. The insights derived from this section contribute richly to the understanding of Research Question 2, shedding light on the myriad ways government communications during a pandemic can sway stock market movements.

2.7.4 Correlations of COVID-19 fatalities and cases vs UK stock market movement

In this section, we present the correlation analysis between the reported number of COVID-19 fatalities and cases versus the opening and closing prices of the FTSE stock index during the three time periods indicated in section 2.7.1. The aim is to understand the influence of COVID-19's progression on the UK stock market during the pandemic, addressing research question three listed in Section 2.3.

Research question 3 examines the potential influence of COVID-19-related fatalities and cases in the UK on the movement of the UK stock market. While at first glance, this question might seem tangential to sentiment theory, which is central to research questions 1 and 2, its inclusion serves a crucial purpose. By addressing this question, we aim to ensure that the observed relationship between COVID-19 briefings and stock market performance, as explored in questions 1 and 2, is not merely a function of the number of cases or fatalities. In essence, we're controlling for the possibility that the stock market's reactions are rooted in the tangible progress of the pandemic (as reflected by cases and fatalities) rather than the sentiment of government announcements. If a significant relationship emerges between COVID-19 cases/fatalities and stock market performance in our analysis for question 3, it could suggest that the correlations observed in questions 1 and 2 might be influenced or even overshadowed by the sheer number of cases/fatalities rather than the sentiment conveyed in government announcements. Subsequently, we discuss the influence of the number of reported COVID-19 fatalities and cases on the UK stock market during the COVID-19 epidemic, providing potential explanations for observed patterns.

2.7.4.1 Correlation between COVID-19 Cases and FTSE Prices

- **Overall Periods (Table 2.15):** there was a positive and statistically significant correlation between the lagged reported number of new COVID-19 cases and both the opening and closing prices of all FTSE stock indexes. The relationship was especially strong for the FTSE250. The observed correlation suggests that the stock market was reacting to more than just the raw numbers of the pandemic. Other factors, such as governmental and monetary interventions, the anticipation of economic recoveries, and the sentiment surrounding these numbers, likely played a role. The resilience of the FTSE250 might indicate that medium-sized enterprises were either beneficiaries of certain pandemic-related trends or were more agile in adapting to the changing landscape. Further details on this is presented in the conclusion section 2.8.

- **Period 1 (Pre-testing, Table 2.17):** The number of lagged COVID-19 cases negatively correlated with FTSE stock indexes. This period marked the onset of the pandemic, characterized by uncertainty and fear. The stock market's decline could be attributed to concerns about the unknown trajectory of the virus, potential lockdowns, and the ensuing economic implications.

- **Period 2 (Peak pandemic, Table 2.18):** The number of lagged COVID-19 cases had a positive relationship with the FTSE stock indexes, especially the FTSE250. This phase might have seen a level of market acclimatization to the pandemic's realities. The positive correlation suggests that investors were possibly banking on robust governmental interventions, potential treatments, and an eventual return to normalcy. The FTSE250's performance again points to the adaptability of medium-sized businesses during this tumultuous period. This also indicates that stock prices rebounded and stabilized in Period 2 after the significant decline in Period 1. This trend is further visualized in Figure 2.1.

- **Period 3 (Vaccine roll-out, Table 2.19):** Similar to Period 1, the number of lagged COVID-19 cases negatively influenced the FTSE stock indexes. The greater the number of cases, the lower the UK FTSE prices. All correlation measures indicate this significant negative association. The roll-out of vaccines, while promising, also brought to light challenges like distribution bottlenecks, vaccine hesitancy, and concerns about new variants. The market might have perceived these as indicators of a prolonged path to full economic recovery.

2.7.4.2 Correlation between COVID-19 Deaths and FTSE Prices

- **Overall Periods (Table 2.16):** There was a negative correlation between the lagged number of COVID-19 deaths and FTSE indexes. The death toll represents a more direct and tangible consequence of the pandemic, potentially exerting a more pronounced psychological impact on investors. This could have influenced market sentiments more consistently.

- **Period 1 (Pre-testing, Table 2.20):** The number of lagged COVID-19 deaths negatively correlated with FTSE stock indexes. As fatalities began to rise, the tangible reality and severity of the pandemic became evident. This grim milestone could have compounded initial market fears, pushing stock prices down.

- **Period 2 (Peak pandemic, Table 2.21):** The number of lagged COVID-19 deaths showed a positive correlation with the FTSE stock indexes. This indicates a recovery and stabilization of stock prices in Period 2, following their significant decline in Period 1. This trend is further visualized in Figure 2.1. While counterintuitive, this suggests that the market had somewhat decoupled from the immediate emotional response to the death toll. Instead, it might have been focused on broader recovery narratives, interventions, and long-term economic implications.

- **Period 3 (Vaccine roll-out, Table 2.22):** The number of lagged COVID-19 deaths had a strong negative relationship with FTSE stock indexes. The resurgence of a negative correlation during vaccine roll-out could have been due to concerns about the vaccines' effectiveness against emerging variants or challenges in achieving global herd immunity, given the continuing fatalities.

Table 2.15: Correlation of Daily COVID-19 Cases with Next Day's FTSE Stock Prices Across All Periods. This table illustrates how the reported number of new cases correlates with the subsequent day's opening and closing stock prices, considering a one-day lag.

Correlations	Period	COVID-19 New Daily Cases (lag=1)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
FTSE100 (**Open** price)	All periods	0.22***	0.23***	0.12***
FTSE100 (**Close** price)	All periods	0.23***	0.24***	0.12***
FTSE250 (**Open** price)	All periods	**0.30*** **	**0.39*** **	**0.21*** **
FTSE250 (**Close** price)	All periods	**0.30*** **	**0.39*** **	**0.22*** **
FTSE350 (**Open** price)	All periods	0.25***	0.25***	0.12***
FTSE350 (**Close** price)	All periods	0.25***	0.25***	0.13***
FTSE All Shares (**Open** price)	All periods	0.25***	0.25***	0.13***
FTSE All Shares (**Close** price)	All periods	0.25***	0.25***	0.13***

* $p - value < 0.1$,** $p - value \leqslant 0.05$, *** $p - value < 0.01$

Table 2.16: Correlation of Daily COVID-19 Deaths with Next Day's FTSE Stock Prices Across All Periods. The table demonstrates the relationship between the reported number of deaths and the following day's stock prices, considering a one-day lag.

Correlations	Period	COVID-19 New Daily Deaths (lag=1)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
FTSE100 (**Open** price)	All periods	-0.01	-0.14***	-0.09***
FTSE100 (**Close** price)	All periods	-0.02	-0.13***	-0.08***
FTSE250 (**Open** price)	All periods	-0.00	-0.16***	-0.11***
FTSE250 (**Close** price)	All periods	0.00	-0.16***	-0.11***
FTSE350 (**Open** price)	All periods	-0.01	-0.16***	-0.10***
FTSE350 (**Close** price)	All periods	-0.01	-0.15***	-0.09***
FTSE All Shares (**Open** price)	All periods	-0.01	-0.16***	-0.10***
FTSE All Shares (**Close** price)	All periods	-0.01	-0.15***	-0.10***

* $p - value < 0.1$,** $p - value \leqslant 0.05$, *** $p - value < 0.01$

Table 2.17: Correlation of Daily COVID-19 Cases with Next Day's FTSE Stock Prices During Period 1 (Pre-testing). This table presents how the reported daily cases relate to the subsequent day's stock prices, using a one-day lag.

Correlations	Period	COVID-19 New Daily Cases (lag=1)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
FTSE100 (**Open** price)	Period 1	-0.28***	-0.42***	-0.32***
FTSE100 (**Close** price)	Period 1	-0.32***	-0.43***	-0.32***
FTSE250 (**Open** price)	Period 1	-0.30***	-0.39***	-0.28***
FTSE250 (**Close** price)	Period 1	-0.32***	-0.40***	-0.29***
FTSE350 (**Open** price)	Period 1	-0.29***	-0.41***	-0.31***
FTSE350 (**Close** price)	Period 1	-0.32***	-0.42***	-0.32***
FTSE All Shares (**Open** price)	Period 1	-0.29***	-0.42***	-0.32***
FTSE All Shares (**Close** price)	Period 1	-0.32***	-0.42***	-0.32***

$^{*}\ p-value < 0.1,^{**}\ p-value \leqslant 0.05,\ ^{***}\ p-value < 0.01$

Table 2.18: Correlation of Daily COVID-19 Cases with Next Day's FTSE Stock Prices During Period 2 (Peak pandemic). This table showcases the relationship between daily reported cases and the subsequent day's stock prices, accounting for a one-day lag.

Correlations	Period	COVID-19 New Daily Cases (lag=1)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
FTSE100 (**Open** price)	Period 2	0.61***	0.38***	0.19***
FTSE100 (**Close** price)	Period 2	0.62***	0.40***	0.22***
FTSE250 (**Open** price)	Period 2	**0.82*** **	**0.76*** **	**0.56*** **
FTSE250 (**Close** price)	Period 2	**0.82*** **	**0.77*** **	**0.58*** **
FTSE350 (**Open** price)	Period 2	0.67***	0.40***	0.22***
FTSE350 (**Close** price)	Period 2	0.68***	0.42***	0.25***
FTSE All Shares (**Open** price)	Period 2	0.68***	0.41***	0.23***
FTSE All Shares (**Close** price)	Period 2	0.69***	0.43***	0.26***

$^{*}\ p-value < 0.1,^{**}\ p-value \leqslant 0.05,\ ^{***}\ p-value < 0.01$

Table 2.19: Correlation of Daily COVID-19 Cases with Next Day's FTSE Stock Prices During Period 3 (Vaccine roll-out). The table indicates how daily cases correspond with the next day's stock prices, factoring in a one-day lag.

Correlations	Period	COVID-19 New Daily Cases (lag=1)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
FTSE100 (**Open** price)	Period 3	-0.25***	-0.24**	-0.21***
FTSE100 (**Close** price)	Period 3	-0.27***	-0.27***	-0.23***
FTSE250 (**Open** price)	Period 3	-0.32***	-0.33***	-0.27***
FTSE250 (**Close** price)	Period 3	-0.32***	-0.35***	-0.28***
FTSE350 (**Open** price)	Period 3	-0.27***	-0.25**	-0.22***
FTSE350 (**Close** price)	Period 3	-0.28***	-0.28***	-0.24***
FTSE All Shares (**Open** price)	Period 3	-0.27***	-0.25**	-0.22***
FTSE All Shares (**Close** price)	Period 3	-0.28***	-0.27***	-0.24***

* $p-value < 0.1$,** $p-value \leqslant 0.05$, *** $p-value < 0.01$

Table 2.20: Correlation of Daily COVID-19 Deaths with Next Day's FTSE Stock Prices During Period 1 (Pre-testing). The table elucidates the relationship between reported daily deaths and the subsequent day's stock prices, given a one-day lag.

Correlations	Period	COVID-19 New Daily Deaths (lag=1)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
FTSE100 (**Open** price)	Period 1	-0.14	-0.26**	-0.21***
FTSE100 (**Close** price)	Period 1	-0.17	-0.26**	-0.22***
FTSE250 (**Open** price)	Period 1	-0.12	-0.23*	-0.18**
FTSE250 (**Close** price)	Period 1	-0.14	-0.24**	-0.19**
FTSE350 (**Open** price)	Period 1	-0.14	-0.25**	-0.20**
FTSE350 (**Close** price)	Period 1	-0.17	-0.26**	-0.22***
FTSE All Shares (**Open** price)	Period 1	-0.14	-0.25**	-0.21***
FTSE All Shares (**Close** price)	Period 1	-0.17	-0.26**	-0.22***

* $p-value < 0.1$,** $p-value \leqslant 0.05$, *** $p-value < 0.01$

Table 2.21: Correlation of Daily COVID-19 Deaths with Next Day's FTSE Stock Prices During Period 2 (Peak pandemic). The table emphasizes the correlation between daily deaths and the next day's stock prices, accounting for a one-day lag.

Correlations	Period	COVID-19 New Daily Deaths (lag=1)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
FTSE100 (**Open** price)	Period 2	0.72***	0.60***	0.40***
FTSE100 (**Close** price)	Period 2	0.72***	0.61***	0.41***
FTSE250 (**Open** price)	Period 2	**0.89*** **	**0.75*** **	**0.54*** **
FTSE250 (**Close** price)	Period 2	**0.82*** **	**0.76*** **	**0.55*** **
FTSE350 (**Open** price)	Period 2	0.76***	0.61***	0.41***
FTSE350 (**Close** price)	Period 2	0.76***	0.62***	0.42***
FTSE All Shares (**Open** price)	Period 2	0.77***	0.61***	0.42***
FTSE All Shares (**Close** price)	Period 2	0.78***	0.63***	0.42***

* $p-value < 0.1$,** $p-value \leqslant 0.05$, *** $p-value < 0.01$

Table 2.22: Correlation of Daily COVID-19 Deaths with Next Day's FTSE Stock Prices During Period 3 (Vaccine roll-out). This table describes how reported daily deaths influence the subsequent day's stock prices, with a one-day lag.

Correlations	Period	COVID-19 New Daily Deaths (lag=1)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
FTSE100 (**Open** price)	Period 3	-0.72***	-0.85***	-0.65***
FTSE100 (**Close** price)	Period 3	-0.73***	-0.85***	-0.67***
FTSE250 (**Open** price)	Period 3	-0.72***	-0.85***	-0.67***
FTSE250 (**Close** price)	Period 3	-0.72***	-0.84***	-0.65***
FTSE350 (**Open** price)	Period 3	-0.72***	-0.85***	-0.67***
FTSE350 (**Close** price)	Period 3	-0.73***	-0.86***	-0.67***
FTSE All Shares (**Open** price)	Period 3	-0.72***	-0.86***	-0.67***
FTSE All Shares (**Close** price)	Period 3	-0.73***	-0.86***	-0.67***

* $p-value < 0.1$,** $p-value \leqslant 0.05$, *** $p-value < 0.01$

2.7.4.3 Conclusion and Significance of the Analysis for Research question 3

Research Question 3 aimed to discern if the number of COVID-19-related fatalities and cases in the UK could elucidate the trends in the UK stock market during various pandemic phases. The findings suggest a noteworthy relationship:

During the pandemic's onset (Period 1), an increase in cases and deaths corresponded with a decline in stock market prices. However, in the pandemic's peak (Period 2), the stock market exhibited an unexpected resilience, with prices stabilizing and even increasing alongside rising cases and fatalities. By the vaccine roll-out phase in Period 3, despite the promise of vaccines, the market displayed sensitivity to the challenges surrounding vaccine distribution, fears of new virus variants, and the potential long-term economic repercussions.

