EXCHANGE TRADED FUNDS AND THEIR ASSOCIATION WITH THE VOLATILITY OF THE DAILY RETURNS OF MAJOR SOUTH AFRICAN BANKS

Author:

Barak Herman (HRMBAR002)

Supervisor:

Darron West (University of Cape Town)

Research dissertation presented for the approval of the University of Cape Town Senate in fulfilment of part of the requirements for the degree of Master of Commerce specialising in Finance (in the field of Financial Management) in approved courses and a minor dissertation. The other part of the requirement for this qualification was the completion of a programme of courses. I hereby declare that I have read and understood the regulations governing the submission of Master of Commerce dissertations, including those relating to length and plagiarism, as contained in the rules of the University, and that this dissertation conforms to those regulations.

Faculty of Commerce

UNIVERSITY OF

CAPE TOWN

February 2024

UNIVERSITY OF CAPE TOWN

FACULTY OF COMMERCE

Igniting Knowledge and Opportunity

Plagiarism Declaration

COMPULSORY DECLARATION:

1. This dissertation has been submitted to Turnitin (or equivalent similarity and originality checking software) and I confirm that my supervisor has seen my report and any concerns revealed by such have been resolved with my supervisor.
2. I certify that I have received Ethics approval (if applicable) from the Commerce Ethics Committee.
3. This work has not been previously submitted in whole, or in part, for the award of any degree in this or any other university. It is my own work. Each significant contribution to, and quotation in, this dissertation from the work, or works of other people has been attributed, and has been cited and referenced.

Student number	HRMBAR002
Student name	Barak Herman
Signature of Student	Signed by candidate
Date:	31 January 2024 | 10:00:18 PM SAST

Abstract

This dissertation examines whether the degree of ownership of major South African Bank shares by the Exchange Traded Fund (ETF) market in South Africa along with ETF's unique trading dynamics (for the purposes of this dissertation, "ETF's unique trading dynamics" refers to the relationship between ETFs and their constituent shares through the ETF share creation and redemption process and the purchase and sale of constituent shares), has an association with the volatility of the South African Major Bank's daily returns. This dissertation determines the degree of ownership by the ETF market and its trading dynamics using publicly available data through the identification and collation of publicly available data obtained from third-party data providers. Using a GARCH model process to estimate the volatility of daily returns, this dissertation then incorporates the degrees of ETF ownership and ETF trading dynamics variables as external regressors to assess whether there was a statistically significant association in explaining the volatility of daily returns of the Major South African Banks shares. The results of the dissertation are mixed. Concerning the degree of ETF ownership, there is evidence of a statistically significant negative association between the degree of ETF Ownership and the volatility of daily returns for two of the six shares, while in the instances of the remaining shares, there is no statistically significant association. With respect to ETF unique trading dynamics, driven by ETF-Specific Events (which, for the purposes of this dissertation, means the introduction of a new ETF into the market, the creation of new ETF shares, and redemption of existing ETF shares), no statistically significant association is found between the volatility of daily returns for any of the shares. This study reveals a nuanced relationship between ETF ownership and stock volatility in major South African banks, demonstrating that while higher ETF ownership may reduce volatility in some bank shares, its impact is not universal across the sector. These findings may be useful for investors, financial analysts and policymakers who are interested in understanding the potential

impact ETFs may have on equity markets, and who may use this information for risk management, advising on investment strategies, and shaping regulations around ETFs. This research adds to the body of knowledge in relation to understanding of the role of ETFs in financial market stability and their potential impact on constituent share price volatility, with specific reference to emerging markets like South Africa, offering valuable insights for various market stakeholders.

Keywords: *ETF market; GARCH model; ETF ownership; ETF trading*

Contents

1 INTRODUCTION

1.1 Background

ETFs have seen a tremendous increase in popularity since their inception. Globally, the number of ETFs and assets under management ("AUM") has grown year on year, and South Africa has seen a similar trend, although on a smaller scale. As can be seen in Figure 1, the global ETF market, and Figure 2, the South African Exchange Traded Product ("ETP") market, global AUM has reached over 9.5 trillion US Dollars at the end of 2022, with South Africa's ETP market reaching a reasonable 129 billion South African Rand at the end of 2022.

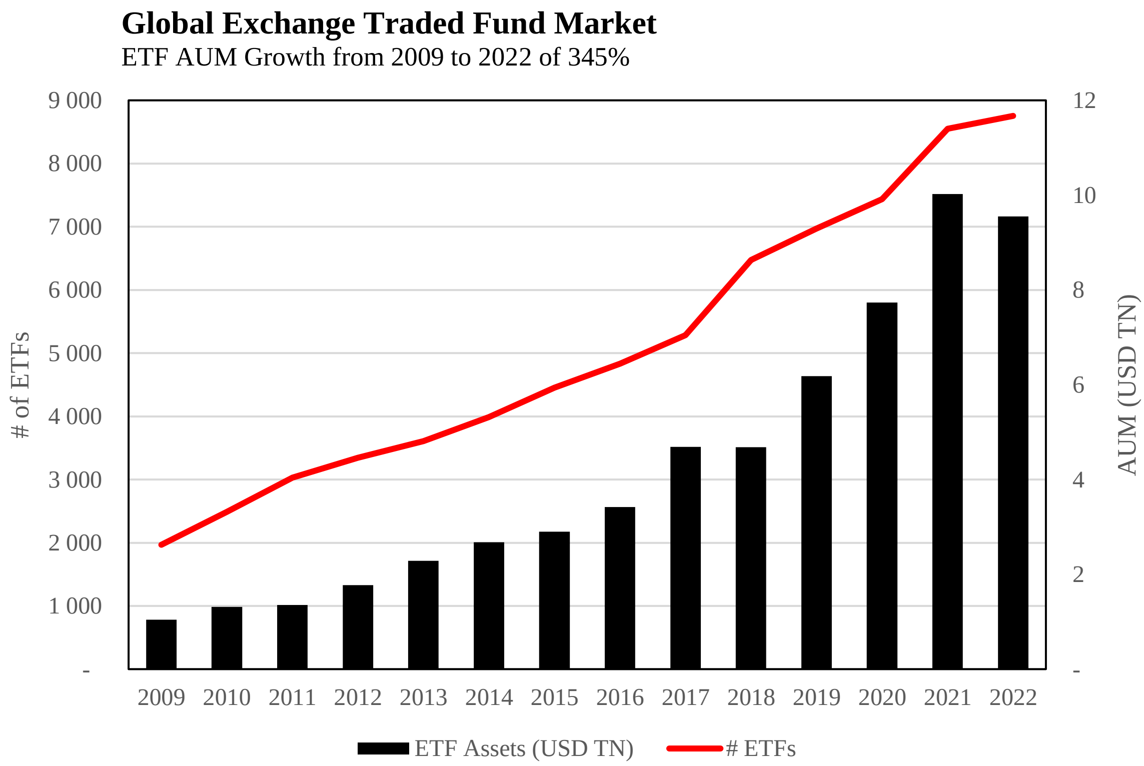


Figure 1: Global Exchange Traded Fund Market: ETF AUM Growth from 2009 to 2022 of 3345% (Source: etfgi.com)

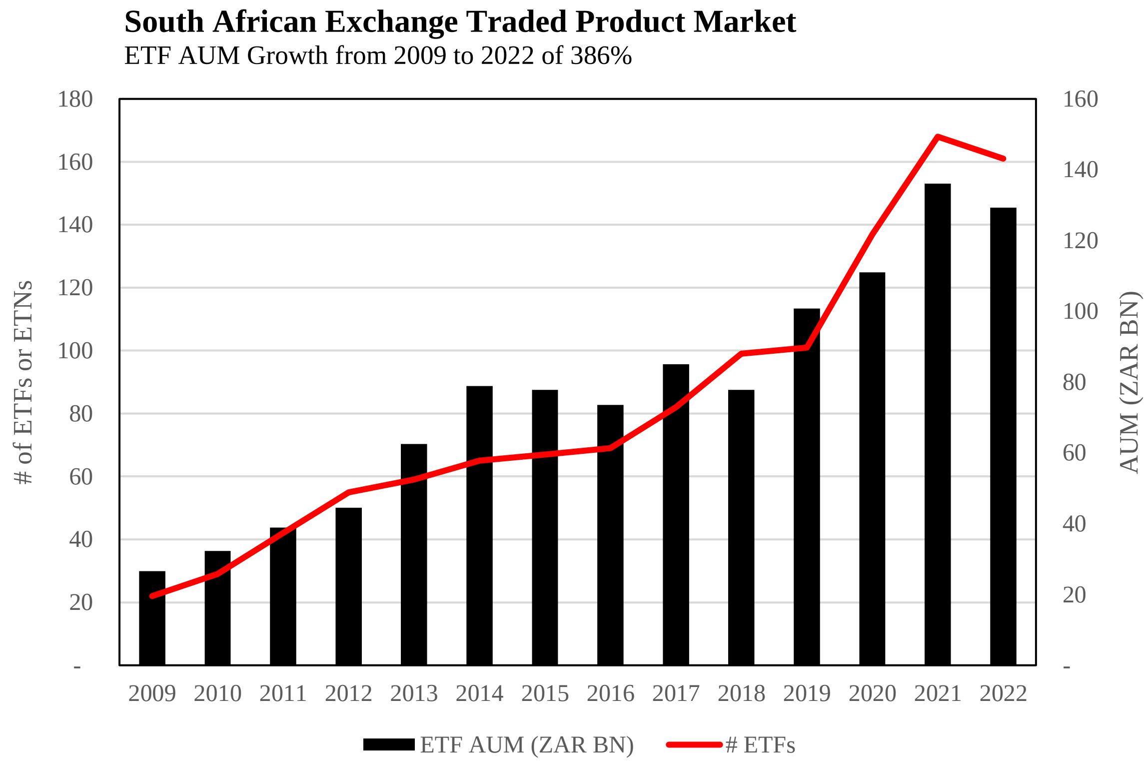


Figure 2: South African Exchange Traded Product Market: ETF AUM Growth from 2009 to 2022 of 386% (Source: etfsa.co.za)

Apart from the increases in AUM, ETFs are significant drivers of trading volume. According to iShares, the world's largest ETF provider, ETFs accounted for 37% of trading by value in the United States in 2018 (BlackRock, 2018) up from 30% in 2016 (Vlastelica, 2017). This indicates how extensively investors use ETFs to gain exposure to various sectors quickly and cheaply.

To begin with, several factors drive ETFs' popularity. Investors are drawn in by ETFs tax efficient structure (this benefit is limited from a South African perspective. ETFs are taxed in the same manner as unit trusts, but some ETFs may be used in a taxpayers' tax fee savings account), investment holdings transparency, cheap and effective diversification, intraday liquidity and low costs (there is some debate on this topic as there are additional transactional costs involved with ETFs that increase the effective annual cost of the investment). Due to these benefits, ETFs are popular with investment professionals and average retail investors. Besides, advancements in technology and online investing platforms are also making it easier for the average retail investor to invest in these products. As more unsophisticated investors utilize ETFs as investment vehicles, there must be transparency regarding the risks, returns, and impact on financial markets.

It is evident that since their creation, ETFs as investment vehicles have seen some innovation their underlying asset base and their mechanics. ETFs expose investors to the equity, fixed-income, currency, and commodity markets. With the introduction of leveraged and inverse ETFs, an investor can gain multiples of exposure to their desired market. The South African ETF market has remained relatively vanilla, the only variation being the underlying exposure. However, with financial innovation and increased popularity comes great scrutiny and attention from market commentators and regulatory bodies.

1.2 Financial innovation and concerns raised by regulators

Financial engineering innovations like program trading and ETFs have been blamed for market crashes like October 1987's Black Monday, where markets fell by as much as 20%, and the Flash Crash of 2010, where the Dow Jones Industrial Average ("DJIA") fell 9% in a few minutes only to recover most of the loss later in the day. In both instances, a breakdown of the relationship between the underlying asset and related derivative market had been noted as the driver of market movements (it should be noted that ETFs are not derivatives due to their "in-kind" nature, although leveraged and inverse ETFs may be classified as derivatives).

For these and other market anomalies, market commentators and financial regulators have raised concerns regarding ETFs and the potential impact they may have on the stability of financial markets. Various regulatory bodies have raised various concerns that centre around ETFs and their mechanics.