These findings underscore that while there is a relationship between the pandemic's progression (in terms of cases and deaths) and stock market movements, it isn't straightforward. Other factors, such as governmental responses, market sentiments, and evolving economic perspectives, heavily influence stock market trends. Thus, while the number of cases and deaths provides some insight into stock market movements, it's one piece of a larger puzzle. The sentiment and contextual environment during each phase played a pivotal role, indicating that the stock market's reactions were not solely driven by the pandemic's direct impacts, but also by the broader socio-economic context.

In the context of this study, the significant correlations found in Research Question 3 emphasize the need to consider the direct impacts of the pandemic alongside other influencing factors when examining the relationship between government announcements, market sentiments, and stock market performance.

2.7.5 Granger Causality analysis

After observing significant correlations in the preceding analysis, it becomes essential to probe the causative relationships between the variables of interest. The Granger causality tests provide a mechanism to understand if one variable can predict another, offering insights beyond mere correlation. In the complex tapestry of financial market behaviour amidst a global health crisis, the dissemination and reception of information play a pivotal role. Granger causality tests offer a means to decipher this complexity, shedding light on the causal pathways and sequencing of information exchange.

- **From COVID-19 Numbers to Sentiment:**
  Research Question 1 postulated a relationship between the sentiment of government announcements and stock market performance. In the context of information flow, it's crucial to determine if the unfolding pandemic scenario, as represented by COVID-19 numbers, 'Granger causes' or influences the sentiment of government announcements. This would suggest that the tangible progression of the pandemic shapes the narrative tone of governmental communications.

- **From Sentiment to Stock Market Performance:**
  Research Question 2 delved into the connection between the sentiment scores from the government's COVID-19 updates and FTSE index prices. Having established a potential causative relationship from COVID-19 numbers to sentiment, it becomes imperative to understand if this sentiment in turn 'Granger causes' stock market movements. This would imply that market dynamics are significantly swayed by the tone and content of government communications, forming the next link in our chain of information flow.

- **Direct Influence of COVID-19 Numbers on Stock Prices:**
  For Research Question 3, while the earlier correlations signposted a relationship between the pandemic's progression and stock market reactions, Granger causality can pinpoint if the number of COVID-19 cases and deaths directly influence stock prices, bypassing the sentiment intermediary. This would be akin to the market reacting directly to raw data about the pandemic's progression, underscoring the magnitude of its impact.

The above analysis can be visualized as a circular mechanism of information flow, reminiscent of information theory: COVID-19 numbers influence government sentiment, which in turn affects stock prices, while also checking if there's a direct pathway from COVID-19 numbers to stock prices.
Before setting this mechanism in motion, it's vital to ensure that our time series data adheres to the principles of stationarity. Non-stationary data can misguide our understanding of causality. The adherence to stationarity is verified using the methods outlined in the methodology section 2.6.9, and the outcomes are consolidated in Table 2.23.

Table 2.23: Unit Root Test Results – It displays the unit root test results for the time series under investigation, thereby testing for stationarity before performing the Granger Causality test, as per the methodology Section 2.6.9. Based on the results from both the Augmented Dickey-Fuller (ADF) and KPSS tests presented in Table 2.23, we can confirm that all the series under consideration are stationary. This ensures that we can proceed with the Granger causality tests without the risk of encountering spurious results due to non-stationary data. The stationarity of the series provides a solid foundation for the subsequent analyses, ensuring that the relationships we identify are genuine and not a mere artifact of non-stationarity.

Variables	ADF Test				KPSS Test	
	Unchanged		Transformed*		Unchanged	
	Statistics	p-value	Statistics	p-value	Statistics	p-value
Sentiment BERT	-5.86	0.00	–	–	–	–
New Cases	-3.43	0.04	-3.89	0.01	–	–
New Deaths	-2.42	0.36	-3.21	0.08	–	–
FTSE100 (P_{ro}**)	-2.31	0.16	–	–	1.54	0.01
FTSE250 (P_{ro})	-1.40	0.58	–	–	1.80	0.01
FTSE350 (P_{ro})	-2.07	0.25	–	–	1.63	0.01
FTSE All-Shares (P_{ro})	-1.97	0.29	–	–	1.65	0.01
FTSE100 (P_{rc}***)	-1.31	0.62	–	–	1.54	0.01
FTSE250 (P_{rc})	-2.03	0.27	–	–	1.80	0.01
FTSE350 (P_{rc})	-1.27	0.64	–	–	1.63	0.01
FTSE All-Shares (P_{rc})	-1.23	0.65	–	–	1.65	0.01
FTSE100 (R_t****)	-18.83	0.00	–	–	–	–
FTSE250 (R_t)	-17.64	0.00	–	–	–	–
FTSE350 (R_t)	-18.08	0.00	–	–	–	–
FTSE All-Shares (R_t)	-18.08	0.00	–	–	–	–
AZN PLC (P_{ro})	-2.93	0.04	–	–	–	–
AZN PLC (P_{rc})	-3.09	0.02	–	–	–	–
AZN PLC (R_t)	-18.97	0.00	–	–	–	–
AZN PLC (V*****)	-3.33	0.01	–	–	–	–

* *Squared*, ** $Price(Open)$, *** $Price(Close)$, **** *Return* , ***** $TradingVolume$

2.7.5.1 Granger causality for the FTSE100 and Sentiment polarity

The Granger causality tests delved into the potential of sentiment polarity from the government's COVID-19 updates as a harbinger for FTSE100's opening and closing prices. It's pivotal to stress that "Granger causality" here pinpoints predictive precedence between time series, rather than a direct cause-and-effect linkage. Table 2.24 showcases the outcomes during the pandemic's peak (Period 2), a time frame characterized by palpable correlations as spotlighted in Table 2.11. In the Granger causality paradigm:

- A left arrow ($\leftarrow$) connotes that the second variable has predictive precedence over the first.
- A bidirectional arrow ($\rightleftarrows$) signals a mutual predictive relationship: both time series seem to hold foresight over each other's future values.
- A dash (-) signifies an absence of any Granger causality between the involved series.

Table 2.24: Granger causality for the FTSE100 Sentiment polarity – Period 2. It displays the Granger Causality test results for the opening and closing prices of the FTSE100 and the UK government updates' Sentiment polarity time series in Period 2 (Peak of the pandemic).

Null Hypothesis - **Period 2**	Lag	Wald Test		
		$Chi - square$	$p - value$	Direction
Sentiment **does not** $Granger - cause$ FTSE100 (open)	$(S_{lag=1})$	4.2412**	0.03	$\leftarrow$
Sentiment **does not** $Granger - cause$ FTSE100 (close)	$(S_{lag=1})$	5.4261**	0.01	$\leftarrow$
Sentiment **does not** $Granger - cause$ FTSE100 (open)	$(S_{lag=10})$	119.8843***	0.00	$\rightleftarrows$
Sentiment **does not** $Granger - cause$ FTSE100 (close)	$(S_{lag=10})$	72.1222***	0.00	$\rightleftarrows$

* $p - value < 0.1$,** $p - value \leqslant 0.05$, *** $p - value < 0.01$

From Table 2.24:

- During Period 2, sentiment polarity with a one-day lag ($S_{lag=1}$) doesn't appear to possess foresight over the FTSE100 prices. Interestingly, the FTSE100 prices might have predictive precedence over sentiment polarity. This could hint at a complex interplay where market movements might be anticipating or reacting to other external factors, which subsequently reflect in government sentiments.
- With a ten-day lag ($S_{lag=10}$), a bidirectional Granger causality emerges. This intriguingly suggests that not only does sentiment polarity from a government update ten days prior hold predictive power over FTSE100's prices but also that market movements might have some predictive capacity over government sentiment released ten days later. While this mutual foresight is statistically robust, its real-world interpretation is intricate. It could allude to a nuanced feedback loop where market reactions and governmental communications are intertwined in a dance of mutual influence. Nevertheless, given the unexpected nature of this result, it is presented primarily for academic rigour and to inspire further probing into this intricate relationship.

2.7.5.2 Granger causality for the FTSE100 and the term frequency "Vaccine"

The Granger causality tests sought to discern if the frequency of the term "Vaccine" in the COVID-19 UK government updates offers predictive insights into the FTSE100's opening and closing prices. The rationale behind the emphasis on the FTSE100 mirrors that elucidated in the exposition of results in Table 2.24. The Granger causality tests are spotlighted during the pandemic's zenith (Period 2), aligning with the epoch that revealed statistically significant correlations as documented in Table 2.12. Table 2.25 elucidates the outcomes of the Granger causality test.Observations from Table 2.25 reveal:

- In Period 2, the term frequency "Vaccine" with a one-day lag ($V_{lag=1}$) doesn't offer predictive value over the FTSE100 prices. Notably, the Granger causality is solely from the FTSE100 prices towards the term frequency "Vaccine". This could indicate that market movements may, directly or indirectly, influence subsequent government communications or the emphasis on the vaccine narrative.

- With a ten-day lag ($V_{lag=10}$), there emerges a bidirectional Granger causality. This suggests that the frequency of the term "Vaccine" in a government update ten days prior might offer foresight into FTSE100's prices. Conversely, the market's dynamics might provide insights into the vaccine narrative ten days hence. While statistically valid, this bidirectional causality might be somewhat counterintuitive, especially in the direction of stock prices predicting future government communications. Such findings could be emblematic of deeper, latent factors influencing both the stock market and government narratives, leading to this observed interplay. Given the complexities and potentially unexpected nature of this result, it's featured primarily for academic completeness and to inspire further nuanced explorations into the relationship between market dynamics and the vaccine discourse.

Table 2.25: Granger causality for the FTSE100 vs COVID-19 death and cases – Period 2. It showcases the Granger Causality test outcomes for the FTSE100's opening and closing prices vis-à-vis the frequency of the term "Vaccine" in COVID-19 UK government updates during the pandemic's peak (Period 2).

Null Hypothesis - **Period 2**	Lag	Wald Test		
		$Chi-square$	$p-value$	Direction
"Vaccine" **does not** $Granger-cause$ FTSE100 (open)	($V_{lag=1}$)	14.6205***	0.00	$\leftarrow$
"Vaccine" **does not** $Granger-cause$ FTSE100 (close)	($V_{lag=1}$)	19.9195***	0.00	$\leftarrow$
"Vaccine" **does not** $Granger-cause$ FTSE100 (open)	($V_{lag=10}$)	25.2945***	0.00	$\rightleftarrows$
"Vaccine" **does not** $Granger-cause$ FTSE100 (close)	($V_{lag=10}$)	69.6700***	0.00	$\rightleftarrows$

* $p-value < 0.1$,** $p-value \leqslant 0.05$, *** $p-value < 0.01$

2.7.5.3 Granger Causality between FTSE100 and COVID-19 New Deaths and Cases for Period 1

The Granger causality tests were conducted to assess if the number of reported COVID-19 cases and deaths could provide predictive insights into the FTSE100's opening and closing prices during Period 1 (Pre-testing). As with the previous analyses, the spotlight is on the FTSE100 stock index due to its analogous behaviour, as depicted in Tables 2.24 and 2.25. The Granger causality tests correspond to Period 1, a juncture that unveiled statistically significant correlations, as captured in Tables 2.15 and 2.16. Table 2.26 distils the outcomes of the Granger causality test. Salient observations from Table 2.26 are:

- During Period 1, the number of COVID-19 cases, whether lagged by one day ($C_{lag=1}$) or ten days ($C_{lag=10}$), is not predictive of the FTSE100's opening or closing prices. Yet, intriguingly, the FTSE100 prices appear to offer predictive value over the reported number of new COVID-19 cases. While this may appear counterintuitive at first blush, it suggests that market dynamics could, directly or indirectly, reflect public anticipation or reaction to emerging health trends, even before they're officially reported.
- The number of reported COVID-19 deaths, when lagged by ten days ($D_{lag=10}$), reveals a bidirectional Granger causality with the FTSE100 prices. This intimates that not only can the number of deaths from ten days prior offer predictive insights into FTSE100 prices, but also the converse. The logical conundrum here, particularly the notion that stock prices could foretell future reported deaths, implies that there might be latent variables or dynamics at play, affecting both the stock market and reported health metrics. It's pivotal to handle such results with caution, recognizing that while statistical models can unveil such patterns, they do not necessarily confer direct causation or practical interpretation.

Table 2.26: Granger causality between FTSE100 prices and reported numbers of COVID-19 cases and deaths for Period 1.

Null Hypothesis - **Period 1**	*Lag*	Wald Test		
		$Chi-square$	$p-value$	Direction
Nb cases **does not** $Granger-cause$ FTSE100 (Open)	($C_{lag=1}$)	10.0867***	0.00	$\leftarrow$
Nb cases **does not** $Granger-cause$ FTSE100 (Close)	($C_{lag=1}$)	12.4075***	0.00	$\leftarrow$
Nb cases **does not** $Granger-cause$ FTSE100 (Open)	($C_{lag=10}$)	26.6129***	0.00	$\leftarrow$
Nb cases **does not** $Granger-cause$ FTSE100 (Close)	($C_{lag=10}$)	24.1751***	0.00	$\leftarrow$
Nb deaths **does not** $Granger-cause$ FTSE100 (Open)	($D_{lag=1}$)	0.2524	0.61	–
Nb deaths **does not** $Granger-cause$ FTSE100 (Close)	($D_{lag=1}$)	0.1998	0.65	–
Nb deaths **does not** $Granger-cause$ FTSE100 (Open)	($D_{lag=10}$)	38.9946	0.90	$\leftarrow$
Nb deaths **does not** $Granger-cause$ FTSE100 (Close)	($D_{lag=10}$)	16.4071*	0.08	$\rightleftarrows$

* $p-value < 0.1$,** $p-value \leqslant 0.05$, *** $p-value < 0.01$

2.7.5.4 Granger Causality between FTSE100 and COVID-19 New Deaths and Cases for Period 2

The Granger causality tests were conducted to investigate if the reported number of COVID-19 cases and deaths could provide predictive insights into the FTSE100's opening and closing prices for Period 2, which is designated as the "Peak of the pandemic". As has been the approach in previous analyses (Tables 2.24, 2.25, and 2.26), the focus remains on the FTSE100 stock index. For a glimpse into the statistically significant correlations identified for this period, we refer to Tables 2.15 and 2.16. The Granger causality results for this period are consolidated in Table 2.27. Salient observations from Table 2.27 are:

- During Period 2, the number of reported COVID-19 cases, when lagged by one day ($C_{lag=1}$) and ten days ($C_{lag=10}$), emerges as a significant predictor for the FTSE100's opening prices. However, only the one-day lag ($C_{lag=1}$) offers predictive value over the FTSE100's closing prices. This indicates that near-term reported COVID-19 case trends might influence stock market movements more immediately than longer-term trends.
- Concerning the number of reported COVID-19 deaths, the FTSE100 prices, both opening and closing, appear to provide predictive insights over the number of reported deaths when lagged by one day ($D_{lag=1}$) and ten days ($D_{lag=10}$). This finding reiterates a pattern observed in the previous period, suggesting that market dynamics might reflect public sentiments, concerns, or reactions to health trends, possibly even before they're documented or officially acknowledged.

Table 2.27: Granger causality for Sentiment polarity vs COVID-19 death and cases – Period 2. It displays the Granger Causality test results for the opening and closing prices of the FTSE100 and the number of COVID-19 new cases and deaths in Period 2 (Peak of the pandemic).

Null Hypothesis - **Period 2**	*Lag*	Wald Test		
		$Chi-square$	$p-value$	Direction
Nb cases **does not** *Granger – cause* FTSE100 (Open)	($C_{lag=1}$)	3.7991**	0.05	→
Nb cases **does not** *Granger – cause* FTSE100 (Close)	($C_{lag=1}$)	3.7606**	0.05	→
Nb cases **does not** *Granger – cause* FTSE100 (Open)	($C_{lag=10}$)	18.9528**	0.04	→
Nb cases **does not** *Granger – cause* FTSE100 (Close)	($C_{lag=10}$)	13.7508	0.18	–
Nb deaths **does not** *Granger – cause* FTSE100 (Open)	($D_{lag=1}$)	15.5653***	0.00	←
Nb deaths **does not** *Granger – cause* FTSE100 (Close)	($D_{lag=1}$)	14.6400***	0.00	←
Nb deaths **does not** *Granger – cause* FTSE100 (Open)	($D_{lag=10}$)	26.9884***	0.00	←
Nb deaths **does not** *Granger – cause* FTSE100 (Close)	($D_{lag=10}$)	19.5921***	0.00	←

* $p-value < 0.1$,** $p-value \leqslant 0.05$, *** $p-value < 0.01$

2.7.5.5 Granger Causality between FTSE100 and COVID-19 New Deaths and Cases for Period 3

For the "Vaccine roll-out" period, designated as Period 3, we again conducted Granger causality tests to determine whether reported COVID-19 cases and deaths had predictive power over the FTSE100's opening and closing prices. Keeping in line with our previous analyses (Tables 2.24, 2.25, 2.26, and 2.27), the spotlight remains on the FTSE100 stock index. Relevant correlation analyses for this period can be revisited in Tables 2.15 and 2.16. The Granger causality results for this phase are documented in Table 2.28. From the findings presented in Table 2.28, we observe:

- The number of reported COVID-19 cases, when lagged by one day ($C_{lag=1}$) and ten days ($C_{lag=10}$), doesn't possess strong predictive power over the FTSE100's opening prices for this period. However, an intriguing exception is seen when the ten-day lagged new COVID-19 cases time series appears to exhibit a bidirectional Granger causality with the FTSE100's closing prices.
- For the number of reported COVID-19 deaths, there's a noticeable bidirectional Granger causality pattern between the deaths time series and the FTSE100 prices (both opening and closing), when the data is lagged by either one day ($D_{lag=1}$) or ten days ($D_{lag=10}$). This suggests that during the vaccine roll-out period, the stock market movements were deeply intertwined with the health scenario, possibly reflecting public sentiments or concerns.

Table 2.28: Granger causality for Sentiment polarity vs COVID-19 death and cases – Period 3. It displays the Granger Causality test results for the opening and closing prices of the FTSE100 and the number of COVID-19 new cases and deaths in Period 3 (Vaccine roll-out)

Null Hypothesis - **Period 3**	*Lag*	Wald Test		
		$Chi-square$	$p-value$	Direction
Nb cases **does not** $Granger-cause$ FTSE100 (Open)	($C_{lag=1}$)	12.7429***	0.00	$\leftarrow$
Nb cases **does not** $Granger-cause$ FTSE100 (Close)	($C_{lag=1}$)	12.8323***	0.00	$\leftarrow$
Nb cases **does not** $Granger-cause$ FTSE100 (Open)	($C_{lag=10}$)	21.5957***	0.00	$\leftarrow$
Nb cases **does not** $Granger-cause$ FTSE100 (Close)	($C_{lag=10}$)	27.4723***	0.00	$\rightleftarrows$
Nb deaths **does not** $Granger-cause$ FTSE100 (Open)	($D_{lag=1}$)	3.2732*	0.07	$\rightleftarrows$
Nb deaths **does not** $Granger-cause$ FTSE100 (Close)	($D_{lag=1}$)	2.7645*	0.09	$\rightleftarrows$
Nb deaths **does not** $Granger-cause$ FTSE100 (Open)	($D_{lag=10}$)	17.6711*	0.06	$\rightleftarrows$
Nb deaths **does not** $Granger-cause$ FTSE100 (Close)	($D_{lag=10}$)	30.8576***	0.00	$\rightleftarrows$

* $p-value < 0.1$,** $p-value \leqslant 0.05$, *** $p-value < 0.01$

2.8 Conclusions, limitations, and future work

This research seeks to understand investor sentiment using a unique approach. We analyse the sentiment found in the UK government's daily COVID-19 updates, using it as a representation of UK investor sentiment. The idea is that these government communications, especially during significant events like a pandemic, can provide insights into the broader mood and perspectives of the population, including investors. Our main objective is to explore how this sentiment relates to fluctuations in the UK stock market.
The analysis is grounded in data spanning from March 3^{rd} 2020 to June 23^{rd} 2021, capturing the range from the first to the last COVID-19-related governmental press briefings.
Our findings depict a multifaceted relationship between the sentiment derived from the updates and stock prices. While certain stock indices displayed clear correlations in specific instances, these links were absent in others. To capture the changing influence of the pandemic, we segmented its progression into three distinct phases, labelled as Periods 1, 2, and 3, as depicted in Figure 2.9. The evolution of governmental responses across the three periods sheds light on the adaptive nature of crisis management.

2.8.1 Period 1: Pre-testing

- **Testing and Cases Dynamics:** This period was marked by a disparity between the reported COVID-19-related cases and the number of new fatalities, largely due to the absence of widespread testing, which potentially led to the underreporting of cases Watson et al. (2020).

- **Sentiment Dynamics:** The government consistently projected optimism despite the rise in fatalities and cases. This strategy of bolstering positivity during crises aligns with the findings of Coombs (2007) on crisis communication strategies.

- **Market Reactions:** Rising fatalities seemed to have a correlation with the sentiment of the subsequent day's governmental updates, suggesting an urgency to uplift the public's and investors' mood during turbulent times, as observed by Tetlock (2007) in his examination of media sentiment and its impact on markets.

2.8.2 Period 2: Peak pandemic

- **Rise in Cases:** The advent of testing facilities, including self-testing kits, allowed for a more accurate depiction of the pandemic's spread. As such, reported cases surged, reflecting a trend observed globally during this time Peeling et al. (2020).

- **Sentiment Dynamics:** The term "vaccine" gained prominence in governmental communications, echoing its public discourse significance. This observation aligns with Bish and Michie (2010), which emphasizes pivotal solutions in public health communications during health crises.

- **Market Reactions:** The interaction between market indicators and governmental sentiment, as described in Table 2.10, is consistent with Baker and Wurgler (2007) findings on the influence of investor sentiment on stock returns.

2.8.3 Period 3: Vaccine roll-out

- **Vaccination Impact:** Following the vaccine roll-out, its success was reflected in governmental communications and market reactions. This mirrors the societal and economic impacts of successful vaccination campaigns highlighted by Larson et al. (2014).

- **Sentiment Dynamics:** The sustained emphasis on vaccination in governmental communications resonates with Slovic (1987) insights into the role of public sentiment in guiding policy responses during crises.

- **Market Reactions:** The dynamics between information, sentiment, and the market, as exemplified in Table 2.28, align with Tetlock (2007) exploration of media sentiment's relationship with stock market movements.

2.8.4 Overall Insights

- **Governmental Strategy:**
  Sentiment Management in crisis: Throughout the analysed periods, governmental communications emerged as more than just an information relay; they became a strategic tool for shaping and steering public sentiment. In the "Pre-testing period" (Period 1), the escalating health crisis did not translate into overtly negative government communications. Instead, there was a discernible push towards optimism (Table 2.8). The sentiment polarity curve's positive tilt (Figure A.3) showcases the government's proactive approach to crisis communications, underlining the vital role of sentiment management in influencing both public perception and market sentiments.

- **Prominence of the Vaccine Narrative:**
  Symbol of Hope: As the pandemic's chapters unfolded, the term "vaccine" emerged as more than a medical solution—it became a symbol of hope. During the pandemic's peak and the subsequent vaccine roll-out (Periods 2 and 3), its mentions echoed its lifesaving potential and its role as an emblem of resilience and recovery. The heightened emphasis on this narrative had palpable ripple effects, notably on investor sentiments (Table 2.10 and Table A.10). Such strategically positioned narratives underscore their power in guiding public behaviour and setting expectations during unprecedented times.

- **Differential Impact on Stock Indices:**
  Sector Vulnerabilities: According to our research, using data from Bloomberg, a striking 74% of FTSE 250 index firms are from sectors that bore the brunt of the COVID-19 pandemic, compared to just 26% for the FTSE 100. The sectors hardest hit encompassed Tourism and Leisure (including air travel), fossil fuel production and distribution, insurance, and non-essential merchants, aligning with findings by (Griffith et al., 2020). Conversely, industries like utilities, high-tech manufacturing, tobacco, food, and pharmaceutical merchants weathered the storm better, even outpacing the market. Notably, medical and biotech research firms, which include giants like AstraZeneca, surged by 6%, countering the overall market decline of 21%. It's noteworthy that 75% of all FTSE-listed medical and biotech entities are housed within the FTSE 100, leaving the FTSE 250 with just 25%. This disparity in sectoral distribution and vulnerability underpins the differential behaviours of the two indices during the pandemic, offering valuable context to our results.

- **Granger Causality Insights:**
  Narrative Dynamics and Market Reactions: The Granger causality tests unveiled intricate predictive patterns, especially around the term "vaccine". This narrative wasn't just a buzzword; it actively influenced market trajectories (Table A.9 and Table 2.28). Such findings resonate with established financial theories like the efficient market hypothesis Fama (1970), underpinning the critical role of timely information and sentiment modulation in shaping stock market movements. In a world increasingly governed by narratives, understanding these temporal causality patterns becomes paramount for both policymakers and market players.

2.8.5 Limitations

- **Indirect Sentiment Analysis:**
  Utilizing governmental communications as a sentiment proxy presents inherent limitations. These communications are curated for broad audiences and might miss the multifaceted, rapidly evolving views of investors. While they provide transparency and manage broader perceptions, they might lag behind real-time market sentiments. Direct sentiment measurements from platforms like StockTwits and Twitter have emerged as invaluable tools (Antweiler and Frank (2004); Bollen et al. (2011)), offering almost instantaneous snapshots of investor sentiments. As Tetlock (2007) illustrated, even the mood of financial news can impact stock market movements. Moreover, the rise of algorithmic trading has underscored the importance of real-time sentiment data. These algorithms, reacting faster than humans, often leverage sentiments from multiple sources for trading decisions. Zhang et al. (2011) further supported this by highlighting the predictive power of social media sentiment on stock market behaviours. While governmental communications offer a broad

perspective on investor sentiment, other platforms, such as investor forums and social media message boards, can provide a more detailed view of specific stocks or sectors. These platforms capture nuanced sentiments in ways that governmental channels might overlook. Nguyen et al. (2015) underscored this by proposing a model that integrates sentiment derived from specific topics discussed on message boards to predict stock price movement. Their research emphasizes the importance of understanding the specific topics of discussion and the associated sentiments, rather than just the overall mood, to better inform stock market predictions.

- **Macro vs. Micro Insights:**
  While our study paints a picture of overarching trends in governmental communications and their impact on stock market indices, it might inadvertently gloss over the subtle intricacies that can be pivotal in understanding investor sentiment. These nuances, often found in discrete events or specific stock behaviours, can be pivotal in shaping the larger trends we observe. Kolari and Pynnönen (2010) underscored the importance of event-driven analyses in financial studies, suggesting that zeroing in on these specific occurrences can reveal granular insights that macro trends might obscure. Further, supporting this,Tetlock (2007) demonstrated how even short-lived media content could significantly impact stock price movements. These micro-analyses can unearth anomalies, identify previously unnoticed correlations, and offer actionable strategies for investors and policy-makers. Balancing the broad strokes of macro analysis with the detail-oriented scrutiny of micro insights ensures a more comprehensive understanding of market dynamics.

- **Broad Information Landscape:**
  The informational ecosystem that investors operate within is multifaceted and ever-evolving. While our study has placed emphasis on governmental communications, it's crucial to acknowledge that these form only one part of the broader narrative. Traditional news outlets, financial analysts, and global events contribute to shaping investor sentiment. Furthermore, the digital age has ushered in a new wave of influencers through platforms like Twitter, LinkedIn, and various financial forums. These digital platforms can amplify or even initiate market movements based on real-time discussions and sentiment shifts. Engelberg and Parsons (2011) elucidated the profound impact of media on financial markets, underlining how timely news stories can be pivotal in steering stock price trajectories. Therefore, while governmental communications are undeniably influential, the broader media landscape plays an equally, if not more, significant role in the financial decision-making process.