The Financial Stability Board's ("FSB") paper on the potentially destabilizing effects of ETFs raised several concerns. The FSB called for increased surveillance on ETF product innovation to ensure that potential investors' (particularly retail investors) understanding of the risks relating to the underlying asset class and the return profile of exotic ETFs (leveraged and inverse ETFs) is as clear as possible. The FSB also referred to the increased counterparty risks a Synthetic ETF investor faces, especially when the ETF and Total Return Swap ("TRS") counterparty is the same bank. This risk is compounded where illiquid assets are posted as collateral for the TRS. The common practice of ETFs engaging in securities lending to improve returns also raised concerns for the FSB. The rationale was that increased levels of securities lending would put additional strain on the ETF, where the fund experienced large outflows or increased default risk by the borrower (FSB, 2011). Likewise, the Bank for International Settlements ("BIS") paper detailed the history of various ETF structures. It discussed more exotic ETFs (synthetic ETFs) and compared to structured products. The paper also highlighted concerns about risks that may impact the stability of financial markets. It was manifested that the inability of investment banks to appropriately manage risks stems from synthetic ETFs and their tracking error risk while managing the risk of their trading book. This additional risk was further enhanced when outflows from these ETFs put additional strain on the TRS counterparties. Further, ETF structures may increase counterparty risk during times of crisis through problems with collateral. Securities lending practices impact the ability to effect calls for collateral, and the use of illiquid assets to cover collateral requirements has a negative impact on collateral values and increases counterparty risk, which in turn increases overall liquidity in the market (Ramaswamy, 2011).

Conversely, the European Central Bank ("ECB") mentioned apprehensions that the transmission of risks throughout financial markets was more significant due to the prevalence and increased usage of ETFs. Like the FSB, the ECB was concerned that there may be increased liquidity risk due to the "on-demand" liquidity offered by ETFs, especially market stress and low liquidity; due to the nature of synthetic ETFs and ETF's securities lending practices, the overall financial markets exposure to counterparty risk was elevated. The ECB's report also mentioned the possibility that the increased popularity and in turn holdings of ETFs may be increasing the volatility of the underlying markets that the ETFs were tracking. The report unfortunately did not go further but referred to the work of Ben-David, Franzoni, and Moussawi whose work was considered in section two of this dissertation (ECB, 2018).

Following the above themes of transferring information across financial markets, there have been several studies addressing the potential contagion effects between markets (most research has addressed derivatives and their underlying market), but only in recent years have ETFs been included in this vein of research.

1.3 Why should ETFs be researched?

It is apparent that ETFs are used by retail and professional investors, owing to their ease with and low cost at which a person can invest in a broad market or sector-specific exposure provided in liquid and illiquid markets. Also, the ability of EFT flows to provide insight into investor sentiment (thusand enable the dissemination of information to underlying markets) make ETFs a financial instrument that is of interest to a broad spectrum of market participants. In this domain, this dissertation investigates the potential impact that ETFs have on their underlying asset volatility (with specific focus on South African equities). Vast numbers of sell-side research, investment firms, and listed companies' use intrinsic valuation methodology to provide investment recommendations, make investment decisions, and make capital budgeting decisions. An integral component of intrinsic valuation is using a company's Weighted Average Cost of Capital ("WACC"), where the cost of capital is directly impacted by the company's volatility relative to the broader market (Gregoriou, 2015). Unwarranted increases in volatility, arising from non-fundamental price changes, may hinder a company's ability to, raise affordable capital (both debt and equity), be seen as a stable and attractive investment option by large institutional investors, provide appropriate share-based compensation to employees, and partake in merger and acquisitions activity. Another consideration was the impact on the company's capital structure and capital allocation decision-making. Varying the cost of capital has the potential to change the target capital structure, decision to buy back shares, and execute other corporate actions. These potential consequences will have a direct impact on the company's and shareholder value.

In summary, although ETFs were not the most material investment vehicle, they were worthy of further research due to their overall growth trajectory and increasing popularity; the type of investors they attract, both sophisticated and unsophisticated; their ability to signal market sentiment; their potentially destabilizing structure that has the ability to disrupt orderly markets; their potential to accelerate and exacerbate the transmission of risk; the material consequences of affecting the characteristics of underlying assets (liquidity, volatility); and the limited amount of research conducted.

1.4 Research Objective

The objective of this study is to determine if there is an association between the volatility of daily returns of the major South African banks and 1.) the level of ETF ownership of a constituent stock and 2.) ETF-Specific events (which, for the purposes of this dissertation means the introduction of a new ETF into the market, the creation of new ETF shares, and redemption of existing ETF shares).

1.5 Study Outline

The remainder of the dissertation is split into five sections. Section 2 encompasses a comprehensive literature review covering the existing body of work on ETFs and their impact on underlying asset volatility. Section 3 details the data gathered and the methodology applied to test the research questions. Section 4 summarizes the findings and results from the tests performed, section 5 provides an overall conclusion and section 6 potential arears for future research.

2 LITERATURE REVIEW

This chapter provided an overview of the origins of ETFs, the common types of ETFs found in global financial markets, and their unique characteristics that set them apart from other financial instruments. It then described how the underlying constituents and indices has been impacted by the introduction of ETFs and trading activity related explicitly to ETFs, including the volatility of the constituents. In the end, it then reviewed the various methodologies employed to measure and test the volatility of the underlying constituents.

2.1 Exchange traded funds

2.1.1 ETFs: a brief history

The idea of having a single basket of securities was not born with the first ETF. From the 1970s, the popularity of Index Tracking Mutual Funds, and later in the 1980s, Portfolio and Program trading continued to grow. Gaining exposure to a broad spectrum of securities through a single instrument with lower associated costs appealed to many investors. With this demand, in 1989, the first US Index Participation Share ("**IPS**"), which tracked the S&P 500 was launched. Its tenor was short-lived due to a federal court ruling and was not substituted for some time (Gastineau, 2010).

During this lull, the Toronto Stock Exchange ("**TSE**") launched the Toronto Index Participation Share ("**TIPs**"), which tracked the TSE 35 (the largest 35 companies listed on the TSE). TIPs was a successful product, and the TSE launched another similar product, which tracked the TSE 100 shortly after. In 2000, the TSE exited their Index Participation Share business and offloaded it to Barclays, which incorporated the holdings into one of their funds.

While the TSE had successfully launched TIPs, the American Stock Exchange was working on its US-based product. In 1993, the S&P 500 Depository Receipt, or SPDR, was introduced to the market. As the name suggested, SPDR tracked the S&P 500 index and was the world's largest ETF with over $300bn in AUM (Gastineau, 2010).

Since the inception of the first ETF, there have been several advances in the types of assets being tracked by ETFs. Initially, ETFs tracked equity indices, but in later years, a broader range of asset classes was added (Bojinov, 2015).

In 2002, the first Fixed Income ETFs were launched by iShares, which looked at tracking varying US Treasuries. The current fixed-income ETF market allowed individual investors to gain access to specific credit market exposure, government bonds, corporate bonds, differing currencies and geographies, and varying levels of credit risk.

Commodities was the class that came into the ETF fold with State Street's launch of the GLD ETF in 2004. GLD offered investors exposure to Physical Gold Bullion and was backed by the physical commodity. Investors use GLD to hedge against various risks and market volatility. Like State Street, other ETF issuers are considering bringing new products to the market.

In 2005, Invesco listed its Euro Currency Trust, which offered exposure to the Euro relative to the US Dollar. Over the years, there has been an increase in the number of currencies provided, but they were limited to major and liquid global currencies. Currency ETFs enable investors to gain exposure to currencies within the equity market. As a long-term investment, currency have not succeeded as the underlying deposits usually earn low interbank rates and incur relatively higher transaction costs.

2006 saw the introduction of Leveraged ETFs into financial markets with ProShares' suite of "Bullish" and "Bearish" products. Leveraged (long) and Inverse (short) ETFs provided investors with a return profile that was a multiple of the underlying index being tracked. Although Leveraged ETFs appeared straightforward, they possessed attributes that add complexity and should only be used by sophisticated investors. In section 2.1.3, we dive deeper into Leveraged ETFs and their mechanics.

Hence, it is evident now where ETFs have come from and the variations in the market, now the next sections would delve into getting a better understanding of the nuances of ETFs.

2.1.2 What is an ETF?

As asserted by Lettau and Madhavan (Lettau and Madhavan, 2018), an ETF is a marketable security that can be bought and sold throughout the trading day. An ETF is similar to an index tracking mutual fund/unit trust in that the objective of an ETF is to track the underlying performance of a basket of securities or index. The main difference between an index-tracking mutual fund/unit trust and an ETF is that an ETF can be traded throughout the day, while a mutual fund/unit trust can be bought and sold at the end of the trading day (Lettau and Madhavan, 2018).

2.1.2.1 ETF Creation and Redemption

ETF shares are created and redeemed through a unique mechanism known as Creation and Redemption units (Ben-David et al., 2018). Figure 3 outlined the dynamics of the relationship between the various market participants.

The ETF Manager will appoint an Authorised Participant (**"AP"**) as its market maker. An AP is usually a large financial services firm that has the necessary skills, resources, and infrastructure to act as the ETF's market maker (Ben-David et al., 2018).

Where ETF shares are created, investors will provide cash to the AP, who will go into the market and source the various securities that make up the basket the ETF is tracking. The AP will then provide the securities to the ETF in exchange for shares in the ETF (Ben-David et al., 2018).

Where ETF shares are redeemed, the investor will provide the AP with the ETF share. In exchange, the investor will receive cash from the AP. The AP will then present the ETF share to the ETF for redemption, and the ETF will take back the ETF share and present the AP with the basket of securities that make up the ETF (Ben-David et al., 2018). .

As stated by Lettau and Madhavan (Lettau and Madhavan, 2018), one of the unique attributes of ETFs is the “in-kind” exchange of securities for ETF shares that enables APs to take advantage of imbalances between the ETF market and the underlying securities market. This enables ETF and underlying securities markets to effectively move to equilibrium through the APs and other market participants taking advantage of the arbitrage opportunity (Lettau and Madhavan, 2018).

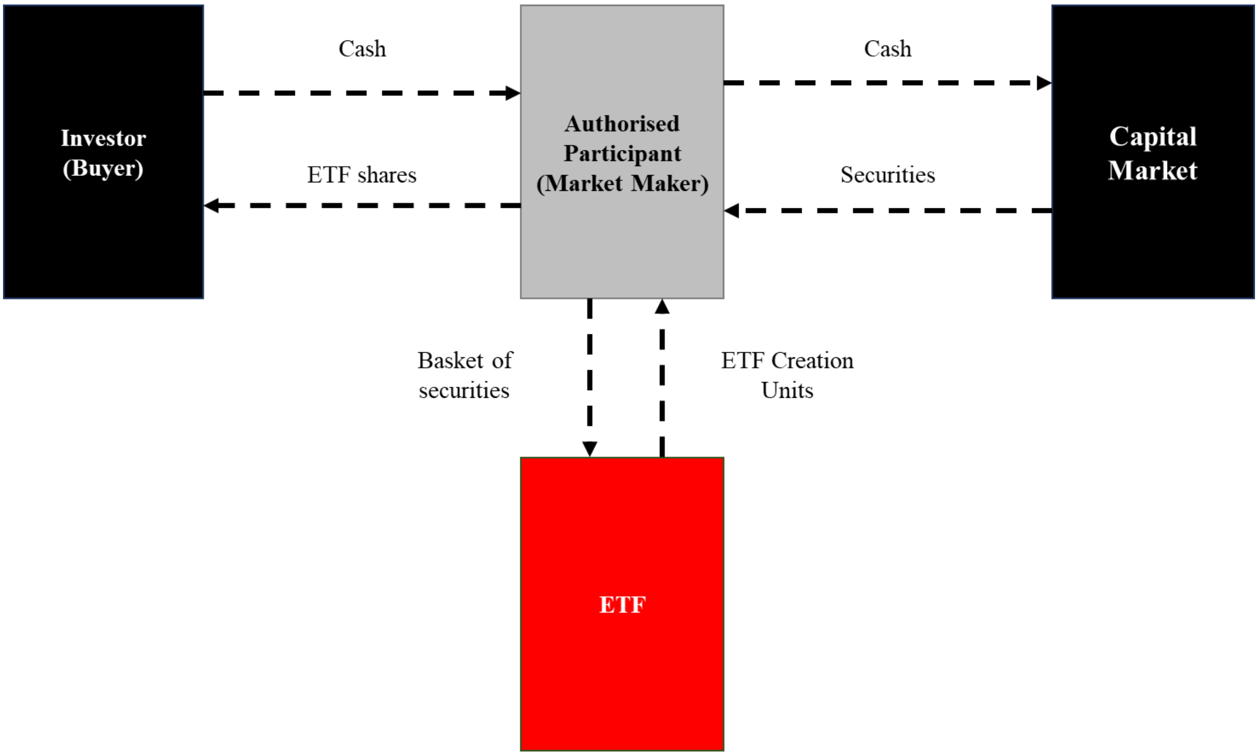


Figure 3: ETF Creation and Redemption (Source:Ben-David et al. (2018))

2.1.2.2 ETF Arbitrage and the theoretical impact on demand/supply of underlying securities

Certainly, the easiest way to understand the arbitrage opportunity between the ETF and underlying securities markets is through an example.