- **Caveats of Granger Causality:**
  While Granger causality offers a structured approach to discerning potential predictive relationships in data, its application must be cautiously undertaken. One of its core

assumptions is data stationarity, which, if violated, can lead to misleading conclusions. Moreover, our analysis encountered instances of bidirectional causality, a phenomenon where both time series appear to 'Granger-cause' each other. This suggests a form of feedback loop or circular influence between the variables, a scenario where changes in one variable can influence changes in another and vice versa over time. Such bidirectional relationships can be particularly challenging to decipher as they defy simplistic causal interpretations. In real-world scenarios, especially in the complex domain of financial markets, this might signify intricate interdependencies between variables, where events or sentiments in one domain reciprocally influence and are influenced by another. As highlighted by Ding et al. (1993), the presence of bidirectional causality often necessitates a deeper investigation into the underlying mechanisms and potential confounding factors. Furthermore, Granger causality essentially captures linear relationships, potentially overlooking intricate nonlinear dynamics in the data. Lütkepohl (2005) has underscored these challenges, emphasizing the need for more advanced models to unravel such complex relationships. In light of evolving financial ecosystems, a hybrid approach integrating Granger causality with other methods might provide a more holistic understanding.

2.8.6 Future Directions

- **Direct Sentiment Analysis as a Path Forward:**
  With the rapid advancements in Natural Language Processing (NLP) and machine learning, there lies a promising avenue in directly harnessing sentiments from financial news, investor forums, and even social media platforms. These platforms can serve as a real-time barometer of investor sentiment, capturing subtle shifts in mood and perception that might not be immediately evident in broader datasets. Tetlock (2007) underscored the potential of such an approach, demonstrating its efficacy in predicting market movements. Furthermore, the works of Sprenger et al. (2014) suggest that Twitter data, when analysed with sophisticated NLP tools, can be particularly revealing of both retail and institutional investor sentiments. Additionally, considering the rise of platforms like Reddit's r/wallstreetbets, there's a burgeoning need to incorporate such unconventional yet influential sources into sentiment models. As Bollen et al. (2011) found, the mood variations discerned from Twitter can even predict the daily ups and downs of the stock market. As we move forward, integrating these diverse data sources and advanced analytical techniques can pave the way for more robust, dynamic, and real-time financial prediction models.

- **Cross-Country Comparative Studies:**
  Investigating the interplay between governmental communications and stock market dynamics across various nations can offer invaluable insights into the global financial ecosystem. Bekaert et al. (2005) underscored the richness of insights this approach can

bring, especially in understanding how different regulatory environments, cultural nuances, and economic structures can influence stock market behaviours. Such studies can also illuminate how global events, like the COVID-19 pandemic, are perceived and addressed differently across countries and how these variances impact investor sentiment and market dynamics. For instance, CHIANG et al. (2011) demonstrated how international news can profoundly affect local market returns and volatilities. By extending our analysis to a cross-border perspective, we can also explore the synchronization of global markets, understanding how events in one country might ripple across borders, affecting markets far removed from the epicentre of the news. In the age of globalization, where markets are intricately interconnected, such an approach is valuable and imperative to unravel the complexities of the global financial tapestry.

- **Sector-Specific Exploration:**
  The COVID-19 pandemic illuminated the varying vulnerabilities across different sectors. For instance, sectors such as tourism and travel faced drastic downturns, while others like pharmaceuticals and technology saw increased interest. Given these disparities, diving deeper into sector-specific analyses becomes crucial to understanding the nuances and unique challenges each faces. In line with this, Pástor and Pietro (2003) emphasized the importance of understanding industries' characteristics when evaluating their stock returns, especially during uncertain times. Similarly, Gompers and Metrick (2001) highlighted how external shocks could lead to varied reactions across sectors, reinforcing the need for tailored analyses during global events.

- **Diversified Causal Models:**
  Traditional causality tests, while foundational, sometimes fall short of capturing the multi-faceted nature of financial data. Vector Autoregression (VAR) offers a more comprehensive perspective, allowing for the analysis of multiple time series and their interdependencies simultaneously. As Hyndman and Athanasopoulos (2018) have endorsed, such advanced methods can bring forth intricate causal relationships. Furthermore, models like Structural VAR (SVAR) provide the advantage of incorporating exogenous shocks, allowing analysts to understand unexpected events' impacts, as discussed by Blanchard and Quah (1988). There's also increasing traction in exploring non-linear causality tests Diks and Panchenko (2006), which can cater to non-linear relationships often found in financial markets. As research evolves, embracing this diverse suite of analytical tools will be essential for more exhaustive insights.

The COVID-19 pandemic has underscored the intricate web of interconnectedness that defines our globalized world. Beyond the immediate health implications, it has reverberated through economies, influenced policy decisions, and shaped public sentiment on an unprecedented scale. The intersection of health crises, economic turbulence, and rapid information dissemination has revealed the delicate balance upon which modern societies operate. This research, by delving into the nuanced interplay between governmental communications, public sentiment, and market dynamics, provides invaluable insights for a myriad of stakeholders. As highlighted by Smales (2014), sentiment, whether derived from traditional news sources or newer platforms, plays a pivotal role in financial markets; moreover, as Baker and Wurgler (2007) discuss, sentiment-driven decisions can have long-standing effects on market outcomes. Navigating this complex landscape, especially in a post-pandemic world, necessitates an understanding of these multifaceted relationships. It is our hope that this study serves not only as an analytical exploration but also as a foundation for future research, policy-making, and strategic decision-making in the face of global challenges.

2.9 Concluding Thoughts

The journey of this thesis began with an exploration into the quantitative realm of financial markets, employing the Irrational Fractional Brownian Motion (IFBM) model to delve into asset price dynamics of the S&P 500 and FTSE 100. The analysis highlighted the model's potential to provide a more nuanced understanding of market dynamics and its capacity to capture leptokurtosis, an aspect observed in real-world financial return distributions. However, the limitations inherent in the model's assumptions, parameter estimation methods, and dependency on historical data were acknowledged, paving the way for an array of future work aimed at model extension, robustness checks, and alternative parameter estimation methodologies.

Transitioning into the behavioural facet, the narrative sailed into the realm of investor sentiment in Chapter 2, with a specific lens on the COVID-19 pandemic era. The interplay between governmental communications and stock market dynamics was dissected, illuminating the nuanced influence of sentiment on market behaviours. The indirect nature of sentiment analysis through governmental communications and the caveats associated with Granger causality was recognized as limitations, urging a deeper dive into direct sentiment analysis and diversified causal models in future explorations.

The synthesis of insights from both chapters underscores the intricate dance between quantitative and behavioural dynamics in financial markets. The IFBM model's quantitative lens and the behavioural lens of investor sentiment analysis together paint a more holistic picture of market dynamics. They unveil a landscape where asset prices are swayed by both empirical statistical patterns and the capricious nature of human sentiment, oftentimes mirrored through external communications and events.

Looking ahead, the path is laid out for a myriad of exploratory directions. The exploration of alternative methodologies for parameter estimation in quantitative models, a broader robustness check across various markets and economic conditions, and the harnessing of real-time sentiment data are among the promising avenues. Cross-country comparative studies and sector-specific explorations could further enrich the understanding, offering a more nuanced glimpse into the global financial ecosystem.

The COVID-19 pandemic served as a catalyst for deeper examination, not only of market dynamics but also of the broader informational ecosystem within which investors operate. The delicate balance between quantitative models, investor sentiment, and real-world events was spotlighted, emphasizing the multi-faceted nature of financial markets. This thesis, through its dual exploration, strives to contribute to the evolving narrative of financial market analysis, offering a foundation upon which further studies can build, explore, and innovate. The boundless potential of intertwining quantitative modelling with behavioural analysis beckons, promising richer insights into the enigmatic behaviour of financial markets and the myriad factors that steer their course.

Appendix A

Appendix

A.1 UK stocks returns distributions throughout the COVID-19 pandemic

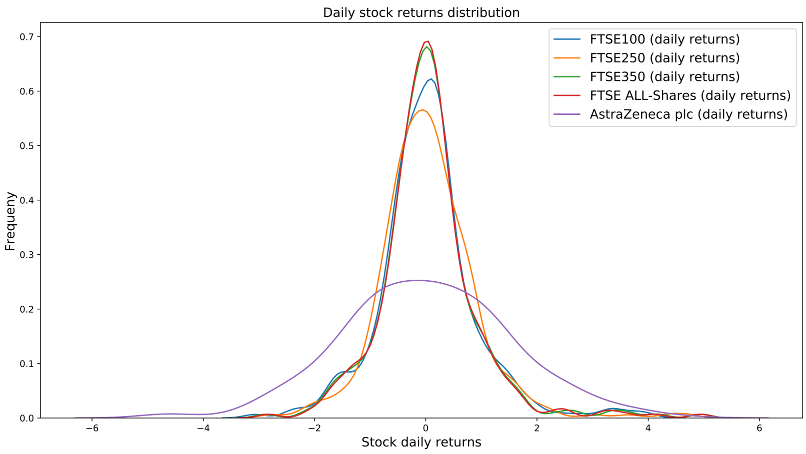


Figure A.1: UK stocks returns distributions throughout the COVID-19 pandemic

A.2 Sentiment Polarity scores— Methods comparison (BERT vs TextBlob vs Vader)

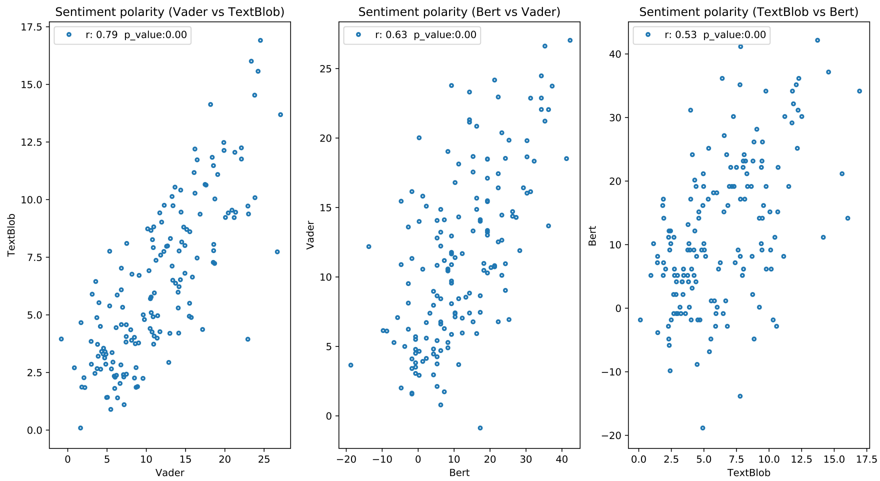


Figure A.2: Sentiment Polarity (BERT vs TextBlob vs Vader)

A.3 Bert-based Sentiment polarity scores distribution

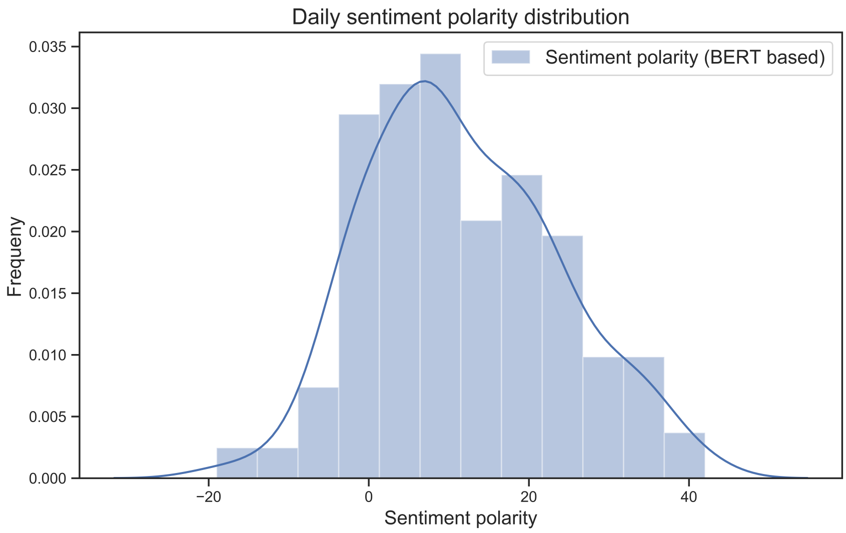


Figure A.3: Bert-based Sentiment polarity scores distibution

A.4 COVID-19 daily new deaths versus new cases lagged

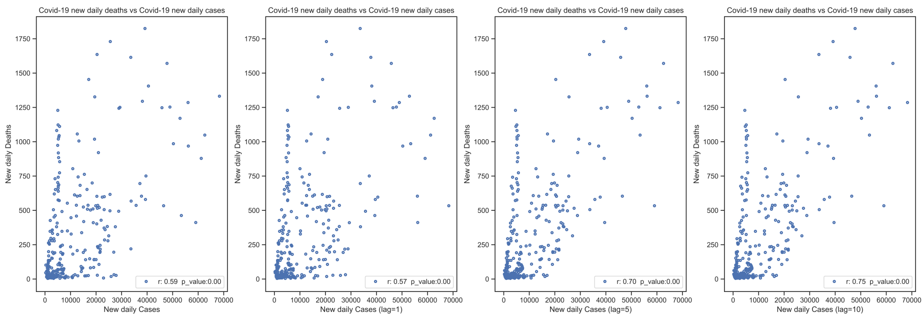


Figure A.4: COVID-19 daily new deaths versus new cases. This figure shows how the number of new deaths is correlated with the number of new cases when cases were announced 1 days ago (lag=1), 5 days ago (lag=5), and 10 days ago (lag=10)

A.5 Word frequency ”Vaccine” vs COVID-19 cases and deaths - Period 2

Table A.1: Word frequency “Vaccine” vs COVID-19 cases and deaths - Period 2

Correlations	Period	Word frequency (keyword=Vaccine)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
Number of cases (lag=1)	Period 1	0.01	-0.02	-0.03
Number of deaths (lag=1)	Period 1	0.05	-0.01	-0.02
Number of cases (lag=1)	Period 2	0.67***	0.71***	0.55***
Number of deaths (lag=1)	Period 2	0.65***	0.72***	0.52***
Number of cases (lag=1)	Period 3	0.23	0.08	0.14
Number of deaths (lag=1)	Period 3	0.31	0.09	0.12