Figure 4 illustrated the scenario where the ETF shares, with a current market price of R90, are lower than the R100 Net Asset Value (**"NAV"**) of the basket of securities the ETF is tracking. In this case, the AP would go into the market and purchase an ETF share from a willing seller for R90. The AP would then present the ETF share to the ETF manager and in return the AP would receive the basket of securities currently worth R100. The AP would then sell the individual securities in the market for R100 (this example does not account for transaction costs and taxes), making a risk-free profit of R10.

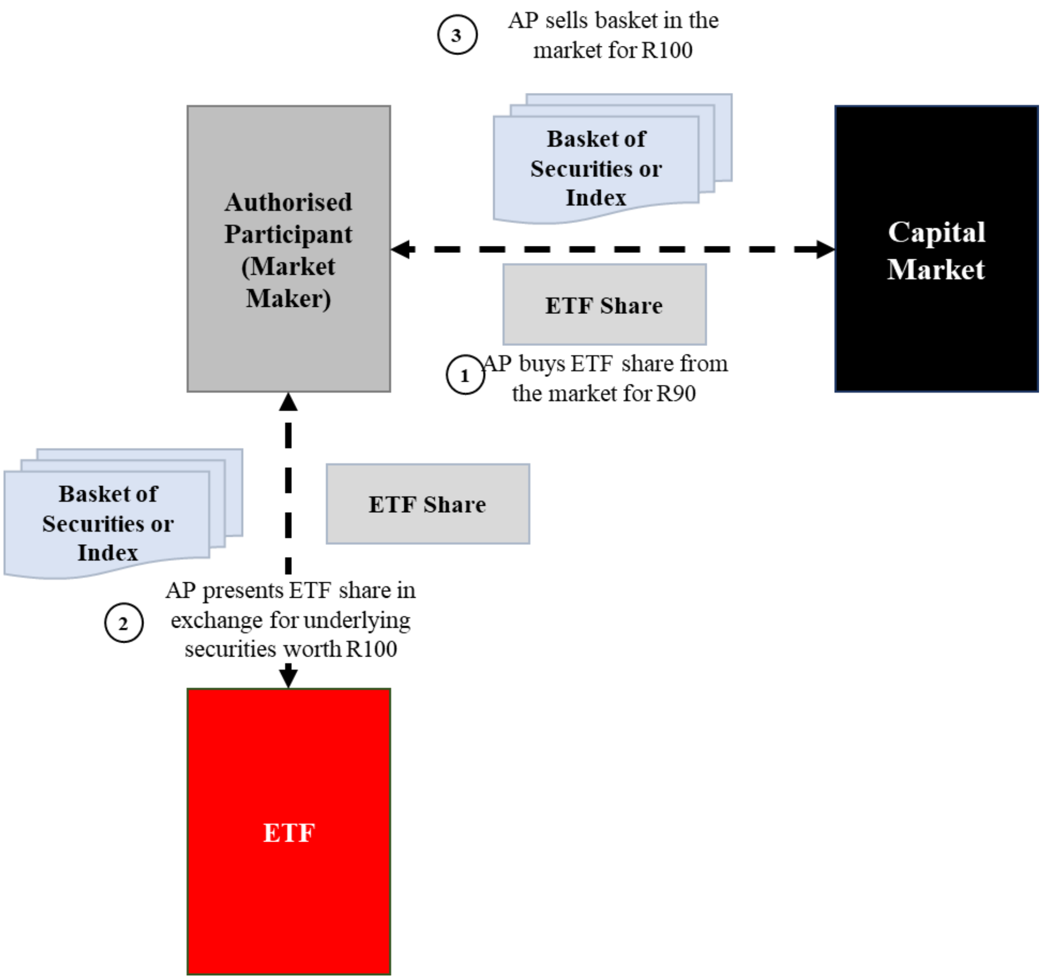


Figure 4: ETF Arbitrage where ETF price is less than ETF NAV (Source: Ben-David et al. (2018))

Figure 5 illustrated the converse scenario where the basket of securities, with a current NAV of R90, is lower than the current ETF share price of R100. The AP would go into the market and purchase the individual securities that make up the basket the ETF is tracking at a cost of

R90. The AP would then present the basket of securities to the ETF, who, in exchange, would provide the AP with the ETF share. The AP would then sell the ETF share in the market for R100 and make a risk-free profit of R10.

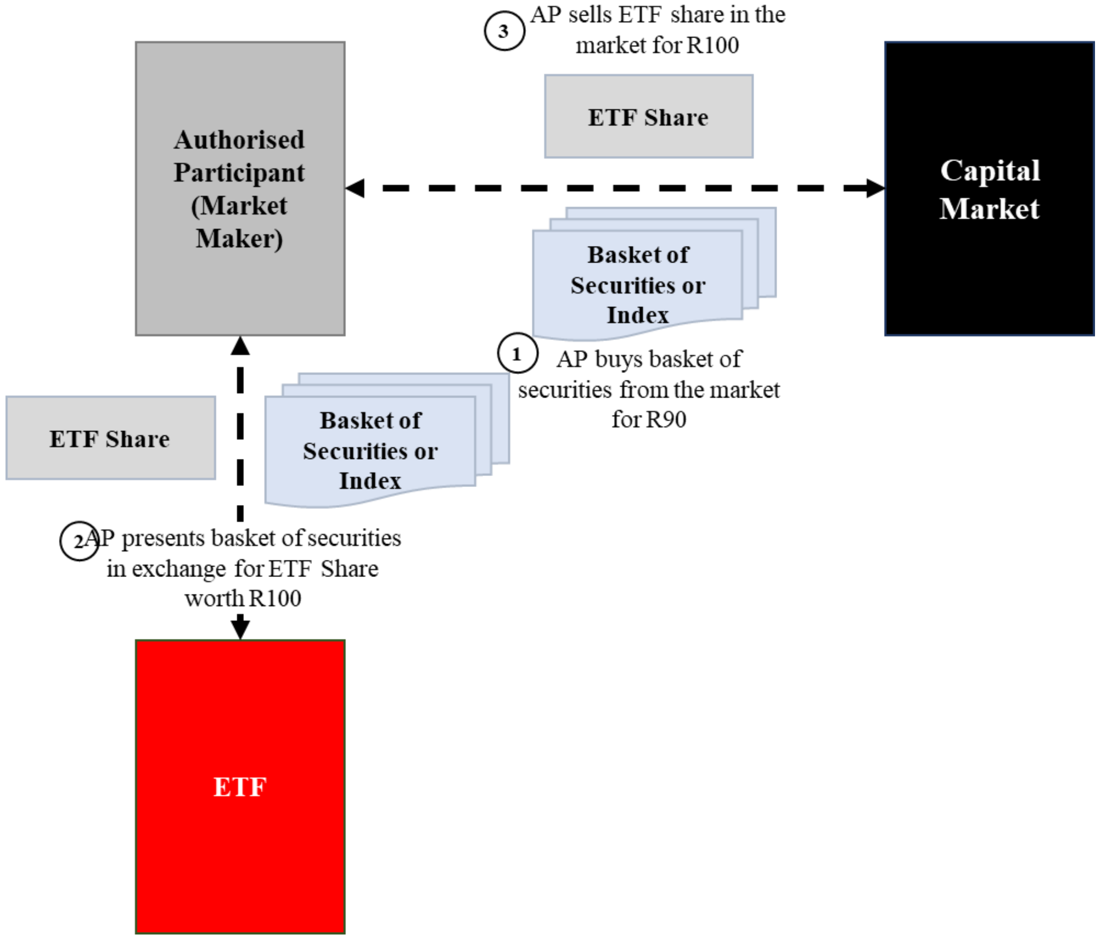


Figure 5: ETF Arbitrage where ETF price is more than ETF NAV (Source: Ben-David et al. (2018))

In these scenarios, the AP is taking advantage of its position as a market maker to capture these profits and bring the markets into equilibrium. Other market participants are able to take advantage of the same imbalances by buying the cheaper assets and short-selling the more expensive ones, and then holding the positions until the asset prices converge and close out their positions.

The arbitrage activity brings non-fundamental price changes in the ETFs and baskets of securities, which move the markets into equilibrium. Using the demand and supply pricing theory, we can understand the dynamics of the markets (Ben-David et al., 2018).

Figure 6 illustrated where the ETF market is priced lower than the underlying securities. As shown in Figure 4, the AP would purchase the ETF share, increasing its demand and driving up the price, and sell the underlying basket, increasing supply and lowering the price (Ben-David et al., 2018).

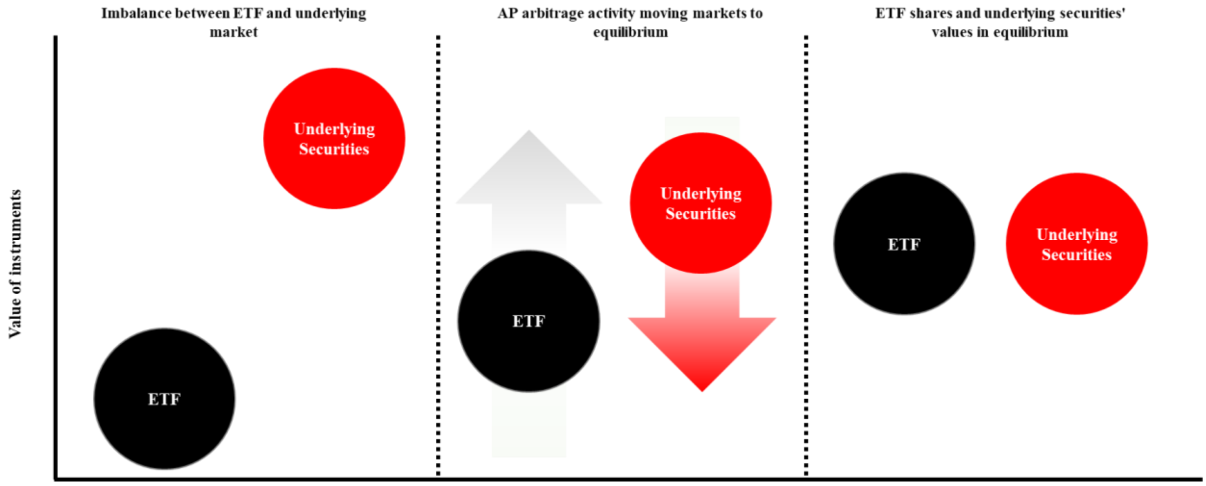


Figure 6: ETF Equilibrium where ETF price is less than NAV (Source: Ben-David et al. (2018))

Figure 7 illustrated the converse, where the ETF market is priced higher than the underlying securities. As outlined in Figure 5, the AP would purchase the underlying basket, increasing demand and driving the price of the underlying basket higher, and selling the ETF share, increasing supply, and moving the price lower (Ben-David et al., 2018).

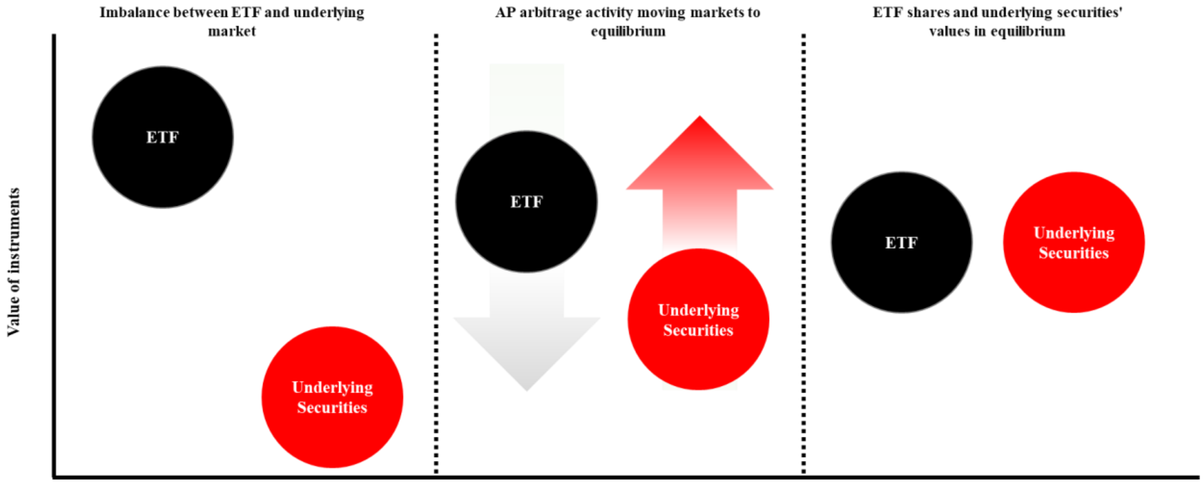


Figure 7: ETF Equilibrium where ETF price is more than NAV (Source:Ben-David et al. (2018))

2.1.3 Leveraged and Inverse ETFs

Similar to plain vanilla ETFs, Leveraged and Inverse ETFs (collectively named Leveraged ETFs for this dissertation) sought to provide investors with an amplified long or short exposure to an index at a mandated order of magnitude. A leveraged ETF amplifies the returns of an index through the use of debt or derivative instruments. To provide these amplified returns, the fund must maintain a leverage ratio consistent with its mandate at the end of each trading day. In order to do this, the fund will reset its underlying exposures at the end of each day to ensure the desired returns (Cheng and Madhavan, 2009).

2.1.3.1 Rebalancing activity of Leveraged and Inverse ETFs and opportunistic traders

As mentioned in section 2.1.3, Leveraged ETFs amplify the underlying index returns using debt or derivative instruments. The Leveraged ETF maintain a constant leverage ratio by resetting its exposure at the end of each day to ensure that the mandated ratio is maintained. Funds rebalance at different frequencies to reduce transactional costs (Yates, 2018).

The next paragraph describes an example to illustrate the mechanics of Leverage ETF rebalancing activities.

A leveraged 2x ETF, where the desired return was two times that of the underlying index, has an NAV of R100 at T_0. The 2x ETF needed to be exposed to R200's worth of the underlying index at T_0 in order for it to deliver its 2x return on T_1. At T_1, the index increased in value by 10% (R10) to an NAV of R110. In turn, the 2x ETF has increased its NAV by 20% (R20), two times the return of the index, to R120.

At the end of T_1, the 2x ETF would need to increase its exposure to the underlying index from R200 to R240, as the current NAV of the ETF is R120, to be able to deliver a 2x return the following day.

On T_2, the index increased by 13.64% (R15) to an NAV of R125. The 2x ETF increased its NAV by 27.27% (R32.73) to an NAV R152.73. At the end of T_2, the 2x ETF would increase its exposure to R305.45 to deliver its 2x return the following day.

On T_3, the index lost 20% or R25 to a NAV of R100. The 2x ETF dropped 40% (R61.09) of its NAV and dropped to a value of R91.64. The 2x ETF would require to reduce its exposure to R183.27 at the end of the T_3 to obtain the required return on T_4 (Yates, 2018).

Table 1 summarizes values, returns, and required exposures for a 2x ETF and 2x Inverse ETF.

Table 1: Example of Rebalancing of 2x and -2x Leveraged ETF

Day	Index Value	Index Return	L ETF 2x (Return)	LI ETF -2x (Return)	L ETF 2x (Value)	LI ETF -2x (Value)	L ETF 2x (Required Exposure)	LI ETF -2x (Required Exposure)
T0	100.00	0.00%	0.00%	0.00%	100.00	100.00	200.00	(200.00)
T1	110.00	10.00%	20.00%	(20.00%)	120.00	80.00	240.00	(160.00)
T2	125.00	13.64%	27.27%	(27.27%)	152.73	58.18	305.45	(116.36)
T3	100.00	(20.00%)	(40.00%)	40.00%	91.64	81.45	183.27	(162.91)

An interesting feature of Leveraged ETF rebalancing was that the ETFs will cause trading "momentum" and follow the direction of the index regardless of being a leveraged or inverse funds. As shown in Table 1, when the index experiences a gain, the Leveraged ETF will acquire more exposure to the underlying index, and the Inverse ETF will reduce short exposure, going long on the underlying. The converse is seen with losses incurred in the underlying index.

Another apparent characteristic was that the performance of Leveraged and Inverse ETFs did not track the underlying index’s performance over time but only over the rebalancing period, in this example, one day. This was due to the compounding nature of rebalancing, where the previous day's performance impacts the return profile of the current days trading (Hill et al., 2015).

In the following sections, the researcher provided an overview of the research conducted to date relating to the introduction of ETFs into financial markets, ETF trading activity, and the related impact on constituent assets. Although the review covered several economic outcomes, the primary focus would be the impact on the volatility of the underlying constituents.

2.2 Market Effects of the Introduction of ETFs

Several studies (Patro (2001); Hegde and McDermott (2004) Richie and Madura (2007); Madura and Ngo (2008); Bae et al. (2012)) have been carried out to investigate the implications of the introduction of ETFs on various factors that impact financial markets and the ETFs underlying constituents.

Firstly, by conducting an event study, Patro (2001) found that the introduction of the World Equity Benchmark Series (“**WEBS**”) ETF’s resulted in a significant and positive impact on average index returns post the introduction of the ETF as well as a decrease in the premiums of closed-end funds tracking the Morgan Stanley Capital International (“**MSCI**”) index. Patro attributed the decrease in premiums to investors viewing WEBS ETF as an alternative investment vehicle, thereby decreasing demand for the closed-end fund Patro (2001).

Moreover, a common point of interest regarding ETFs was the impact ETFs have on the liquidity and price discovery of the ETFs underlying assets. Hegde and McDermott (2004) examined whether the introduction of the DIAMONDS and Q's ETFs (ETFs that tracked the Dow Jones Industrial Average and the Nasdaq 100 indices) had an impact on the liquidity of the underlying securities. The researchers argued that the introduction of ETFs would increase arbitrage activity between the ETF and the underlying securities market, reducing adverse selection risk for market participants. It was noted that the liquidity of the underlying securities increased after introducing the ETFs, which was consistent with their arbitrage hypothesis Hegde and McDermott (2004).

Additionally, Richie and Madura (2007), Madura and Ngo (2008) and Bae et al. (2012) all conducted similar tests to determine whether the introduction of ETFs into financial markets had an impact on the liquidity of the underlying securities (the researchers also tested whether other economic outcomes where impacted by ETFs, with those findings discussed below). The results of these research papers were consistent in that the researchers found that the introduction of ETFs led to an increase in the liquidity of the constituent securities.

Richie and Madura (2007) attributed the increase in liquidity of constituent securities to the introduction of new market participants, bringing the additional trading volume and liquidity. Also, Bae et al. (2012) found that the introduction of ETFs led to an increase in the short selling of securities, which result in an increase in the liquidity of the constituent security. It is interesting to note how these three studies reached similar conclusions, each with its explanation for the increase in liquidity.

The impact of the introduction of ETFs on market volatility and its underlying assets has been a focus in several studies, the results of which vary significantly between authors.

Correspondingly, Lin and Chiang (2006) also examined the volatility of the constituent securities of the Taiwan 50 index before and after the introduction of the Taiwan Top 50 Tracker Fund ETF ("**TTT**"). Using a realised volatility and GARCH model approach 64 days pre- and post-the introduction of the TTT as volatility measures as inputs into a Variance Difference Ratio ("**VDR**") (a VDR is a ratio of the variance after the defined event to the variance before the defined event), they found that regardless of volatility measure used the volatility of the underlying securities increased post the introduction of the TTT ETF Lin and Chiang (2006).

In the same vein as Lin and Chiang (2006), Curcio et al. (2012) researched if the introduction of vanilla and leveraged ETFs had an impact on the volatility of the underlying securities in the US real estate sector. The authors research differed from Lin and Chiang (2006)'s research in terms of geographic and sector-specific focus, but the authors utilised the same testing methodology and employed the VDR test for the 64 trading days prior to and post introduction of the various ETFs on the US real estate sector constituent stocks. Like Lin and Chiang (2006), the authors found statistically significant increases in the volatility of the underlying real estate securities after the introduction of ETFs.

Following the research carried out by Lin and Chiang (2006), Matarutse (2014) investigated the potential implications of the introduction of ETFs on the volatility of the underlying shares in the South African context. The study specifically addressed the South African equity market and used Satrix Top 40 ETF, South Africa's first ETF, as its sample. The methodology used to test the hypothesis differed from Lin and, Chiang and Curcio's in that the author compared the volatility, measured using an E-GARCH model, of the daily returns of the underlying securities for the five years preceding the introduction of the Satrix Top 40 and the ten years post the introduction of the Top 40. In line with Lin and Chiang, contrary to Curcio, the author concluded that there was an overall decrease in the volatility of the underlying securities. The author also noted a decrease in volatility persistence of the underlying shares post the introduction of the Satrix Top 40 ETF. The author attributed the changes to the increased trading volumes and the reduction in asymmetric information associated with the introduction of the Satrix Top 40 ETF.

Although not direct, Richie and Madura (2007) investigated if the introduction of the PowerShares QQQ ETF ("**QQQ**") (currently known as Invesco QQQ or colloquially as "Cubes"), which tracks the Nasdaq 100, had an impact on the risk of the index's underlying securities. Using univariate regression models, they concluded that the constituents of QQQ experienced lower systemic risk due to improved liquidity and price stabilisation arising from the related ETF arbitrage trading activity from QQQ's introduction.

In a later study, Madura and Ngo (2008) explored whether the inception of ETFs had an impact on the valuation and trading volume, and with it the volatility, of constituent stocks. They found that due to the increase in trading volume and number of participants interested in the constituent stocks from the introduction of ETFs, there would a reduction in stock price volatility and lead to an increase in the value of the constituent stocks. Unlike the Richie and Madura (2007) research, which was limited to the QQQ, the authors extended their research to the ten most expensive component stocks in ETFs created between 1996 and 2004 in the USA. Using multivariate models to explain the variation of the returns, they found that inception of ETFs led to an overall increase in the value of the constituent stocks, partly due to a reduction in the constituent stock's volatility.

Contrary to the positive impacts on constituent securities valuation and volatility found by Richie and Madura (2007) and Madura and Ngo (2008), Bae et al. (2012) found that ETFs have a negative and statistically significant impact on the underlying securities firm value, this finding was especially pronounced with smaller securities. Using US equity market data for eight years, the authors investigated the effects of ETF holdings on the volatility, short interest, and liquidity of underlying securities. The authors referred to the generally accepted assertion that these three characteristics impact the value of a firm (underlying constituent security). Although the liquidity of the securities increased, a positive attribute, the researchers argued that the liquidity was driven by short selling, which negates any positive outcome. From a volatility perspective. the authors noted a positive relationship between the level of ETF holdings and the systemic risk of the underlying securities, i.e., with an increase in ETF holdings, there was a corresponding increase in systemic risk and volatility.

While the primary focus of this dissertation is the impact of equity markets, it is worthwhile to review similar research relating to ETFs of other asset classes. Corbet and Twomey (2014) looked at testing whether the introduction of commodity-based ETFs had a negative impact on the volatility of US commodity markets or improved the liquidity and the efficiency of the transfer of information. The authors cited concerns raised by commodity market regulators about ETFs' extensive holdings in commodities as well as the potential introduction of cross-commodity correlations that may have appeared with the introduction of mixed commodity ETFs. Similar to the research detailed above, the authors compared the volatility of daily returns in the commodity spot market pre- and post-the introduction of commodity ETFs. The authors found varying results based on the market size of the commodity. It was found that the volatility of smaller markets decreased due to the increase in liquidity, which echoes the findings of Richie and Madura (2007) and Madura and Ngo (2008), while the volatility of larger markets increased because of dominant ownership forces of ETFs.

2.3 The Link Between ETF Trading and Market Volatility

As mentioned in sections 2.1.2 and 2.1.3, ETFs have unique mechanisms in terms of their structure that drive specific trading activity. In the sections below, the researcher described what research has been conducted relating to the potential impact of these mechanisms. The first part of the review examines traditional ETFs, while the second examines Leveraged ETFs.