* $p - value < 0.1$,** $p - value \leqslant 0.05$, *** $p - value < 0.01$

A.6 Correlation of sentiment polarity scores (VADER-based) vs AstraZeneca plc overnight stock returns

Table A.2: Correlation of sentiment polarity (VADER) vs AstraZeneca plc overnight stock returns

Correlations	Period	Sentiment Polarity (VADER)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
AstraZeneca plc (Next day **Open** price)	Period 1	0.21**	0.20**	0.16**
AstraZeneca plc (Next day **Close** price)	Period 1	0.23**	0.19**	0.14**
AstraZeneca plc (Next day **Open** price)	Period 2	0.35**	0.40**	0.29***
AstraZeneca plc (Next day **Close** price)	Period 2	0.32**	0.36**	0.26**

$^{*}\ p-value < 0.1,^{**}\ p-value \leqslant 0.05,\ ^{***}\ p-value < 0.01$

A.7 Correlations of Covid-19 cases vs FTSE250 prices - All periods

Table A.3: Correlations of Covid-19 cases vs FTSE250 prices - All periods

Correlations	Period	Covid-19 New Daily Cases (lag=1)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
FTSE250 (**Open** price)	Period 1	-0.30***	-0.39***	-0.28***
FTSE250 (**Close** price)	Period 1	-0.32***	-0.40***	-0.29***
FTSE250 (**Open** price)	Period 2	0.82***	0.76***	0.56***
FTSE250 (**Close** price)	Period 2	0.82***	0.77***	0.58***
FTSE250 (**Open** price)	Period 3	-0.32***	-0.33***	-0.27***
FTSE250 (**Close** price)	Period 3	-0.32***	-0.35***	-0.28***

$^{*}\ p-value < 0.1,^{**}\ p-value \leqslant 0.05,\ ^{***}\ p-value < 0.01$

A.8 Correlations of Covid-19 deaths vs FTSE250 prices - All periods

Table A.4: Correlations of Covid-19 deaths vs FTSE250 prices - All periods

Correlations	Period	Covid-19 New Daily Deaths (lag=1)		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
FTSE250 (**Open** price)	Period 1	-0.12	-0.23*	-0.18**
FTSE250 (**Close** price)	Period 1	-0.14	-0.24**	-0.19**
FTSE250 (**Open** price)	Period 2	0.89***	0.75***	0.54***
FTSE250 (**Close** price)	Period 2	0.82***	0.76***	0.55***
FTSE250 (**Open** price)	Period 3	-0.72***	-0.85***	-0.67***
FTSE250 (**Close** price)	Period 3	-0.72***	-0.84***	-0.65***

* $p - value < 0.1$,** $p - value \leqslant 0.05$, *** $p - value < 0.01$

A.9 Correlations of Covid-19 cases vs AstraZeneca plc prices - All periods

Table A.5: Correlations of Covid-19 cases vs AstraZeneca plc prices - All periods

Correlations	Period	Covid-19 New Daily Cases		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
AstraZeneca plc (**Close** price)	Period 1	-0.22*	-0.23**	-0.12
AstraZeneca plc (**Close** price) ($C_{lag=1}$)	Period 1	-0.14	-0.15	-0.06
AstraZeneca plc (**Close** price) ($C_{lag=10}$)	Period 1	0.56***	0.48***	0.35***
AstraZeneca plc (**Close** price)	Period 2	-0.80***	-0.77***	-0.60***
AstraZeneca plc (**Close** price) ($C_{lag=1}$)	Period 2	-0.79***	-0.77***	-0.58***
AstraZeneca plc (**Close** price) ($C_{lag=10}$)	Period 2	-0.68***	-0.72***	-0.54***
AstraZeneca plc (**Close** price)	Period 3	0.34***	0.08	0.06
AstraZeneca plc (**Close** price) ($C_{lag=1}$)	Period 3	0.25**	0.03	0.03
AstraZeneca plc (**Close** price) ($C_{lag=10}$)	Period 3	-0.31***	-0.45***	-0.28***

* $p-value < 0.1$,** $p-value \leqslant 0.05$, *** $p-value < 0.01$

A.10 Correlations of Covid-19 deaths vs AstraZeneca plc prices - All periods

Table A.6: Correlations of Covid-19 deaths vs AstraZeneca plc prices - All periods

Correlations	Period	Covid-19 New Daily Deaths		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
AstraZeneca plc (**Close** price)	Period 1	-0.12	-0.08	-0.04
AstraZeneca plc (**Close** price) ($D_{lag=1}$)	Period 1	-0.03	0.00	0.11
AstraZeneca plc (**Close** price) ($D_{lag=10}$)	Period 1	0.54***	0.54***	0.40***
AstraZeneca plc (**Close** price)	Period 2	-0.69***	-0.71***	-0.51***
AstraZeneca plc (**Close** price) ($D_{lag=1}$)	Period 2	-0.69***	-0.70***	-0.50***
AstraZeneca plc (**Close** price) ($D_{lag=10}$)	Period 2	-0.66***	-0.67***	-0.46***
AstraZeneca plc (**Close** price)	Period 3	-0.38***	-0.71***	-0.48***
AstraZeneca plc (**Close** price) ($D_{lag=1}$)	Period 3	-0.39***	-0.71***	-0.49***
AstraZeneca plc (**Close** price) ($D_{lag=10}$)	Period 3	-0.47***	-0.78***	-0.55***

* $p-value < 0.1$,** $p-value \leqslant 0.05$, *** $p-value < 0.01$

A.11 Correlations of Covid-19 cases vs AstraZeneca plc daily returns - All periods

Table A.7: Correlations of Covid-19 cases vs AstraZeneca plc daily returns - All periods

Correlations	Period	Covid-19 New Daily Cases		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
AstraZeneca plc (**Daily** Rt)	Period 3	0.05	0.09	0.06
AstraZeneca plc (**Daily** Rt) ($C_{lag=1}$)	Period 3	0.24*	0.12	0.08
AstraZeneca plc (**Daily** Rt) ($C_{lag=10}$)	Period 3	-0.31**	0.15	0.10

* $p-value < 0.1$,** $p-value \leqslant 0.05$, *** $p-value < 0.01$

A.12 Correlations of Covid-19 cases vs AstraZeneca plc daily returns - All periods

Table A.8: Correlations of Covid-19 cases vs AstraZeneca plc daily returns - All periods

Correlations	Period	Covid-19 New Daily Deaths		
		Pearson (ρ_p)	Spearman (ρ_s)	Kendall(τ)
AstraZeneca plc (**Daily** Rt)	Period 3	0.21**	0.23**	0.16**
AstraZeneca plc (**Daily** Rt) ($D_{lag=1}$)	Period 3	0.21**	0.14	0.08
AstraZeneca plc (**Daily** Rt) ($D_{lag=10}$)	Period 3	0.21**	0.15	0.09

* $p-value < 0.1$,** $p-value \leqslant 0.05$, *** $p-value < 0.01$

A.13 Granger causality for Vaccine mentions vs AstraZeneca plc - Period 2

Table A.9: Granger causality for Vaccine mentions vs AstraZeneca plc - Period 2

Null Hypothesis - **Period 2**	Lag	Wald Test		
		$Chi-square$	$p-value$	Direction
"Vaccine" **does not** $Granger-cause$ AstraZeneca (open)	(lag=1)	4.6001**	0.03	$\leftarrow$
"Vaccine" **does not** $Granger-cause$ AstraZeneca (close)	(lag=1)	4.5667**	0.03	$\leftarrow$
"Vaccine" **does not** $Granger-cause$ AstraZeneca (open)	(lag=10)	27.9460***	0.00	$\rightleftarrows$
"Vaccine" **does not** $Granger-cause$ AstraZeneca (close)	(lag=10)	24.7242 ***	0.00	$\rightleftarrows$
"Vaccine" **does not** $Granger-cause$ AstraZeneca (Volume)	(lag=1)	4.7372 **	0.02	$\rightarrow$
"Vaccine" **does not** $Granger-cause$ AstraZeneca (Volume)	(lag=10)	36.7324 ***	0.00	$\rightleftarrows$

* $p-value < 0.1$,** $p-value \leqslant 0.05$, *** $p-value < 0.01$

A.14 Granger causality for Vaccine mentions vs Sentiment polarity scores - All periods

Table A.10: Granger causality for Vaccine mentions vs Sentiment polarity scores - All periods

Null Hypothesis - **All periods** - (lag=5)	Period	Wald Test		
		$Chi - square$	$p - value$	Direction
"Vaccine" **does not** $Granger - cause$ Sentiment	Period 1	11.7164**	0.03	→
"Vaccine" **does not** $Granger - cause$ Sentiment	Period 2	6.6907	0.24	–
"Vaccine" **does not** $Granger - cause$ Sentiment	Period 3	16.3603***	0.00	→

* $p - value < 0.1$,** $p - value \leqslant 0.05$, *** $p - value < 0.01$

A.15 Granger causality for Sentiment polarity vs Covid-19 death and cases - Period 1

Table A.11: Granger causality for Sentiment polarity vs Covid-19 death and cases - Period 1

Null Hypothesis - **Period 1**	Lag	Wald Test		
		$Chi - square$	$p - value$	Direction
Nb deaths **does not** $Granger - cause$ Sentiment	(lag=1)	11.7287	0.00	→
Nb deaths **does not** $Granger - cause$ Sentiment	(lag=2)	11.9569	0.00	→
Nb deaths **does not** $Granger - cause$ Sentiment	(lag=3)	13.9670	0.00	→
Nb deaths **does not** $Granger - cause$ Sentiment	(lag=4)	15.8184	0.00	→
Nb deaths **does not** $Granger - cause$ Sentiment	(lag=5)	15.7543	0.00	→
Nb deaths **does not** $Granger - cause$ Sentiment	(lag=6)	18.8931	0.00	→
Nb deaths **does not** $Granger - cause$ Sentiment	(lag=7)	23.7781	0.00	→
Nb deaths **does not** $Granger - cause$ Sentiment	(lag=8)	26.7464	0.00	→
Nb deaths **does not** $Granger - cause$ Sentiment	(lag=9)	27.3277	0.00	→
Nb deaths **does not** $Granger - cause$ Sentiment	(lag=10)	31.1145	0.00	→
Nb cases **does not** $Granger - cause$ Sentiment	(lag=1)	10.5551	0.00	→
Nb cases **does not** $Granger - cause$ Sentiment	(lag=2)	13.6411	0.00	→
Nb cases **does not** $Granger - cause$ Sentiment	(lag=3)	13.9461	0.00	→
Nb cases **does not** $Granger - cause$ Sentiment	(lag=4)	13.5265	0.00	→
Nb cases **does not** $Granger - cause$ Sentiment	(lag=5)	13.1633	0.02	→
Nb cases **does not** $Granger - cause$ Sentiment	(lag=6)	16.5727	0.01	→
Nb cases **does not** $Granger - cause$ Sentiment	(lag=7)	18.9116	0.00	→
Nb cases **does not** $Granger - cause$ Sentiment	(lag=8)	21.4870	0.00	→
Nb cases **does not** $Granger - cause$ Sentiment	(lag=9)	22.8110	0.00	→
Nb cases **does not** $Granger - cause$ Sentiment	(lag=10)	28.2457	0.00	→

Granger causality for Sentiment polarity vs Covid-19 death and cases - Period 2

Table A.12: Granger causality for Sentiment polarity vs Covid-19 death and cases - Period 2

Null Hypothesis - **Period 2**	Lag	Wald Test		
		$Chi-square$	$p-value$	Direction
Nb deaths **does not** $Granger-cause$ Sentiment	(lag=1)	5.1042	0.02	←
Nb deaths **does not** $Granger-cause$ Sentiment	(lag=2)	2.8197	0.24	–
Nb deaths **does not** $Granger-cause$ Sentiment	(lag=3)	4.5781	0.20	–
Nb deaths **does not** $Granger-cause$ Sentiment	(lag=4)	7.2432	0.12	–
Nb deaths **does not** $Granger-cause$ Sentiment	(lag=5)	3.5368	0.61	–
Nb deaths **does not** $Granger-cause$ Sentiment	(lag=6)	6.2156	0.39	←
Nb deaths **does not** $Granger-cause$ Sentiment	(lag=7)	7.4649	0.38	←
Nb deaths **does not** $Granger-cause$ Sentiment	(lag=8)	13.2562	0.10	←
Nb deaths **does not** $Granger-cause$ Sentiment	(lag=9)	29.2533	0.00	⇄
Nb deaths **does not** $Granger-cause$ Sentiment	(lag=10)	29.2515	0.00	⇄
Nb cases **does not** $Granger-cause$ Sentiment	(lag=1)	4.3255	0.03	→
Nb cases **does not** $Granger-cause$ Sentiment	(lag=2)	5.2018	0.07	→
Nb cases **does not** $Granger-cause$ Sentiment	(lag=3)	6.8947	0.07	→
Nb cases **does not** $Granger-cause$ Sentiment	(lag=4)	19.5149	0.00	→
Nb cases **does not** $Granger-cause$ Sentiment	(lag=5)	12.3200	0.03	→
Nb cases **does not** $Granger-cause$ Sentiment	(lag=6)	11.2393	0.08	→
Nb cases **does not** $Granger-cause$ Sentiment	(lag=7)	11.4795	0.11	–
Nb cases **does not** $Granger-cause$ Sentiment	(lag=8)	14.5968	0.06	→
Nb cases **does not** $Granger-cause$ Sentiment	(lag=9)	22.3988	0.00	→
Nb cases **does not** $Granger-cause$ Sentiment	(lag=10)	23.0361	0.01	→

A.16 Granger causality for Sentiment polarity vs Covid-19 death and cases - Period 3

Table A.13: Granger causality for Sentiment polarity vs Covid-19 death and cases - Period 3

Null Hypothesis - **Period 3**	*Lag*	Wald Test		
		$Chi - square$	$p - value$	Direction
Nb deaths **does not** $Granger - cause$ Sentiment	(lag=1)	0.0145	0.90	$-$
Nb deaths **does not** $Granger - cause$ Sentiment	(lag=2)	5.6072	0.06	$\leftarrow$
Nb deaths **does not** $Granger - cause$ Sentiment	(lag=3)	8.4151	0.03	$\rightleftarrows$
Nb deaths **does not** $Granger - cause$ Sentiment	(lag=4)	11.6645	0.02	$\rightleftarrows$
Nb deaths **does not** $Granger - cause$ Sentiment	(lag=5)	11.2548	0.04	$\rightleftarrows$
Nb deaths **does not** $Granger - cause$ Sentiment	(lag=6)	22.2400	0.00	$\rightleftarrows$
Nb cases **does not** $Granger - cause$ Sentiment	(lag=1)	0.1967	0.65	$-$
Nb cases **does not** $Granger - cause$ Sentiment	(lag=2)	0.0045	0.99	$-$
Nb cases **does not** $Granger - cause$ Sentiment	(lag=3)	4.1608	0.24	$-$
Nb cases **does not** $Granger - cause$ Sentiment	(lag=4)	12.3233	0.01	$\rightarrow$
Nb cases **does not** $Granger - cause$ Sentiment	(lag=5)	20.1144	0.00	$\rightarrow$
Nb cases **does not** $Granger - cause$ Sentiment	(lag=6)	20.9968	0.00	$\rightleftarrows$

Bibliography

H. Abdi.
The kendall rank correlation coefficient.
Encyclopedia of measurement and statistics, 2:508–510, 2007.

H. Akoglu.
User's guide to correlation coefficients.
Turkish journal of emergency medicine, 18(3):91–93, 2018.

A. M. Al-Awadhi, K. Alsaifi, A. Al-Awadhi, and S. Alhammadi.
Death and contagious infectious diseases: Impact of the covid-19 virus on stock market returns.
Journal of behavioral and experimental finance, 27:100326, 2020.