2.3.1 Impact of vanilla ETFs

Traditional ETFs initiate trading activity through the in-kind creation and redemption process, described in section 2.1.2.1, and ETF arbitrage activity, described in section 2.1.2.2. With their paper, Wermers and Xue (2015) decomposed ETF trading activity into "informed trades", trades that were effected based on the fundamental valuation of the stock, and "noise trades", trades not related to the fundamentals of the related stock, to determine if there is a relationship between noise trades (non-fundamental movements) relating to ETFs and the volatility of the underlying stocks. The authors suggested that noise trade would not impact the long-run fundamental value of the underlying security, while an informed trade would. Using this logic, the authors determined the concentration of noise trades to total trades and measured the impact ETF noise trades have on the volatility of the underlying index. The authors concluded that ETF noise trades impact the underlying index's volatility. However, the effects were minor and decay rapidly.

Following on the idea of decomposing trades, Xu et al. (2018) investigated whether the motivation behind ETF trading activity has an impact on the fundamental volatility (measured by the Variance Efficient Price Innovation measure) and total volatility (measured by the realised variance) of the underlying index. The authors analysed data from six major US indexes and their related ETFs and decomposed the ETF trades into three trading motives: (i) trades motivated by private information ("**Privately Informed**"), (ii) trades motivated by differences between investors in interpreting information ("**Disagreement-Induced**") and (iii) trades motivated by external liquidity forces ("**Liquidity Trades**"). The authors found statistically significant correlations between Privately Informed and Disagreement-Induced ETF trades and the fundamental volatility of the indexes. In contrast, the correlation between Liquidity Trades and the underlying indexes was found to be weak. Furthermore, there was little correlation between Privately Informed ETF trades and total volatility but a significant correlation between Disagreement-Induced and Liquidity Trades and total volatility. These findings indicated that the ETF trading and the motives driving ETF trades help explain changes in both fundamental and total volatility of the underlying indexes.

Extending the research carried out in their 2018 paper, Xu et al. (2020) examined if the different components of ETF trading had an impact on the volatility of the underlying index. Similar to Xu et al. (2018), the authors categorised the ETF trades into three categories: (i) Private Information, (ii) disagreement among investors, and (iii) investor impatience ("**Liquidity Trades**") and then determined if the trading volumes of each category could explain the volatility of the underlying index. The authors' study differed from Xu et al. (2018) in that it had a much narrower scope, limited to the CSI 300 and its three related ETFs. It only investigated the relationship between the ETF trade categories and total volatility (realised volatility), excluding fundamental volatility. It was declared that disagreement among investors and Liquidity Trades were significantly correlated (with disagreement among investors showing the strongest relationship) to the total volatility of the underlying index, while Private Information trades were not correlated. The results were consistent with those of Xu et al. (2018).

Furthermore, Wang and Xu (2019) also applied a similar categorization methodology to that employed by Xu et al. (2018). However, authors categorized ETF flows as either "backwards looking" or "forward-looking" flows and determined if these flows had an impact on the fundamental and/or total volatility of the underlying indexes. Backwards-looking flows referred to the authorised partner's ("AP") ETF creation and redemptions in response to market demand forces, while forward-looking flows were flows/trades over and above market demand and represented instances where the AP was actively trading in the ETF (AP arbitrage activity and trading based on information known by the AP). Using regression analysis, the researchers test if ETF flows, aggregated as well as categorised as backwards looking and forward-looking, were useful as an explanatory variable for the index's total volatility. It was observed that aggregated ETF flows were statistically significant in explaining the total volatility of the index and forward-looking ETF flows were more meaningful in predicting the index's total volatility than backwards-looking flows. Moreover, results indicated that the aggregated and decomposed ETF flows have an impact on the index's next trading day's fundamental volatility. The aggregated ETF flows were found to be positively related to the index's volatility, and the forward-looking flows were incredibly significant in determining the fundamental volatility of the index. These results aligned with the authors' expectations regarding the "active" nature of forward-looking trades made by APs' based on new information acquired. These findings supported the results of Xu et al. (2018) and Xu et al. (2020) in that informed trades (Privately Informed or Forward-Looking) impact the fundamental volatility of the underlying securities, while Noise or Liquidity Trades have a more significant impact on total volatility.

Moving away from the decomposition and motivation of ETF trades, Xu and Yin (2017) investigated the relationship between an ETF's trading volume and the underlying index's volatility and constituent securities' volatility. The authors selected the S&P 500 index and its related ETFs as they accounted for a significant percentage of the US equity market. Using time series analysis, the authors uncovered that ETF trading volumes had a statistically significant effect on the volatility of the underlying index. The authors also included an analysis of the lagged effect of ETF trading volumes on the volatility of the S&P 500 index. Using a Vector Autoregressive estimation methodology, they identified two-way Granger Causality, where ETF trading volumes Granger Causes an increase in the volatility of the S&P 500 index, which in turn causes an increase in ETF trading volume. The evidence of a two-way Granger Causality was consistent with research on the relationship between derivative markets and their underlying markets Xu and Yin (2017).

The following section discussed the potential impact of Leveraged ETF's rebalancing activities on the ETF's underlying constituents.

2.3.2 Impact of leveraged ETFs

After the crash of 1987, Black Monday, the market looked to blame the increase in volatility on Index futures. However, most research on this concluded that index futures had little, if any, impact on market volatility. In response to the increased volatility during the 2008 Global Financial Crisis ("GFC"), the market turned its attention to leveraged ETFs as the cause of excessive volatility. The research summary below looks at determining the validity of the assertion above.

In this regard, Cheng and Madhavan (2009) revealed the dynamics of, and the effects leveraged and inverse ETFs have on financial markets. The paper investigated whether the daily rebalancing of these ETFs had an impact on market volatility. Due to the nature of leveraged and inverse ETFs investment mandate, the authors assumed that the ETFs would most likely carry out rebalancing trades as close to the end of the day as possible to ensure the following day's return matched its mandate. The authors ran regression analyses to determine if trading from the rebalancing caused variability in returns at the end of the trading day. The authors found strong evidence that the direction and magnitude of returns, as well as the volatility of returns, were positively related to the volume of the leveraged and inverse ETFs rebalancing activity Cheng and Madhavan (2009).

Shum et al. (2016) extended the work of Cheng and Madhavan (2009) by investigating whether leveraged ETFs intensified end-of-day volatility and determining if there was a profitable front-running strategy that opportunistic traders could have employed. In agreement with Cheng and Madhavan (2009), Shum et al. (2016) found that the rebalancing trading activity of leveraged ETFs resulted in increased volatility at the end of the trading day. They also found that where the proportion of leveraged ETF rebalancing trading volume to total trading volume was greatest, the more severe the impact on the volatility for the period would have been. Regarding their profitable trading strategy, the authors found that predatory traders could have employed a profitable front-running strategy by taking advantage of the known rebalancing activities of leveraged ETFs.

Furthering Cheng and Madhavan (2009)'s quest for clarity, Flores (2015)'s study on leverage ETFs, their mechanics, returns, and potential impact on the volatility of the underlying shares of the DAX index brought a European perspective to the body of research. The author investigated if leveraged ETFs increase the daily volatility of the underlying assets as well as if the rebalancing mechanism of leveraged ETFs is causing increases in volatility near the end of the trading day. Although the methodology employed was inspired by the works of Cheng and Madhavan (2009) and Shum et al. (2016), the author used a difference-in-differences approach to determine changes in the daily volatility of underlying stocks before and after the introduction of the sample leveraged ETFs. To test whether leveraged ETFs increase the volatility of underlying stocks at market closing, they split the trading day into 18 different periods and used the volume of rebalancing flows to explain changes in volatility. The results showed that there was a statistically significant positive effect on volatility where a leveraged ETF had been introduced and that there was a relationship between the proportion of leveraged ETF rebalancing trades and the underlying stocks' volatility Flores (2015).

Contrary to this, Trainor (2010) found that there had not been any significant increases in volatility after the introduction of leveraged ETFs. While volatility increased during the GFC, it returned to normal levels, although the prevalence of leveraged ETFs increased. Trainor conceded that the S&P 500 was a large market, and if leveraged ETFs formed a greater proportion of the flows, the results of the study may have been different Trainor (2010).

By narrowing the scope of tests to the US Real Estate Investment Trust ("**REIT**") sector, Boney-Dutra et al. (2013) asserted that the impact of Leveraged ETFs rebalancing on the volatility of the constituents of the REIT index relative to the broad market index was more pronounced than the broad market index.

Like Trainor, Li and Zhao (2014) determined whether leveraged ETFs had an impact on the liquidity and volatility of their constituent stocks. To test this, the authors conducted two types of tests, an event study, where a series of liquidity and volatility measures were measured before and after the introduction of the sampled leveraged ETFs at varying periods and a regression analysis on daily and intraday returns to determine how the trading activity of leveraged ETFs may have influenced the liquidity and volatility in the last hour of trading. Consistent with Trainor (2010)'s results: the event study approach indicated no economically significant change to the volatility and only a slight widening of the bid-ask spread, indicating a potential decrease in liquidity. The regression analysis, using the standard deviation of the ETF on the day and the day before as explanatory variables for the standard deviation of returns for the underlying stock, yielded statistically insignificant relationships between the volatility of the ETF and the volatility of the underlying stock at a daily level, again supporting Trainor (2010) findings.

2.4 Summary

At the start of the chapter, the history of ETFs and the innovations in product development over the years were briefly described. Then, it drilled down into the mechanics of traditional and leveraged ETFs to provide context for some of the context covered in the sections that investigated the impact of introducing ETFs into financial markets as well as the impact of ETF-specific trading activity on the underlying assets. The next chapter covered the accumulation and selection of data as well as the methodology chosen to test the research questions.

3 DATA AND METHODOLOGY

3.1 Introduction

This chapter summarized the data-gathering process, including where the data was sourced from, the tools used to extract the data, and the limitations and assumptions of the dataset used for this dissertation's analysis. It also provided an overview of the methodology used in testing the paper's hypothesis, the literature that inspired the tests and how the methodology differs from that literature.

The study aimed to assess the potential influence of non-fundamental pricing pressure, attributable to ETF's unique dynamics, on the volatility of individual stock prices. In particular, the researcher investigated whether such pressure is associated with ETF-specific events of ETFs that hold these stocks and if there is an association between the degree of ETF ownership and the underlying stock's volatility.

The analysis was based on time series data from major South African banks, namely ABSA Bank Limited (ABG), Capitec Bank Holdings Ltd (CPI), FirstRand Ltd (FSR), Investec Ltd (INL), Nedbank Group Ltd (NED), and Standard Bank Group Ltd (SBK).

3.2 Data Gathering

The researcher identified and sourced various datasets to perform the tests outlined in section 3.3. This section described the data aggregation and validation process followed, and provided the reader with a guide on how the dataset can be recreated.

3.2.1 Selection and aggregation

3.2.1.1 Scope of ETFs

Two sources were utilised to obtain a list of ETFs that would form part of this dissertation's testing: 1.) Bloomberg ETF Screen and 2.) The Exchange Traded Fund list was obtained from ETFsa.co.za (https://www.etfsa.co.za/ETFs.htm).

The following search criteria for the Bloomberg ETF screen were selected: Fund Type - All; Asset Class - Equity & Mixed Allocation; Geo Focus - South Africa; Exchange - South Africa.

The list obtained from ETFsa.co.za included a region and class indicator which enabled the identification of domestic equity ETFs (foreign equity, domestic and foreign fixed-income and domestic commodity ETFs were excluded), which speaks to the scope of this dissertation.

A reconciliation between the two lists was performed in an effort to make the list as complete as possible. The following variances between the lists were identified:

1. ABSA NewFunds Volatility Managed Defensive Equity, ABSA NewFunds Volatility Managed High Growth Equity, ABSA NewFunds Volatility Managed Moderate Equity, and Coreshares Preftrax ETFs were included in the ETFsa.co.za list but not on the Bloomberg ETF screen. These securities are included in the scope of this dissertation as they are ETFs that track South African equity securities.
2. Fortress REIT was included in the Bloomberg ETF screen but not on the ETFsa.co.za list. This security was excluded from the scope of this dissertation as it is a REIT and not an ETF.

The above scoping exercise results in 32 ETFs (this dissertation's "ETF universe") being included in the scope of this dissertation; refer to Appendix 1 - Applicable South African Equity

ETFs for the list of ETFs. The list of ETFs allowed for identifying the underlying indices and equity securities required for testing this dissertation's hypotheses.

3.2.1.2 Sample Period

The purpose of this dissertation is to investigate whether there is an association between the volatility of the identified stocks and the degree of ETF ownership as well as the volatility of the identified stocks and ETF-related trading activity. The sample period for this dissertation covers the period 1 January 2009 to 31 March 2020.

3.2.1.3 ETF Specific Event Data

The inception dates of the individual ETFs were obtained from the MDD documents obtained from the ETF managers' website.

With the ETF trading data, it was not possible to identify the trading days when the ETFs were going into the market to acquire/dispose of shares as part of the ETF share creation and redemption process ETFs, however, publish an announcement on the Stock Exchange New Service ("SENS") notifying the market that the ETF has launched, added shares (creation) or de-listed shares (redemption). For this dissertation, these SENS announcements are a proxy for the timing of the ETF creation and redemption trades.

The ETF SENS announcements were obtained from Bloomberg by executing function "CN" for the ETF's ticker and screening the subject of the announcements for listing announcements, additions and de-listing of shares, as well as portfolio rebalancing announcements. All available SENS announcements relating to the identified stocks on Bloomberg were reviewed from the earliest possible date to 31 December 2019. Where Bloomberg's data did not include announcements from earlier periods, from the inception of the ETF to the earliest date on Bloomberg, searches on moneyweb.co.za and sharedata.co.za were carried out to source earlier announcements. Where there was still a gap between the inception date and the last SENS announcement identified by Bloomberg, moneyweb.co.za, and sharedata.co.za, a request was sent to the JSE for the remaining announcements, however, no response was received from the JSE.

3.2.1.4 Underlying Constituent Security Data

This dissertation is concerned with the potential impact ETFs have on the South African equity market and its constituent securities, with a specific focus on the major South African Banks, ABSA Bank Limited (ABG), Capitec Bank Holdings Ltd (CPI), FirstRand Ltd (FSR), Investec Ltd (INL), Nedbank Group Ltd (NED), and Standard Bank Group Ltd (SBK).

The identification of the individual securities that formed part of the ETFs was needed to carry out this dissertation's various tests.

ETF managers were contacted and requested to provide the ETF's historical holding data at an individual security level from inception to 31 March 2020. Except for one, most ETF managers could not provide the requested information, citing capacity constraints or legal constraints in providing index-linked data without the JSE's or Financial Times Stock Exchange's ("FTSE") permission.

Morningstar, Inc. is a third-party data provider that provides fund data to investors. A request was sent to the South African Morningstar, Inc. sales team requesting ETF's historical holding data for the same period. Morningstar, Inc. was able to provide the historical holdings data from 2008/2009 onward.

Bloomberg provided historical fund holdings data. The ETF constituent data was available using the "PORT HD" function and selecting the trend date option. Similarly, the Morningstar, Inc. dataset was limited by the available information period. For the ETFs with inception dates pre-2008/2009, a request was sent to the JSE data team to provide the holdings data for the period covering inception to the first date available from the Morningstar, Inc. and Bloomberg datasets. No response was received.

In 2007, Richie and Madura (2007) assumed that the ETF holdings would reflect the constituents of the underlying index. For this dissertation, where the holdings data was not available, the ETF constituent holdings were estimated using the index constituent weights and multiplied the weights to the market capitalization of the ETF to get an estimated Rand value of each constituent security in the portfolio of the ETF.

Daily closing prices of the identified constituent securities were extracted from Thomson Reuters ("Reuters") Datastream using the Time Series Request function. The required data was extracted by inputting the security's Reuters symbol in the "Series/List" field, inputting "P" (which refers to the "Price-Trade" Data Type) in the "Data Type field, selecting the required start date and end date and set the Frequency to "Daily". Daily data was selected for the analysis as the cadence of ETF event data, and ETF Ownership data is only captured daily and monthly. Where there were missing values in the daily closing pricing data, the daily closing prices were imputed using the last observation carried forward method Wongoutong (2021), where the most recent available observation was used to replace the missing value.

3.2.1.5 ETF Ownership Data

In order to determine whether there is an association between the degree of ETF ownership of a constituent stock and the constituent stock volatility, the degree of ETF ownership variable needs to be determined. For clarity, the ETF Ownership variable reflects the percentage ownership of the identified stocks held by ETFs within this dissertation's ETF universe. As this metric is not readily available in the South African context, for the study, this variable was constructed by applying the following process:

The percentage holdings of each constituent stock in each ETF in the ETF universe were obtained from the data from Morningstar and Bloomberg (described above) over the sample period. To determine the completeness of the holdings data, each period of ETF holding data provided was reconciled to 100%. Where immaterial variances were identified, the difference was allocated to the cash holdings of the ETF. Where material variances were identified, the data provider was contacted to determine what caused the variance. Additionally, spot checks against either the MDD of the ETF or data received directly from the Investment Manager of the ETFs were performed to check the validity of the data provided. Instances were identified where the labels of constituent stocks were not included in the dataset provided by the third-party data provider, missing labels were identified and appropriate stock ticker labels were added to the dataset. Instances were identified where different ticker labels were used between different ETFs, multiple tickers for single constituents were grouped under a single ticker to prevent the incorrect percentage holdings by ETFs was determined. All values relating to fees in the ETF were allocated to "Cash" as it did not have an impact on the percentage of constituent stocks held by ETFs. Where ETF holdings data was not available for a specific date, the ETF holdings data from the previous data point was used as the value for the missing date.

The time series market capitalisation data of each ETF was extracted from Bloomberg data services. Where Bloomberg did not have a data point, Refinitiv Eikon was used to source the missing data points. Where the data was unavailable from Bloomberg or Refinitiv Eikon, the ETF market capitalization data was sourced from the ETFs listing/launch announcement.

By multiplying the market capitalisation value of each ETF by the percentage holding of the constituent stock percentages, an approximate Rand value held by each ETF of the constituent stock was calculated. Summing the Rand values of all of the ETFs in the ETF universe provided an approximate Rand value of the identified stock held by the ETF market at each point.

The market capitalisation of each identified stock over the test period was extracted from Bloomberg. By dividing the approximate Rand value of the identified stocks held by the ETF market by the market capitalisation of the identified stock, the approximate percentage of the identified stock held by the ETF market has been determined.

3.2.2 Limitations of Data

The data-gathering process undertaken for this dissertation identifies several limitations and issues with source data. The frequency of data between variables was different, which required simplifying assumptions to be made in order to run our analysis. The availability of the ETF constituent stock data limited the sample period chosen for this study.

Regarding the potential association of ETF Specific events on constituent stocks, trading level data of the ETFs was unavailable, so being able to determine the impact on the volatility of non-fundamental trades, as followed in Ben-David's paper (Ben-David et al., 2018), was not possible. With specific reference to ETF-specific event data points, the completeness and timing of the ETF-specific event dataset are questionable due to there being no requirement for ETFs to report-/announce events, and the announcement only included an effective date, which may not be the same as the trade date.

As noted in 3.2.1.4, the constituent dataset was reliant on the accuracy of third-party data providers, although accurate for the most part, included inconsistencies and data points inconsistent with ETF-reported data.

3.3 Data Analysis

This research carried out two tests: firstly, to assess if there is an association between the level of ownership of ETFs in constituent stocks and the underlying constituent stocks volatility, and secondly, to assess if there is an association between ETF-specific events and the underlying constituent stocks volatility. This section outlined this study's research design and references to research papers that inspired this dissertation's methodology. The specific objectives of this dissertation designed to achieve the research goals are outlined in the following three steps:

Step 1: Estimating constituent stock price volatility for each bank. GARCH (Generalised Autoregressive Conditional Heteroskedasticity) time series models were developed to analyse the volatilities of each bank using daily historical stock pricing data.

Step 2: Investigating the relationship between ETF-specific events and bank stock volatility. GARCH models were constructed that included ETF-specific events (presence or absence of ETF event on a particular day) as an external regressor to explain or predict the volatility of South African banks' stock returns.

Step 3: Analysing the relationship between ETF ownership percentage and constituent stock volatility of South African major banks. We constructed GARCH models that included ETF Ownership Percentage and ETF Specific Event's as an external regressors to explain the volatility of South African bank stock returns.

3.3.1 Measuring Volatility

It was aimed to determine the association between ETF ownership levels and ETF-specific events with the volatility of constituent underlying stocks. As such, the determination of the underlying stocks' volatility is a critical input into this study. This dissertation uses a GARCH model to determine the volatility.

Financial time series, including variables such as stock prices, tend to exhibit a characteristic known as volatility clustering. As the name suggested, volatility clustering involves periods where the level of volatility of a data series, in this study, daily closing stock prices, persists. Noted in Panel 2 and Panel 3 , daily closing stock prices for bank stocks experience periods of high volatility and periods of low volatility that are grouped by Hanck et al. (2023). Cont (2001) referenced numerous empirical studies analysing financial data such as stock prices, indicating that the autocorrelation function, a metric often used to measure volatility clustering, of stock prices “remains positive and decays slowly, remaining significantly positive over several days, sometimes weeks” indicating the presence of volatility clustering.

As noted by Ramos (2021) and Engle (1982) the Autoregressive Conditional Heteroskedasticity model determines “time-dependent volatility as a function of the previously observed volatility.” by taking an autoregressive approach to model volatility, estimating current volatility as a function of prior period volatility.

Ramos (2021) noted that the Generalised Autoregressive Conditional heteroskedasticity model (GARCH) is an extension of Engle’s ARCH model by Bollerslev (1986) to extend the ARCH model from a single lag term to an ARCH model that can accommodate multiple lag terms.

Engle (2001) also asserted that econometricians historically utilised the Ordinary Least Squares model in the determination of volatility in financial applications; however, due to the requirement for econometricians to provide forecasted data, the OLS model is not fit for purpose, and the GARCH family of models provides a more appropriate estimate of volatility.

OLS models are widely used due to their simplicity and ability to be applied to a broad range of applications Ramos (2021). The OLS approach aids in quantifying the extent to which one variable responds to a change in other variables Engle (2001). These models, however, have several restrictive assumptions that limit the models’ ability to provide accurate results. OLS assumptions include linearity (assumption that there is a linear relationship between the result and the explanatory variable), independence (errors are not connected or dependent on each other or predictors), homoscedasticity (variance across a range of error terms is constant), normality (error terms are normally distributed), and no multicollinearity (there is no correlation between independent variables). Engle (2001) notes that these assumptions often do not fit real-world data, the result of which may be distorted model outputs. Engle (2001) stated that applying an OLS model with its limiting homoscedasticity assumptions to a dataset with heteroskedasticity properties results in unbiased OLS regression coefficients. However, it comes at the cost of the standard errors and confidence intervals becoming overly narrow, providing researchers with a "false sense of precision". Ramos (2021) declared that the GARCH family of models, namely ARCH (Autoregressive Conditional Heteroscedasticity) and GARCH (Generalized Autoregressive Conditional Heteroscedasticity) models, treat the variance in the error terms as a variable to be modelled results in deficiencies of OLS models to be remedied while providing a prediction of the variance.

For this reason, using a GARCH model in the context of real-world financial time series data is preferred over the use of a simplified OLS model.

Overview of ARCH / GARCH Family of models

Engle (1982) introduced ARCH models in a study focusing on inflation rates to tackle heteroscedasticity issues. The primary goal of these models is to estimate time-dependent volatility by considering the past volatility observed in the data. The original ARCH model characterized the variance of errors in a regression model as a linear function of lagged values of squared regression errors (Engle, 1982).

In mathematical terms, this can be represented as (Ramos, 2021):

$Rt = \sigma_t \varepsilon_t$ (conditional mean) (Equation 1)

$\sigma^2_t = \alpha_0 + \alpha_1 R^2_{t-1} + \ldots + \alpha_q R^2_{t-m}$ (conditional variance) (Equation 2)

Where:

R_t represents the returns of a financial asset at a specific time period

σ^2_t represents the conditional volatility at time t, and

α_q denotes the various parameters of the ARCH models, estimated from empirical data.

The ARCH model has a specification for both the conditional average and the conditional variance. The ARCH model is a statistical model designed for time series data that describes the volatility of the error term as a function of previously observed error terms.

ARCH models may be simplistic; many cases, they require significant parameters or lags to explain the volatility of share price returns (Ramos, 2021). Bollerslev extended Engle's ARCH model to incorporate an ARMA process, making this new model, GARCH, a generalized ARCH model, which allowed for many lags.