M. Alanyali, H. S. Moat, and T. Preis.
Quantifying the relationship between financial news and the stock market.
Scientific reports, 3(1):1–6, 2013.

C. T. Albulescu.
Covid-19 and the united states financial markets' volatility.
Finance research letters, 38:101699, 2021.

L. Alfaro, A. Chari, A. N. Greenland, and P. K. Schott.
Aggregate and firm-level stock returns during pandemics, in real time.
Technical report, National Bureau of Economic Research, 2020.

F. Allen, R. A. Brealey, and S. C. Myers.
Principles of corporate finance.
McGraw-Hill/Irwin, 2011.

T. G. Andersen and T. Bollerslev.
Answering the skeptics: Yes, standard volatility models do provide accurate forecasts.
International economic review, pages 885–905, 1998.

W. Antweiler and M. Z. Frank.
Is all that talk just noise? the information content of internet stock message boards.
The Journal of finance, 59(3):1259–1294, 2004.

A. Atkeson.
What will be the economic impact of covid-19 in the us? rough estimates of disease scenarios.
Technical report, National Bureau of Economic Research, 2020.

A. Atkins, M. Niranjan, and E. Gerding.
Financial news predicts stock market volatility better than close price.
The Journal of Finance and Data Science, 4(2):120–137, 2018.

A. Azimli.
The impact of covid-19 on the degree of dependence and structure of risk-return relationship: A quantile regression approach.
Finance Research Letters, 36:101648, 2020.

D. Bahdanau, K. Cho, and Y. Bengio.
Neural machine translation by jointly learning to align and translate.
arXiv preprint arXiv:1409.0473, 2014.

M. Baker and J. Wurgler.
Investor sentiment and the cross-section of stock returns.
The journal of Finance, 61(4):1645–1680, 2006.

M. Baker and J. Wurgler.
Investor sentiment in the stock market.
Journal of economic perspectives, 21(2):129–151, 2007.

S. R. Baker, N. Bloom, and S. J. Davis.
Measuring economic policy uncertainty.
The quarterly journal of economics, 131(4):1593–1636, 2016.

S. R. Baker, N. Bloom, S. J. Davis, K. Kost, M. Sammon, and T. Viratyosin.
The unprecedented stock market reaction to covid-19.
The review of asset pricing studies, 10(4):742–758, 2020a.

S. R. Baker, N. Bloom, S. J. Davis, and S. J. Terry.
Covid-induced economic uncertainty.
Technical report, National Bureau of Economic Research, 2020b.

G. Bakshi and L. Wu.
Investor irrationality and the nasdaq bubble.
Unpublished working paper. University of Maryland and Baruch College, 2006.

R. Baldwin and B. W. Di Mauro.
Economics in the time of covid-19: A new ebook.
VOX CEPR Policy Portal, 2(3), 2020.

A. V. Banerjee.
A simple model of herd behavior.
The quarterly journal of economics, 107(3):797–817, 1992.

B. M. Barber and T. Odean.
All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors.
The review of financial studies, 21(2):785–818, 2008.

N. Barberis.
Richard thaler and the rise of behavioral economics.
The Scandinavian Journal of Economics, 120(3):661–684, 2018.

N. Barberis, A. Shleifer, and R. Vishny.
A model of investor sentiment.
Journal of financial economics, 49(3):307–343, 1998.

R. J. Barro, J. F. Ursúa, and J. Weng.
The coronavirus and the great influenza pandemic: Lessons from the "spanish flu" for the coronavirus's potential effects on mortality and economic activity.
Technical report, National Bureau of Economic Research, 2020.

BBC.
Uk government plans to hold daily white house-style televised press briefings.
BBC.
URL `https://www.bbc.co.uk/news/uk-politics-53275395`.

R. M. Beetsma, F. De Jong, M. Giuliodori, and D. Widijanto.
Spread the news: How the crisis affected the impact of news on the european sovereign bond markets.
2012.

G. Bekaert, C. R. Harvey, and C. Lundblad.
Does financial liberalization spur growth?
Journal of Financial economics, 77(1):3–55, 2005.

S. Bikhchandani, D. Hirshleifer, and I. Welch.
A theory of fads, fashion, custom, and cultural change as informational cascades.
Journal of political Economy, 100(5):992–1026, 1992.

A. E. Biondo, A. Pluchino, and A. Rapisarda.
Modeling financial markets by self-organized criticality.
Physical Review E, 92(4):042814, 2015.

K. Birmingham, R. Skelding, M. Kent, G. Viña, J. Lund, M. Meixell, T. K. Larsen, J. Afolabi, C. Marks, H. Wheeler, et al.
Astrazeneca plc (azn).
Update, 7:00, 2018.

A. Bish and S. Michie.
Demographic and attitudinal determinants of protective behaviours during a pandemic: A review.
British journal of health psychology, 15(4):797–824, 2010.

F. Black.
Studies of stock market volatility changes.
Proceedings of the American Statistical Association, Business & Economic Statistics Section, 1976, 1976.

F. Black and M. Scholes.
The pricing of options and corporate liabilities.
Journal of political economy, 81(3):637–654, 1973.

O. J. Blanchard and D. Quah.
The dynamic effects of aggregate demand and supply disturbances, 1988.

J. Bollen, H. Mao, and X. Zeng.
Twitter mood predicts the stock market.
Journal of computational science, 2(1):1–8, 2011.

T. Bollerslev.
Generalized autoregressive conditional heteroskedasticity.
Journal of econometrics, 31(3):307–327, 1986.

P. Bordalo, N. Gennaioli, and A. Shleifer.
Diagnostic expectations and credit cycles.
The Journal of Finance, 73(1):199–227, 2018.

L. Calvet and A. Fisher.
Multifractality in Asset Returns: Theory and Evidence.
The Review of Economics and Statistics, 84(3):381–406, 08 2002.
ISSN 0034-6535.
doi: 10.1162/003465302320259420.

URL https://doi.org/10.1162/003465302320259420.

E. Cambria, D. Das, S. Bandyopadhyay, and A. Feraco.
Affective computing and sentiment analysis.
In *A practical guide to sentiment analysis*, pages 1–10. Springer, 2017.

J. Y. Campbell, A. Lo, and C. MacKinlay.
The econometrics of financial markets princeton university press princeton.
New Jersey: MacKinlay, 1997.

M. A. Carlson.
A brief history of the 1987 stock market crash with a discussion of the federal reserve response.
2007.

J. C. Caruso and N. Cliff.
Empirical size, coverage, and power of confidence intervals for spearman's rho.
Educational and psychological Measurement, 57(4):637–654, 1997.

R. Casarin and F. Squazzoni.
Being on the field when the game is still under way. the financial press and stock markets in times of crisis.
PLoS One, 8(7):e67721, 2013.

R. Castellano, R. Cerqueti, and G. Rotundo.
Exploring the financial risk of a temperature index: A fractional integrated approach.
Annals of Operations Research, 284(1):225–242, 2020.

C.-O. Cepoi.
Asymmetric dependence between stock market returns and news during covid-19 financial turmoil.
Finance Research Letters, 36:101658, 2020.

R. Cerqueti and V. Ficcadenti.
Anxiety for the pandemic and trust in financial markets.
arXiv preprint arXiv:2008.01649, 2020.

W. S. Chan.
Stock price reaction to news and no-news: drift and reversal after headlines.
Journal of Financial Economics, 70(2):223–260, 2003.

Y. Chen, A. Rajabifard, S. Sabri, K. E. Potts, F. Laylavi, Y. Xie, and Y. Zhang.
A discussion of irrational stockpiling behaviour during crisis.
Journal of Safety Science and Resilience, 1(1):57–58, 2020.

T. C. CHIANG, B. N. JEON, and H. LI.
Dynamic correlation analysis of financial contagion.
Financial contagion, page 103, 2011.

A. Chouliaras.
The effect of information on financial markets: A survey.
Available at SSRN 2780921, 2016.

M. Cinelli, W. Quattrociocchi, A. Galeazzi, C. M. Valensise, E. Brugnoli, A. L. Schmidt, P. Zola, F. Zollo, and A. Scala.
The covid-19 social media infodemic.
Scientific reports, 10(1):1–10, 2020.

M. Cinelli, V. Ficcadenti, and J. Riccioni.
The interconnectedness of the economic content in the speeches of the us presidents.
Annals of Operations Research, 299:593–615, 2021.

N. Colnerič and J. Demšar.
Emotion recognition on twitter: Comparative study and training a unison model.
IEEE transactions on affective computing, 11(3):433–446, 2018.

R. Cont.
Empirical properties of asset returns: stylized facts and statistical issues.
Quantitative finance, 1(2):223, 2001.

W. T. Coombs.
Protecting organization reputations during a crisis: The development and application of situational crisis communication theory.
Corporate reputation review, 10:163–176, 2007.

T. M. Cover.
Elements of information theory.
John Wiley & Sons, 1999.

N. Crafts and T. C. Mills.
Six centuries of british economic growth: a time-series perspective.
European Review of Economic History, 21(2):141–158, 2017.

F. Critchley and M. Jones.
Asymmetry and Gradient Asymmetry Functions: Density-Based Skewness and Kurtosis.
Scandinavian journal of statistics, 35(3):415–437, 2008.

Z. Da, X. Huang, and L. J. Jin.

Extrapolative beliefs in the cross-section: What can we learn from the crowds?
Journal of Financial Economics, 140(1):175–196, 2021.

K. Daniel, D. Hirshleifer, and A. Subrahmanyam.
Investor psychology and security market under-and overreactions.
the Journal of Finance, 53(6):1839–1885, 1998.

G. A. Darbellay and D. Wuertz.
The entropy as a tool for analysing statistical dependences in financial time series.
Physica A: Statistical Mechanics and its Applications, 287(3-4):429–439, 2000.

M. Davis and A. Etheridge.
Louis bachelier's theory of speculation.
URL: https://f-origin. hypotheses. org/wp-content/blogs. dir/1596/files/2014/12/Mark-Davis-Talk. pdf, 2006.

W. F. De Bondt and R. Thaler.
Does the stock market overreact?
The Journal of finance, 40(3):793–805, 1985.

L. T. DeCarlo.
On the meaning and use of kurtosis.
Psychological methods, 2(3):292, 1997.

J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova.
Bert: Pre-training of deep bidirectional transformers for language understanding.
arXiv preprint arXiv:1810.04805, 2018.

G. Dhesi and M. Ausloos.
Modelling and measuring the irrational behaviour of agents in financial markets: Discovering the psychological soliton.
Chaos, Solitons & Fractals, 88:119–125, 2016.

G. Dhesi, M. Emambocus, and M. Shakeel.
Semi-closed simulated stock market: an investigation of its components.
International Finance and Banking Society (2011), 2011.

G. Dhesi, M. B. Shakeel, and L. Xiao.
Modified brownian motion approach to modeling returns distribution.
Wilmott, 2016(82):74–77, 2016.

G. Dhesi, B. Shakeel, and M. Ausloos.
Modelling and forecasting the kurtosis and returns distributions of financial markets: irrational fractional brownian motion model approach.

Annals of Operations Research, pages 1–14, 2019.

T. Di Matteo, T. Aste, and M. M. Dacorogna.
Long-term memories of developed and emerging markets: Using the scaling analysis to characterize their stage of development.
Journal of Banking & Finance, 29(4):827–851, 2005.

D. A. Dickey, D. P. Hasza, and W. A. Fuller.
Testing for unit roots in seasonal time series.
Journal of the American statistical Association, 79(386):355–367, 1984.

C. Diks and V. Panchenko.
A new statistic and practical guidelines for nonparametric granger causality testing.
Journal of Economic Dynamics and Control, 30(9-10):1647–1669, 2006.

Z. Ding, C. W. Granger, and R. F. Engle.
A long memory property of stock market returns and a new model.
Journal of empirical finance, 1(1):83–106, 1993.

P. G. Drazin and R. S. Johnson.
Solitons: an introduction, volume 2.
Cambridge university press, 1989.

P. Eachempati, P. R. Srivastava, and P. K. Panigrahi.
Sentiment analysis of covid-19 pandemic on the stock market.
American Business Review, 24(1):8, 2021.

T. Economist.
Only the world wars have rivalled covid-19 for news coverage, 12 2020.
URL `https://www.economist.com/graphic-detail/2020/12/19/only-the-world-wars-have-rivalled-covid-19-for-news-coverage`.

A. Elnahas, D. Kim, and I. Kim.
Natural disaster risk and corporate leverage.
Available at SSRN 3123468, 2018.

C. Emmerson, P. Johnson, I. Mitchell, and D. Phillips.
Brexit and the uk's public finances.
Technical report, IFS Report, 2016.