In mathematical terms, a GARCH (p,q) model is defined as follows (Ramos, 2021):

$R_t = \sigma_t \varepsilon_t$ (conditional mean) (Equation 3)

$\sigma^2_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \, R_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$ (conditional variance) (Equation 4)

where

ε_t is a sequence of independent and identically distributed random variables with mean zero, variance equal to one, and $\alpha_0 > 0$ for $i > 0$. ε_t is assumed to either the standard normal, standardized student's t, or generalized error distribution (Ramos, 2021)

By providing a generalized approach, the GARCH model provides a simpler model with fewer parameters, which is more favorable and understandable than complex models with numerous parameters.

Although these models improved the measurement of the volatility of financial time series, a shortcoming of the ARCH-GARCH family raised by Francq and Zakoian (2019) is that the models model the conditional variance as a linear function of the squared past observations and their implied oversimplification of real-world data symmetry by assuming that positive and negative shocks impact volatility in the same way.

Alternative volatility models have been developed to address various stylized financial time series data facts to solve this problem. These include the exponential GARCH model (EGARCH) proposed by Nelson (1991) and the GJR-GARCH model Glosten et al. (1993). EGARCH and GJR-GARCH models both allow for asymmetric effects and follow the idea that stock price volatility may increase as a result of bad news and stock price volatility can decrease after good news (Nelson, 1991).

The EGARCH(p,q) is mathematically expressed as per Ramos (2021):

$$\ln(\sigma|t^2) = \alpha_0 + \sum_{i=1}^{q} \frac{|R_{t-i}| + \delta_i R_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^{p} \beta_j \, ln(\sigma_{t-j}^2) \qquad \text{(Equation 5)}$$

Where

δ_i parameter implies the leverage effect of R_{t-i}, where volatility is more sensitive to negative returns that positive returns (Black, 1976).

The GJR-GARCH model is expressed as follows per Ramos (2021):

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q}(\alpha_i + \gamma_i N_{t-i}) \, R_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \qquad \text{(Equation 6)}$$

N_{t-i} is an indicator variable where a value of one is given if $R_{t-i} < 0$ and zero for all other instances. Therefore, a positive α_{t-i} and negative R_{t-i} adds to volatility, (σ_t^2) and $(\alpha_i + \gamma_i)R_{t-i}^2$ respectively. Where $\gamma_i > 0$, the GJR model will apply zero as a hurdle discern between the consequences of past shocks (Tsay, 2005).

The analysis carried out in this dissertation is based on real-world financial time series data that possesses heteroskedastic properties. In this circumstance, using a GARCH-based approach to estimating volatility is appropriate.

Likewise, the process of estimating the volatility of the Ibovespa, a market index for the Brazilian equity market, applying a GARCH volatility model using the statistical program "R". Ramos (2021) also referenced the various R scripts that were used in their study and where they can be accessed. This dissertation follows the process described by Ramos (2021) by following the various steps to determine the volatility of the identified stocks, further noted as the “Estimation Process”.

The Estimation Process in Ramos (2021)’s paper was split into three stages: Data, Model Identification, and Forecasting. For this study, the process stops at the end of the Model Identification stage as this dissertation does not require the volatility of the identified stocks to be forecasted.

The first stage of the Estimation Process, the Data stage, describes the steps to be followed to import the stock price data into R, generate a time series of returns by determining the daily log returns, calculating descriptive statistics of the share price data and summarising the results of the identified stocks into tables and graphs that provide an overview of the pricing and daily return data, refer to Panel 1 and Panel 2 for plot graphs of share price and daily log return data.

The second stage of the Estimation Process, the Model Identification stage, explained the process of identifying which GARCH is most appropriate for our dataset.

Ramos (2021) explained that before a GARCH model can be used to determine the volatility of a dataset, the dataset needs to be tested to see if the dataset possesses ARCH characteristics. A Lagrange Multiplier test is used with the null hypothesis being that the dataset has no ARCH effects and looks at a p-value close to zero.

Assuming that the Lagrange Multiplier test results in a dataset with ARCH effects, this research can move forward in determining what GARCH model is best-suited to the dataset. Ramos (2021) asserted that the selection of the best-suited GARCH model can be done using “goodness-of-fit indicators”. The paper referenced the Akaike Information Criteria (AIC) and Bayesian Information Criterion (BIC) indicators as the measures to be used in determining which GARCH model included in R's library is best-suited. In determining the goodness-of-fit, this research would be looking for the model that results in the lowest AIC and BIC coefficients.

Having identified the most appropriate models through the described process, we are now able to use the fitted models to address our research objectives

3.3.2 ETF Ownership and Volatility

As described in 2.1.2.1, ETFs are uniquely related to underlying equity market. In the research paper by Ben-David et al. (2018), the authors investigated the relationship between the level of ownership of constituent securities by ETFs and the volatility of the underlying securities. This dissertation seeks to replicate the experiment for the South African market.

Ben-David et al. (2018) performed an OLS regression analysis to determine if the degree of ETF ownership of constituent stocks explained the volatility of constituent stocks.

Moreover, Ben-David et al. (2018) used ETF ownership as the independent variable to explain the variation in the volatility of the constituent security. This dissertation uses the level of ETF ownership as an external regressor to explain the volatility of the banking stocks that forms part of this study. The level of ETF ownership was determined using the following formula:

$$ETF\ Ownership_{i,t}\ (EO) = \frac{\sum_{j=1}^{J} w_{i,j,t} AUM_{j,t}}{Mkt\ Cap_{i,t}} \qquad \text{(Equation 7)}$$

$J = the\ ETFs\ that\ hold\ security\ i\ in\ their\ portforlio\ at\ the\ end\ of\ the\ quarter$

$W_{i,j,t} =\ the\ weight\ of\ the\ security\ in\ ETF\ j\ portfolio$

$AUM_{j,t} =\ Assets\ Under\ Management\ of\ ETF\ j\ at\ the\ end\ of\ the\ quarter$

$Mkt\ Cap_{i,t} =\ the\ market\ captilisation\ of\ the\ security\ at\ the\ end\ of\ the\ quarter$

As noted in 4.3 and 4.5, GARCH models were developed that included ETF ownership as an external regressor to explain or predict the volatility of major South African banks' stock returns. By estimating the parameters of the GARCH model with external regressors, one can assess whether these factors are statistically significant in explaining the volatility of the major South African bank stocks returns series. This helped understand the potential relationship between the external factors and the stock volatility.

3.3.3 ETF Specific Events

Following on from testing whether the level of ETF ownership impacts on the volatility of the underlying security, another area of interest was to investigate whether trading activity unique to ETFs impact the volatility of the underlying security through non-fundamental pricing pressure. This dissertation investigates whether the launch of an ETF, creation and redemption of ETF shares and rebalancing activities of ETFs impacts the volatility of the underlying volatility.

A significant amount of research has been done in introducing ETFs on equity markets which have used several methods to test their hypotheses. For instance, Lin and Chiang (2006) used a Variance Difference Ratio to compare the volatility of constituent securities before and after the introduction of ETFs in the Taiwanese equity market. Likely, Li and Zhao (2014) used a univariate regression analysis in an event study approach. They compared the volatility of the constituent securities at varying event windows (-5 days; +5 days) and (-50 days; +50 days) pre and post the introduction of the Leveraged ETFs. Li and Zhao (2014) also employed the standard deviation of intraday trade price returns as the volatility measure. Flores (2015), on the other hand, utilised difference-in-differences to determine if the inception of leveraged ETFs impacted the volatility of the constituent securities. The authors investigated whether the volatility of the DAX (German equity stock index) of constituent stocks relative to other stocks in the German equity market and relative to the historical volatility of the constituent stocks before and after the introduction of leveraged ETFs.

Furthermore, Ben-David et al. (2018) performed a similar test to ascertain whether ETF arbitrage activity, an activity unique to ETFs, acted as a catalyst that transferred volatility between the ETF and the underlying equity market. Using the ETF Specific Events data, which incorporates ETF creations, redemptions, and rebalances, as a proxy for ETF trading, as an external regressor to explain the volatility of the banking stocks that formed part of this study.

As noted in 4.3 and 4.4, GARCH models that included ETF-Specific Events as an external regressor to explain or predict the volatility of major South African banks' stock returns were constructed. By estimating the parameters of the GARCH model with external regressors, one can assess whether these factors are statistically significant in explaining the volatility of the major South African bank stock returns series. This analysis aids in discerning the potential relationship between the external factors and the stock volatility.

3.4 Conclusion

This chapter detailed the various data points used for testing the potential impact ETFs has on the volatility of the underlying securities, where the data was sourced from, data validation checks, and the steps followed to construct the datasets.

A summary of the various methods employed by other researchers in related topics was provided, where those methods were used in this dissertation, deviations from those researchers were noted and the rationale for the deviation was provided. Brief descriptions of the methods employed to test this dissertation's hypothesis were provided and relevant research was referenced.

The next chapter detailed summary statistics of the datasets and provided the results of whether the introduction of ETFs and ETF-specific events have an impact on the volatility of the ETF constituent securities.

4 RESEARCH FINDINGS, ANALYSIS AND DISCUSSION

4.1 Introduction

This chapter presents the findings of this study concerning the literature presented in Chapter 2 and methodology presented in Chapter 3.

As described in Chapter 3, this study follows three steps in answering whether there is an association between the volatility of constituent stocks and 1.) the level of ETF ownership of a constituent stock and 2.) ETF Specific events.

Step 1: Estimating constituent stock price volatility for each bank. The researcher developed GARCH time series models to analyse the volatilities of each bank using daily historical stock pricing data. The methodology applied in this study follows the approach outlined in Chapter 3 (Ramos, 2021).

Step 2: Investigate the relationship between ETF-specific events and bank stock volatility. The researcher constructed GARCH (Generalized Autoregressive Conditional Heteroskedasticity) models that included ETF-specific events (presence or absence of ETF event on a particular day) as an external regressor to explain or predict the volatility of South African banks' stock returns.

Step 3: Analysing the relationship between ETF Ownership Percentage and constituent stock volatility of South African major banks. GARCH models were constructed that included ETF Ownership Percentage as an external regressor to explain or predict the volatility of South African banks stock returns.

4.2 Estimating constituent stock price volatility for each bank

For each major South African bank - ABSA Bank Limited (ABG), Capitec Bank Holdings Ltd (CPI), FirstRand Ltd (FSR), Investec Ltd (INL), Nedbank Group Ltd (NED), and

Standard Bank Group Ltd (SBK) - we estimated stock price volatility using daily closing prices extracted from Thomson Reuters. Daily data was chosen for its alignment with the frequency of ETF event and ownership data, which are recorded daily and monthly, respectively. In cases of missing values in the daily closing price data, we employed the Last Observation Carried Forward method (Wongoutong, 2021), where the most recent available observation replaced the missing value.

The figures in Panel 1 below showed the plots of each bank's daily share price over the sample period 1 January 2009 to 31 March 2020. As noted above, where there were missing data points in the Last Observation Carried Forward method, these instances were marked in red and labelled as imputed values in Panel 1 below. Missing values were attributable to incomplete data obtained from data providers as well as public holidays. It is noted that the treatment of public holidays as trading days, by imputing share prices on public holidays, in this dissertation is not consistent with the methodology used in other studies. The shortcoming of the inconsistent methodology used in this paper is included Chapter 6 as an area for further study where future researchers could re-run the tests with a methodology consistent with other studies.