J. E. Engelberg and C. A. Parsons.
The causal impact of media in financial markets.
the Journal of Finance, 66(1):67–97, 2011.

R. F. Engle.
Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation.
Econometrica: Journal of the econometric society, pages 987–1007, 1982.

R. F. Engle and V. K. Ng.
Measuring and testing the impact of news on volatility.
The journal of finance, 48(5):1749–1778, 1993.

O. Erdem.
Freedom and stock market performance during covid-19 outbreak.
Finance Research Letters, 36:101671, 2020.

E. Ernawati, N. Nugraha, I. Waspada, and M. Sari.
The investor sentiment and market reaction before and during the covid-19 pandemic on indonesia stock exchange.
In *6th Global Conference on Business, Management, and Entrepreneurship (GCBME 2021)*, pages 118–124. Atlantis Press, 2022.

L. S. Exchange.
Ftse indices, n.d.
URL `https://www.londonstockexchange.com/indices?tab=ftse-indices`.

R. Fahlenbrach, K. Rageth, and R. M. Stulz.
How valuable is financial flexibility when revenue stops? evidence from the covid-19 crisis.
The Review of Financial Studies, 34(11):5474–5521, 2021.

E. F. Fama.
Efficient capital markets: A review of theory and empirical work.
The journal of Finance, 25(2):383–417, 1970.

S. A. Farimani, M. V. Jahan, A. M. Fard, and S. R. K. Tabbakh.
Investigating the informativeness of technical indicators and news sentiment in financial market price prediction.
Knowledge-Based Systems, 247:108742, 2022.

R. Feldman and J. Sanger.
The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data.
Cambridge University Press, 2007.

N. Fernandes.
Economic effects of coronavirus outbreak (covid-19) on the world economy.
SSRN Electronic Journal, 3557504, 2020.

J. T. A. S. Ferreira, F. A. Quintana, and M. F. J. Steel.
Flexible univariate continuous distributions.
Bayesian Analysis, 4(3):497 – 521, 2009.
doi: 10.1214/09-BA418.
URL `https://doi.org/10.1214/09-BA418`.

T. Fetzer.
Subsidising the spread of covid-19: Evidence from the uk's eat-out-to-help-out scheme.
The Economic Journal, 132(643):1200–1217, 2022.

M. Fischer and I. Klein.
Kurtosis modelling by means of the J-transformation.
Allgemeines Statistisches Archiv, 88(1):35–50, 2004.

I. for Government.
Timeline of uk government coronavirus lockdowns and restrictions.
`https://www.instituteforgovernment.org.uk/charts/uk-government-coronavirus-lockdowns`,
2022.
Accessed 29-July-2022.

K. J. Forbes and R. Rigobon.
No contagion, only interdependence: measuring stock market comovements.
The journal of Finance, 57(5):2223–2261, 2002.

P. Ghasiya and K. Okamura.
Investigating covid-19 news across four nations: a topic modeling and sentiment analysis approach.
Ieee Access, 9:36645–36656, 2021.

G. Giangreco.
Case fatality rate analysis of italian covid-19 outbreak.
Journal of medical virology, 92(7):919–923, 2020.

P. Gogas and T. Papadimitriou.
Machine learning in economics and finance.
Computational Economics, 57(1):1–4, 2021.

I. Goldstein, R. S. Koijen, and H. M. Mueller.
Covid-19 and its impact on financial markets and the real economy.
The Review of Corporate Finance Studies, 9(2):221–228, 2020.

I. Goldstein, R. S. Koijen, and H. M. Mueller.
Covid-19 and its impact on financial markets and the real economy.

The Review of Financial Studies, 34(11):5135–5148, 2021.

P. A. Gompers and A. Metrick.
Institutional investors and equity prices.
The quarterly journal of Economics, 116(1):229–259, 2001.

J. W. Goodell.
Covid-19 and finance: Agendas for future research.
Finance research letters, 35:101512, 2020.

J. W. Goodell and T. L. D. Huynh.
Did congress trade ahead? considering the reaction of us industries to covid-19.
Finance Research Letters, 36:101578, 2020.

C. Gourieroux, A. Holly, and A. Monfort.
Likelihood ratio test, wald test, and kuhn-tucker test in linear models with inequality constraints on the regression parameters.
Econometrica: journal of the Econometric Society, pages 63–80, 1982.

U. Government.
Coronavirus (covid-19) in the uk.
2021.
URL `https://coronavirus.data.gov.uk/details/vaccinations`.
Accessed: 2023-02-25.

C. W. Granger.
Investigating causal relations by econometric models and cross-spectral methods.
Econometrica: journal of the Econometric Society, pages 424–438, 1969.

C. W. Granger.
Testing for causality: a personal viewpoint.
Journal of Economic Dynamics and control, 2:329–352, 1980.

G. Grefenstette.
Tokenization.
In *Syntactic Wordclass Tagging*, pages 117–133. Springer, 1999.

R. Griffith, P. Levell, and R. Stroud.
The impact of covid-19 on share prices in the uk.
Fiscal Studies, 41(2):363–369, 2020.

R. A. Groeneveld and G. Meeden.
Measuring Skewness and Kurtosis.
Journal of the Royal Statistical Society: Series D (The Statistician), 33(4):391–399, 1984.

S. J. Grossman and J. E. Stiglitz.
On the impossibility of informationally efficient markets.
The American economic review, 70(3):393–408, 1980.

B. F. P. Group.
Covid-19 timeline.
https://bfpg.co.uk/2020/04/covid-19-timeline, 2022.
Accessed 29-July-2022.

T. Guardian.
Boris johnson's covid-19 address is one of the most-watched tv programmes ever, Mar 2020.
URL https://www.theguardian.com/tv-and-radio/2020/mar/24/boris-johnsons-covid-19-address-is-one-of-most-watched-tv-programmes-ever.

J. F. Hair Jr, G. T. M. Hult, C. M. Ringle, M. Sarstedt, N. P. Danks, and S. Ray.
Partial least squares structural equation modeling (pls-sem) using r: A workbook, 2021.

J. D. Hamilton.
Oil and the macroeconomy since world war ii.
Journal of political economy, 91(2):228–248, 1983.

L. P. Hansen.
Large sample properties of generalized method of moments estimators.
Econometrica: Journal of the econometric society, pages 1029–1054, 1982.

O. Haroon and S. A. R. Rizvi.
Covid-19: Media coverage and financial markets behavior—a sectoral inquiry.
Journal of Behavioral and Experimental Finance, 27:100343, 2020.

T. A. Hassan, S. Hollander, L. Van Lent, M. Schwedeler, and A. Tahoun.
Firm-level exposure to epidemic diseases: Covid-19, sars, and h1n1.
Technical report, National Bureau of Economic Research, 2020.

Q. He, J. Liu, S. Wang, and J. Yu.
The impact of covid-19 on stock markets.
Economic and Political Studies, 8(3):275–288, 2020.

M. C. Herron.
Estimating the economic impact of political party competition in the 1992 british election.
American Journal of Political Science, pages 326–337, 2000.

D. Hirshleifer, A. Subrahmanyam, and S. Titman.
Feedback and the success of irrational investors.
Journal of financial economics, 81(2):311–338, 2006.

H. Hong and J. C. Stein.
A unified theory of underreaction, momentum trading, and overreaction in asset markets.
The Journal of finance, 54(6):2143–2184, 1999.

A. Huang, W. Wu, and T. Yu.
Textual analysis for china's financial markets: a review and discussion.
China Finance Review International, 2019.

G. Huberman and T. Regev.
Contagious speculation and a cure for cancer: A nonevent that made stock prices soar.
The Journal of Finance, 56(1):387–396, 2001.

J. Hull.
Options, futures, and other derivatives (global ed. utg.), 2012.

C. Hutto and E. Gilbert.
Vader: A parsimonious rule-based model for sentiment analysis of social media text.
In *Proceedings of the international AAAI conference on web and social media*, volume 8, pages 216–225, 2014.

R. J. Hyndman and G. Athanasopoulos.
Forecasting: principles and practice.
OTexts, 2018.

M. Jazbec, B. Pàsztor, F. Faltings, N. Antulov-Fantulin, and P. N. Kolm.
On the impact of publicly available news and information transfer to financial markets.
Royal Society Open Science, 8(7):202321, 2021.

N. Jebril.
World health organization declared a pandemic public health menace: a systematic review of the coronavirus disease 2019 "covid-19".
Available at SSRN 3566298, 2020.

M. C. Jones.
Generating distributions by transformation of scale.
Statistica Sinica, 24(2):749–771, 2014.
ISSN 10170405, 19968507.
URL `http://www.jstor.org/stable/24310875`.

P. Jorion.
Value at risk: the new benchmark for managing financial risk.
The McGraw-Hill Companies, Inc., 2007.

D. Kahneman and A. Tversky.
Prospect theory: An analysis of decision under risk.
Econometrica, 47(2):263–291, 1979.

D. Kahneman and A. Tversky.
Prospect theory: An analysis of decision under risk.
In *Handbook of the fundamentals of financial decision making: Part I*, pages 99–127. World Scientific, 2013.

M. Kalaitzake.
Brexit for finance? structural interdependence as a source of financial political power within uk-eu withdrawal negotiations.
Review of International Political Economy, 28(3):479–504, 2021.

J. Kalyani, P. Bharathi, P. Jyothi, et al.
Stock trend prediction using news sentiment analysis.
arXiv preprint arXiv:1607.01958, 2016.

N. Kansal, L. Goel, and S. Gupta.
A literature review on cross domain sentiment analysis using machine learning.
International Journal of Artificial Intelligence and Machine Learning (IJAIML), 10(2):43–56, 2020.

G. Kaplanski and H. Levy.
Sentiment and stock prices: The case of aviation disasters.
Journal of financial economics, 95(2):174–201, 2010.

J. M. Keynes.
The general theory of interest, employment and money, 1936.

A. Khanthavit.
World and national stock market reactions to covid-19.
ABAC Journal, 40(2), 2020.

J. W. Kolari and S. Pynnönen.
Event study testing with cross-sectional correlation of abnormal returns.
The Review of financial studies, 23(11):3996–4025, 2010.

H. J. Larson, C. Jarrett, E. Eckersberger, D. M. Smith, and P. Paterson.
Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007–2012.
Vaccine, 32(19):2150–2159, 2014.

S. C. Lee and E. T. Lee.
Fuzzy sets and neural networks.
1974.

C. Ley.
Flexible modelling in statistics: past, present and future.
Journal de la Société Française de Statistique, 156(1):76–96, 2015.

C. Ley and D. Paindaveine.
Multivariate skewing mechanisms: A unified perspective based on the transformation approach.
Statistics & probability letters, 80(23-24):1685–1694, 2010.

Y. Li, C. Liang, F. Ma, and J. Wang.
The role of the idemv in predicting european stock market volatility during the covid-19 pandemic.
Finance research letters, 36:101749, 2020.

H. Liu, A. Manzoor, C. Wang, L. Zhang, and Z. Manzoor.
The covid-19 outbreak and affected countries stock markets response.
International Journal of Environmental Research and Public Health, 17(8):2800, 2020.

Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov.
Roberta: A robustly optimized bert pretraining approach.
arXiv preprint arXiv:1907.11692, 2019.

A. W. Lo.
Risk management for hedge funds: Introduction and overview.
Financial Analysts Journal, 57(6):16–33, 2001.

A. W. Lo.
The adaptive markets hypothesis: Market efficiency from an evolutionary perspective.
Journal of Portfolio Management, Forthcoming, 2004.

H. Lütkepohl.
New introduction to multiple time series analysis.
Springer Science & Business Media, 2005.

Š. Lyócsa, E. Baumöhl, T. Vỳrost, and P. Molnár.
Fear of the coronavirus and the stock markets.
Finance research letters, 36:101735, 2020.

A. C. MacKinlay.
Event studies in economics and finance.
Journal of economic literature, 35(1):13–39, 1997.

A. Maligkris.
Political speeches and stock market outcomes.
In *30th Australasian Finance and Banking Conference*, 2017.

B. G. Malkiel.
The efficient market hypothesis and its critics.
Journal of economic perspectives, 17(1):59–82, 2003.

M. L. Manapat, M. A. Nowak, and D. G. Rand.
Information, irrationality, and the evolution of trust.
Journal of Economic Behavior & Organization, 90:S57–S75, 2013.

B. Mandelbrot.
e variation of certain speculative prices, e journal of business, 36 (4), 394-419.
DOI: http://dx. doi. org/10.1086/294632, 1963.

B. B. Mandelbrot and R. L. Hudson.
The (mis) behaviour of markets: a fractal view of risk, ruin and reward.
Profile books, 2010.

K. H. Manguri, R. N. Ramadhan, and P. R. M. Amin.
Twitter sentiment analysis on worldwide covid-19 outbreaks.
Kurdistan Journal of Applied Research, pages 54–65, 2020.

H. Mao, S. Counts, and J. Bollen.
Predicting financial markets: Comparing survey, news, twitter and search engine data.
arXiv preprint arXiv:1112.1051, 2011.

M. Maziarz.
A review of the granger-causality fallacy.
The journal of philosophical economics: Reflections on economic and social issues, 8(2):86–105, 2015.

M. Mazur, M. Dang, and M. Vega.
Covid-19 and the march 2020 stock market crash. evidence from s&p1500.
Finance research letters, 38:101690, 2021.

W. McKibbin and R. Fernando.
The global macroeconomic impacts of covid-19: Seven scenarios.
Asian Economic Papers, 20(2):1–30, 2021.

W. Medhat, A. Hassan, and H. Korashy.
Sentiment analysis algorithms and applications: A survey.
Ain Shams engineering journal, 5(4):1093–1113, 2014.

R. C. Merton.
Theory of rational option pricing.
The Bell Journal of economics and management science, pages 141–183, 1973.

R. C. Merton.
Option pricing when underlying stock returns are discontinuous.
Journal of financial economics, 3(1-2):125–144, 1976.

J. S. Mill.
Essays on some unsettled questions of political economy.
Longmans, Green, Reader and Dyer, 1874.

N. Mirza, B. Naqvi, B. Rahat, and S. K. A. Rizvi.
Price reaction, volatility timing and funds' performance during covid-19.
Finance Research Letters, 36:101657, 2020.

A. Mudinas, D. Zhang, and M. Levene.
Market trend prediction using sentiment analysis: lessons learned and paths forward.
arXiv preprint arXiv:1903.05440, 2019.

K. Nam and N. Seong.
Financial news-based stock movement prediction using causality analysis of influence in the korean stock market.
Decision Support Systems, 117:100–112, 2019.

B. News.
Coronavirus: Chancellor unveils £350bn lifeline for economy.

T. H. Nguyen, K. Shirai, and J. Velcin.
Sentiment analysis on social media for stock movement prediction.
Expert Systems with Applications, 42(24):9603–9611, 2015.

M. Nicola, Z. Alsafi, C. Sohrabi, A. Kerwan, A. Al-Jabir, C. Iosifidis, M. Agha, and R. Agha.
The socio-economic implications of the coronavirus pandemic (covid-19): A review.
International journal of surgery, 78:185–193, 2020.

V. Niederhoffer.
The analysis of world events and stock prices.
The Journal of Business, 44(2):193–219, 1971.

R. Nielsen, R. Fletcher, N. Newman, J. Brennen, and P. Howard.
Navigating the 'infodemic': How people in six countries access and rate news and information about coronavirus.
Reuters Institute for the Study of Journalism, 2020.

R. Nunzio Mantegna.
Lévy walks and enhanced diffusion in Milan stock exchange.
Physica A: Statistical Mechanics and its Applications, 179(2):232–242, 1991.
ISSN 0378-4371.
doi: https://doi.org/10.1016/0378-4371(91)90061-G.
URL `https://www.sciencedirect.com/science/article/pii/037843719190061G`.

E. Onali.
Covid-19 and stock market volatility.
Available at SSRN 3571453, 2020.

D. W. Otter, J. R. Medina, and J. K. Kalita.
A survey of the usages of deep learning for natural language processing.
IEEE transactions on neural networks and learning systems, 32(2):604–624, 2020.

L. Pástor and V. Pietro.
Stock valuation and learning about profitability.
The Journal of Finance, 58(5):1749–1789, 2003.

L. Pástor and P. Veronesi.
Political uncertainty and risk premia.
Journal of financial Economics, 110(3):520–545, 2013.

R. W. Peeling, C. J. Wedderburn, P. J. Garcia, D. Boeras, N. Fongwen, J. Nkengasong, A. Sall, A. Tanuri, and D. L. Heymann.
Serology testing in the covid-19 pandemic response.
The Lancet Infectious Diseases, 20(9):e245–e249, 2020.

S. M. Pincus.
Approximate entropy as a measure of system complexity.
Proceedings of the National Academy of Sciences, 88(6):2297–2301, 1991.

M. Piškorec, N. Antulov-Fantulin, P. K. Novak, I. Mozetič, M. Grčar, I. Vodenska, and T. Šmuc.
Cohesiveness in financial news and its relation to market volatility.
Scientific reports, 4(1):1–8, 2014.

M. Potters, R. Cont, and J.-P. Bouchaud.
Financial markets as adaptive systems.

Europhysics letters, 41(3):239, 1998.

R. R. Prechter Jr.
Unconscious herding behavior as the psychological basis of financial market trends and patterns.
The Journal of Psychology and Financial Markets, 2(3):120–125, 2001.

V. A. Profillidis and G. N. Botzoris.
Modeling of transport demand: Analyzing, calculating, and forecasting transport demand.
Elsevier, 2018.

S. T. Rachev, C. Menn, and F. J. Fabozzi.
Fat-tailed and skewed asset return distributions: implications for risk management, portfolio selection, and option pricing.
John Wiley & Sons, 2005.

S. T. Rachev, S. Mittnik, F. J. Fabozzi, S. M. Focardi, et al.
Financial econometrics: from basics to advanced modeling techniques.
John Wiley & Sons, 2007.

S. Ramelli and A. Wagner.
What the stock market tells us about the consequences of covid-19.
Mitigating the COVID Economic Crisis: Act Fast and Do Whatever, 63, 2020a.

S. Ramelli and A. F. Wagner.
Feverish stock price reactions to covid-19.
The Review of Corporate Finance Studies, 9(3):622–655, 2020b.

G. Ranco, D. Aleksovski, G. Caldarelli, M. Grčar, and I. Mozetič.
The effects of twitter sentiment on stock price returns.
PloS one, 10(9):e0138441, 2015.

M. S. Rizwan, G. Ahmad, and D. Ashraf.
Systemic risk: The impact of covid-19.
Finance Research Letters, 36:101682, 2020.

O. A. Rosso, H. Larrondo, M. T. Martin, A. Plastino, and M. A. Fuentes.
Distinguishing noise from chaos.
Physical review letters, 99(15):154102, 2007.

F. J. Rubio and M. F. J. Steel.
Bayesian modelling of skewness and kurtosis with Two-Piece Scale and shape distributions.
Electronic Journal of Statistics, 9(2):1884 – 1912, 2015.

doi: 10.1214/15-EJS1060.
URL `https://doi.org/10.1214/15-EJS1060`.

A. Samitas, E. Kampouris, and S. Polyzos.
Covid-19 pandemic and spillover effects in stock markets: A financial network approach.
International Review of Financial Analysis, 80:102005, 2022.

N. A. Sansa.
The impact of the covid-19 on the financial markets: Evidence from china and usa.
Electronic Research Journal of Social Sciences and Humanities, 2, 2020.

P. Schober, C. Boer, and L. A. Schwarte.
Correlation coefficients: appropriate use and interpretation.
Anesthesia & Analgesia, 126(5):1763–1768, 2018.

R. P. Schumaker and N. Maida.
Analysis of stock price movement following financial news article release.
Communications of the IIMA, 16(1):1, 2018.

Z. Şenol and F. Zeren.
Coronavirus (covid-19) and stock markets: The effects of the pandemic on the global economy.
Avrasya Sosyal ve Ekonomi Araştırmaları Dergisi, 7(4):1–16, 2020.

C. E. Shannon.
A mathematical theory of communication.
The Bell system technical journal, 27(3):379–423, 1948.

A. Sharif, C. Aloui, and L. Yarovaya.
Covid-19 pandemic, oil prices, stock market, geopolitical risk and policy uncertainty nexus in the us economy: Fresh evidence from the wavelet-based approach.
International Review of Financial Analysis, 70:101496, 2020.

W. F. Sharpe.
Capital asset prices: A theory of market equilibrium under conditions of risk.
The journal of finance, 19(3):425–442, 1964.

H. Shefrin.
Beyond greed and fear: Understanding behavioral finance and the psychology of investing.
Oxford University Press, 2002.

K. Shehzad, L. Xiaoxing, and H. Kazouz.
Covid-19's disasters are perilous than global financial crisis: A rumor or fact?
Finance research letters, 36:101669, 2020.

R. C. Shiller.
Irrational exuberance.
Philosophy and Public Policy Quarterly, 20(1):18–23, 2000.

R. J. Shiller.
From efficient markets theory to behavioral finance.
Journal of economic perspectives, 17(1):83–104, 2003.

R. J. Shiller, S. Fischer, and B. M. Friedman.
Stock prices and social dynamics.
Brookings papers on economic activity, 1984(2):457–510, 1984.

R. J. Shiller et al.
Do stock prices move too much to be justified by subsequent changes in dividends?
1981.

Y. Shin and P. Schmidt.
The kpss stationarity test as a unit root test.
Economics Letters, 38(4):387–392, 1992.

A. Shleifer and L. H. Summers.
The noise trader approach to finance.
Journal of Economic perspectives, 4(2):19–33, 1990.

P. Slovic.
Perception of risk.
Science, 236(4799):280–285, 1987.

L. A. Smales.
News sentiment and the investor fear gauge.
Finance Research Letters, 11(2):122–130, 2014.

L. A. Smales.
Better the devil you know: The influence of political incumbency on australian financial market uncertainty.
Research in International Business and Finance, 33:59–74, 2015.

E. Snowberg, J. Wolfers, and E. Zitzewitz.
Partisan impacts on the economy: evidence from prediction markets and close elections.
The Quarterly Journal of Economics, 122(2):807–829, 2007.

M. K. So, A. M. Chu, and T. W. Chan.
Impacts of the covid-19 pandemic on financial market connectedness.
Finance Research Letters, 38:101864, 2021.

D. Sornette and Y. Malevergne.
From rational bubbles to crashes.
Physica A: Statistical Mechanics and its Applications, 299(1-2):40–59, 2001.

M. G. Sousa, K. Sakiyama, L. de Souza Rodrigues, P. H. Moraes, E. R. Fernandes, and E. T. Matsubara.
Bert for stock market sentiment analysis.
In *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 1597–1601. IEEE, 2019.

T. O. Sprenger, A. Tumasjan, P. G. Sandner, and I. M. Welpe.
Tweets and trades: The information content of stock microblogs.
European Financial Management, 20(5):926–957, 2014.

J. Stankevičienė and S. Akelaitis.
Impact of public announcements on stock prices: Relation between values of stock prices and the price changes in lithuanian stock market.
Procedia-Social and Behavioral Sciences, 156:538–542, 2014.

Statista.
Sources of information about the covid-19/corona pandemic worldwide as of march 2020, 2020.
URL `https://www.statista.com/statistics/1108009/sources-of-information-about-the-covid`

Statista.
Changes in media consumption due to the coronavirus (covid-19) in the uk, 2021.
URL `https://www.statista.com/statistics/1105790/covid19-media-consumption-change-uk/`.

J. H. Stock and M. W. Watson.
New indexes of coincident and leading economic indicators.
NBER macroeconomics annual, 4:351–394, 1989.

Y. Su, Y. Yip, and R. W. Wong.
The impact of government intervention on stock returns: Evidence from hong kong.
International Review of Economics & Finance, 11(3):277–297, 2002.

P. Sudhir and V. D. Suresh.
Comparative study of various approaches, applications and classifiers for sentiment analysis.
Global Transitions Proceedings, 2(2):205–211, 2021.

I. Takes.
Mitigating the covid economic crisis: act fast and do whatever, 2020.

P. O. Takyi and I. Bentum-Ennin.
The impact of covid-19 on stock market performance in africa: A bayesian structural time series approach.
Journal of Economics and Business, 115:105968, 2021.

N. N. Taleb.
The black swan: The impact of the highly improbable, volume 2.
Random house, 2007.

P. C. Tetlock.
Giving content to investor sentiment: The role of media in the stock market.
The Journal of finance, 62(3):1139–1168, 2007.

C. Thomas.
Ftse all-share index performance 2015-2021.
UK Investor Magazine, 2021.
URL `https://ukinvestormagazine.co.uk/ftse-all-share-index-performance-2015-2021/`.
Accessed: 2023-02-25.

M. Tikkanen.
Predicting the ftse all-share index daily close-to-close price direction using sentiment analysis on tweets from uk.
Master's thesis, Lappeenranta-Lahti University of Technology LUT, 2021.

M. Topcu and O. S. Gulal.
The impact of covid-19 on emerging stock markets.
Finance Research Letters, 36:101691, 2020.

R. S. Tsay.
Analysis of financial time series.
John wiley & sons, 2005.

A. Tversky and D. Kahneman.
Advances in prospect theory: Cumulative representation of uncertainty.
Journal of Risk and uncertainty, 5:297–323, 1992.

T. H. Van Hoang and Q. R. Syed.
Investor sentiment and volatility prediction of currencies and commodities during the covid-19 pandemic.
Asian Economics Letters, 1(4), 2021.

A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin.

Attention is all you need.
Advances in neural information processing systems, 30, 2017.

G. Vogel and K. Kupferschmidt.
Side effect worry grows for astrazeneca vaccine, 2021.

E.-J. Wagenmakers.
A practical solution to the pervasive problems ofp values.
Psychonomic bulletin & review, 14(5):779–804, 2007.

H. Wang, C. Li, B. Gu, and W. Min.
Does ai-based credit scoring improve financial inclusion? evidence from online payday lending.
In *40th International Conference on Information Systems, ICIS 2019*. Association for Information Systems, 1984.

J. Watson, P. F. Whiting, and J. E. Brush.
Interpreting a covid-19 test result.
Bmj, 369, 2020.

H. White.
Economic prediction using neural networks: The case of ibm daily stock returns.
In *ICNN*, volume 2, pages 451–458, 1988.

W. WHO.
Covid-19 strategy update.
Geneve, Schweiz, 2020.

Wikipedia contributors.
Gov.uk — Wikipedia, the free encyclopedia, 2022.
URL `https://en.wikipedia.org/w/index.php?title=Gov.uk&oldid=1084162235`.
[Online; accessed 22-July-2022].

J. Williams and D. Wright.
Ambiguity, responsibility and political action in the uk daily covid-19 briefings.
Critical Discourse Studies, pages 1–16, 2022.

T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz, et al.
Huggingface's transformers: State-of-the-art natural language processing.
arXiv preprint arXiv:1910.03771, 2019.

S. S. Wulff.
Time series analysis: Forecasting and control.
Journal of Quality Technology, 49(4):418, 2017.

J. Ye, D. Li, and Y. Cao.
Investor irrational selection bias in stock market based on cognitive psychology: Evidence from herding behaviour.
Revista Argentina de Clínica Psicológica, 29(1):90, 2020.

A. Zaenen and A. van den Bosch.
Proceedings of the 45th annual meeting of the association of computational linguistics.
In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, 2007.

A. Zaremba, R. Kizys, D. Y. Aharon, and E. Demir.
Infected markets: Novel coronavirus, government interventions, and stock return volatility around the globe.
Finance Research Letters, 35:101597, 2020.

F. Zeren and A. Hizarci.
The impact of covid-19 coronavirus on stock markets: evidence from selected countries.
Muhasebe ve Finans İncelemeleri Dergisi, 3(1):78–84, 2020.

D. Zhang, M. Hu, and Q. Ji.
Financial markets under the global pandemic of covid-19.
Finance research letters, 36:101528, 2020.

X. Zhang, H. Fuehres, and P. A. Gloor.
Predicting stock market indicators through twitter “i hope it is not as bad as i fear”.
Procedia-Social and Behavioral Sciences, 26:55–62, 2011.

L. Zunino, M. Zanin, B. M. Tabak, D. G. Pérez, and O. A. Rosso.
Complexity-entropy causality plane: A useful approach to quantify the stock market inefficiency.
Physica A: Statistical Mechanics and its Applications, 389(9):1891–1901, 2010.