Panel 1- Plots of South African Major Banks Share Prices after data imputation

Share Prices of ABG

Total nominal arithmetic return over the sample period -30.72%

Red circles represent missing values where LOCF method has been applied

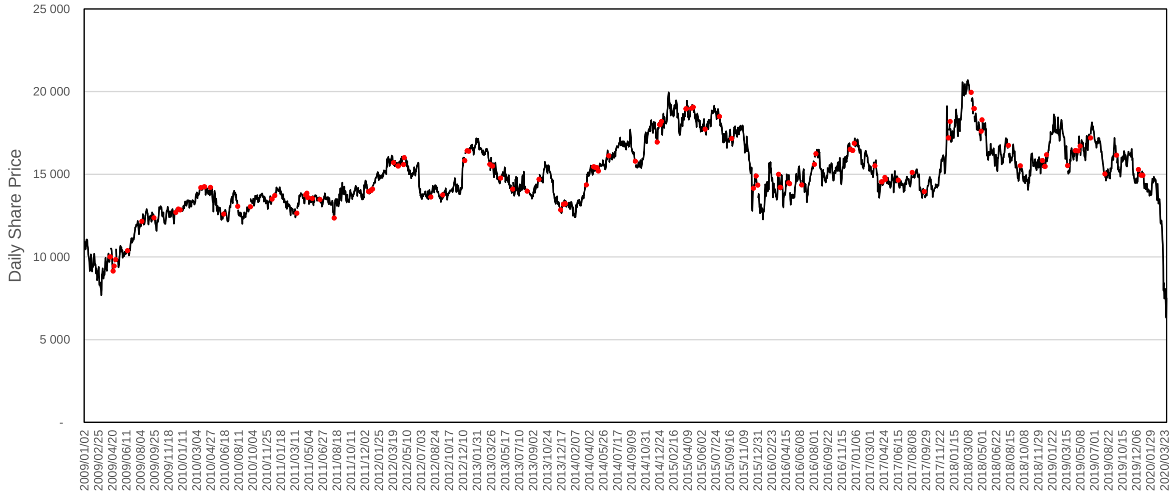


Figure 8: Share Prices of ABG (Source: Reuters)

Share Prices of CPI

Total nominal arithmetic return over the sample period 3112.91%

Red circles represent missing values where LOCF method has been applied

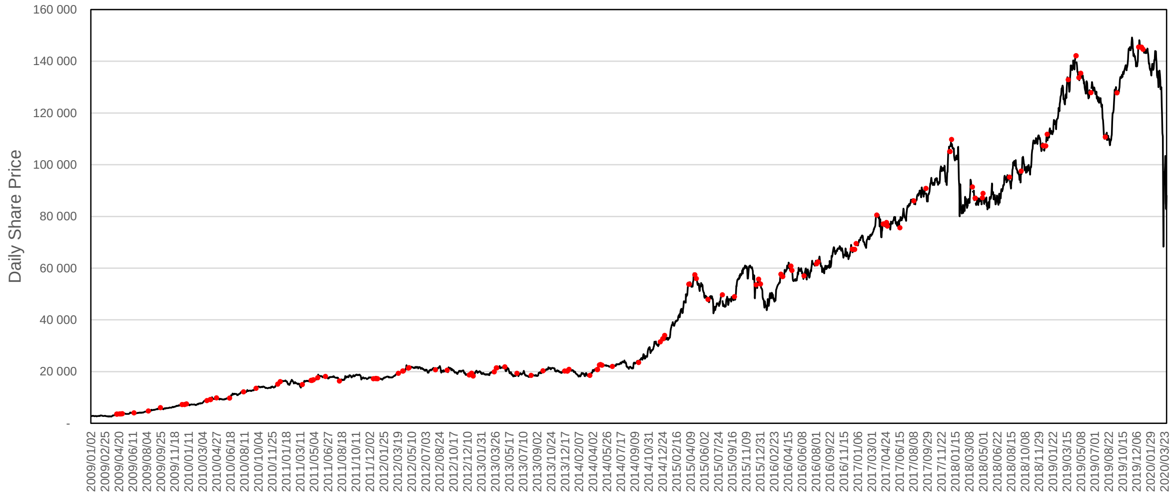


Figure 9: Share Prices of CPI (Source: Reuters)

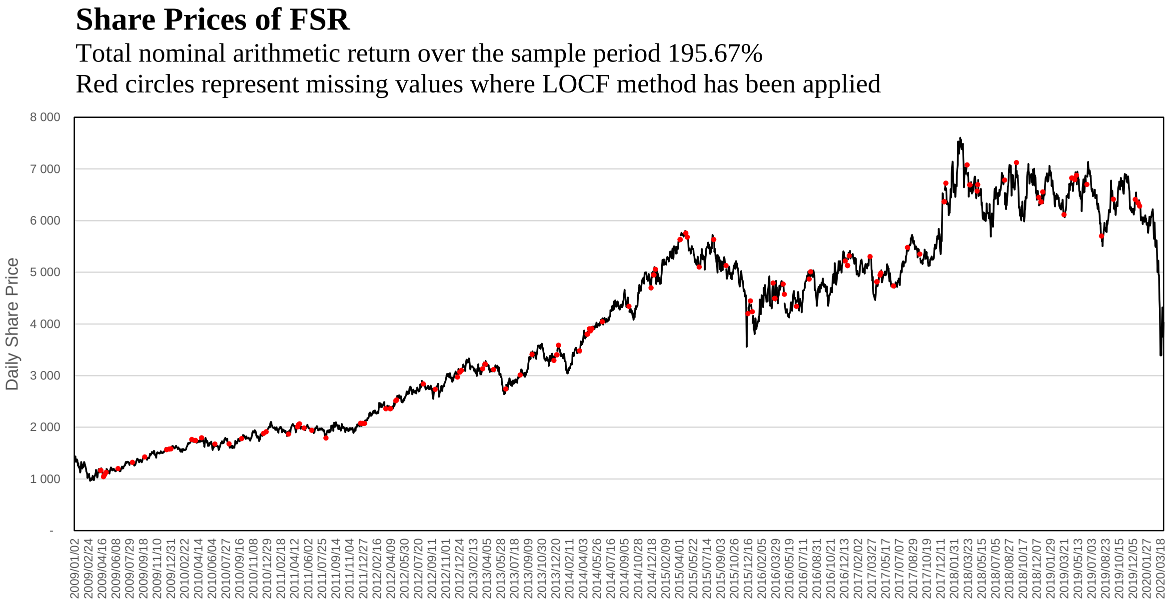


Figure 10: Share Prices of FSR (Source: Reuters)

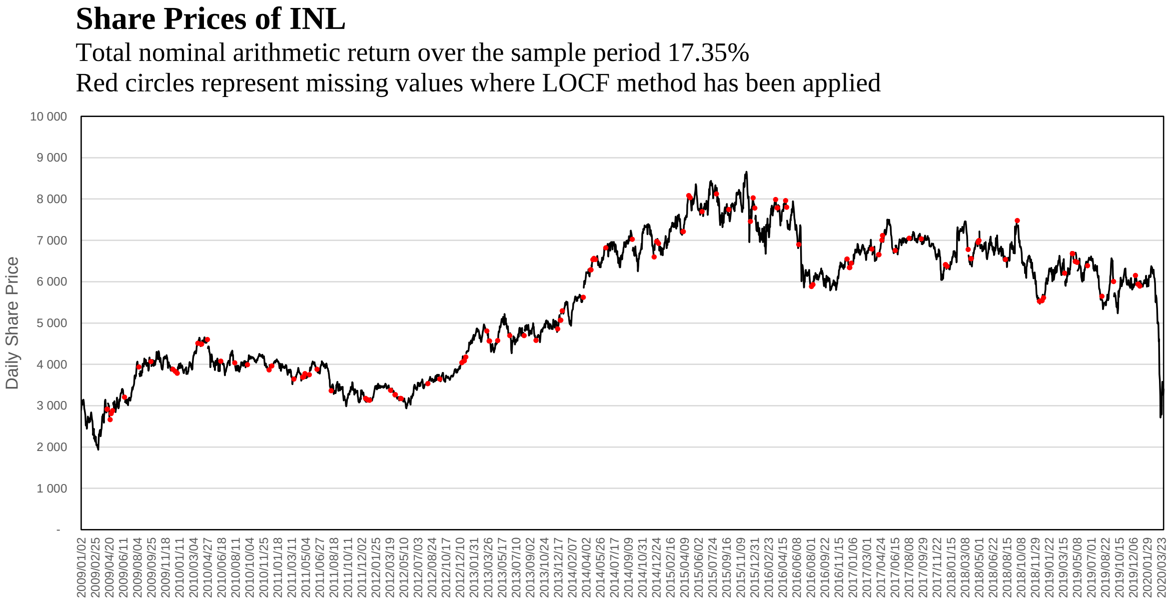


Figure 11: Share Prices of INL (Source: Reuters)

Share Prices of NED

Total nominal arithmetic return over the sample period -12.62%

Red circles represent missing values where LOCF method has been applied

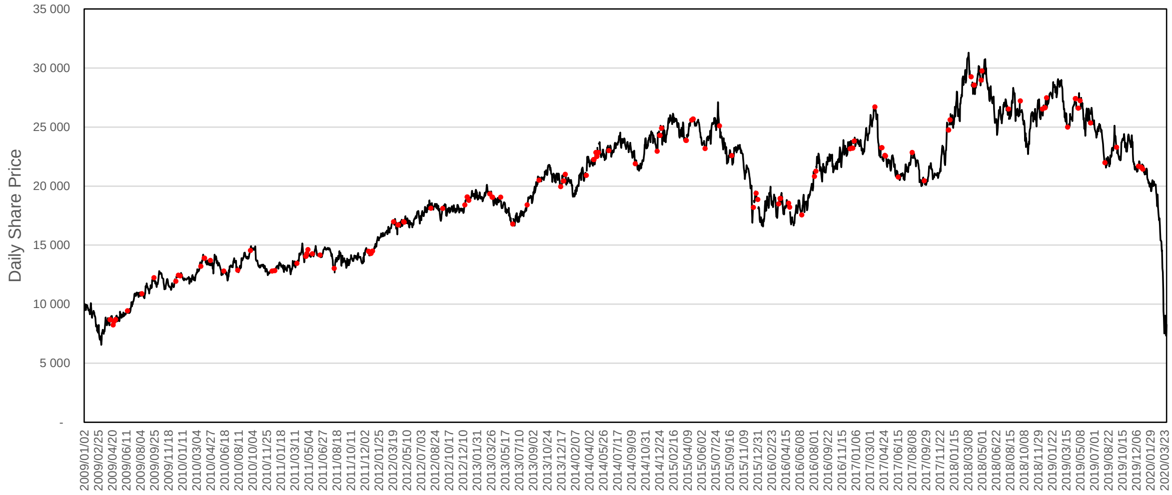


Figure 12: Share Prices of NED (Source: Reuters)

Share Prices of SBK

Total nominal arithmetic return over the sample period 25.73%

Red circles represent missing values where LOCF method has been applied

Figure 13: Share Prices of SBK (Source: Reuters)

Once the daily share price datasets had been updated to include missing data, the researcher then computed the daily log returns of each share for the analysis. Daily log returns were calculated by taking the log of the difference between t's (today) share price and t-1's (yesterday) share price, as set out in the formula below:

$$R_t = log\left(\frac{P_t}{P_{t-1}}\right) \qquad \text{(Equation 8)}$$

The figures in Panel 2 below showed the plots of each bank's daily log returns for each share over the sample period 1 January 2009 to 31 March 2020. It was noted that in Panel 2 below that daily log returns of the bank shares tend to have a period of high and low volatility, volatility clustering described in Chapter 3 above, which is consistent with our understanding of real-world time series data and tend to be grouped around zero.

Reviewing the historical share prices and daily log returns of the South African Major Banks, the share price performance between the shares varied greatly over the sample period. The share prices of ABG and NED decreased over the period, -30.72% and -12.62%, respectively, while CPI, FSR, INL, and SBK all increased 3,112.91%, 195.67%, 17.35%, and 25.73% respectively. Capitec Bank (CPI) and FirstRand Bank outperformed the other banks considerably over the sample period. Although each bank's share price performance over the sample period varied, there were instances of commonality between the banks, specifically where the largest absolute price variances. This was interesting as some of the most significant movements in the bank's share price identified were attributable to events in local and global markets and not the company's business performance.

In December 2015, South Africa's finance minister, Nhlanhla Nene, was unexpectedly removed by President Jacob Zuma and replaced with David Van Rooyen. The removal and replacement were not well received by international investors and credit agencies and resulted in a considerable depreciation of the Rand against major currencies and a 13.5% decrease in the JSE's banking index (Treanor, 2015). Shortly after Nene's removal and David Van Rooyen's appointment, the president appointed Pravin Gordhan as finance minister; Gordhan held the position previously. The appointment of Gordhan partially restored confidence in the South African market, which resulted in a recovery of banking stocks.

Another market event that impacted bank stocks during the sample period was Cyril Ramaphosa's December 2017 victory as the new ANC leader. Ramaphosa's victory provided the international market with some semblance of future stability for South Africa, which resulted in the JSE all-share index increasing by 1.67% and the JSE banking sector gaining 6.63% (Mahlangu, 2017).

Moving to the start of 2020, all shares in the study experienced notable decreases from the start of 2020 to the end of the sample period, 31 March 2020. The decrease in the share prices can be attributed to the overall JSE sell-off driven by the Coronavirus pandemic, which led to one of the fastest market declines in history (Kaplan, 2021).

Panel 2 - Plots of South African Major Banks Nominal Daily Log Returns

Nominal Daily Log Returns of ABG

Average nominal daily return over the sample period -0.030%

Red Circles represent the largest absolute price variations over the sample period

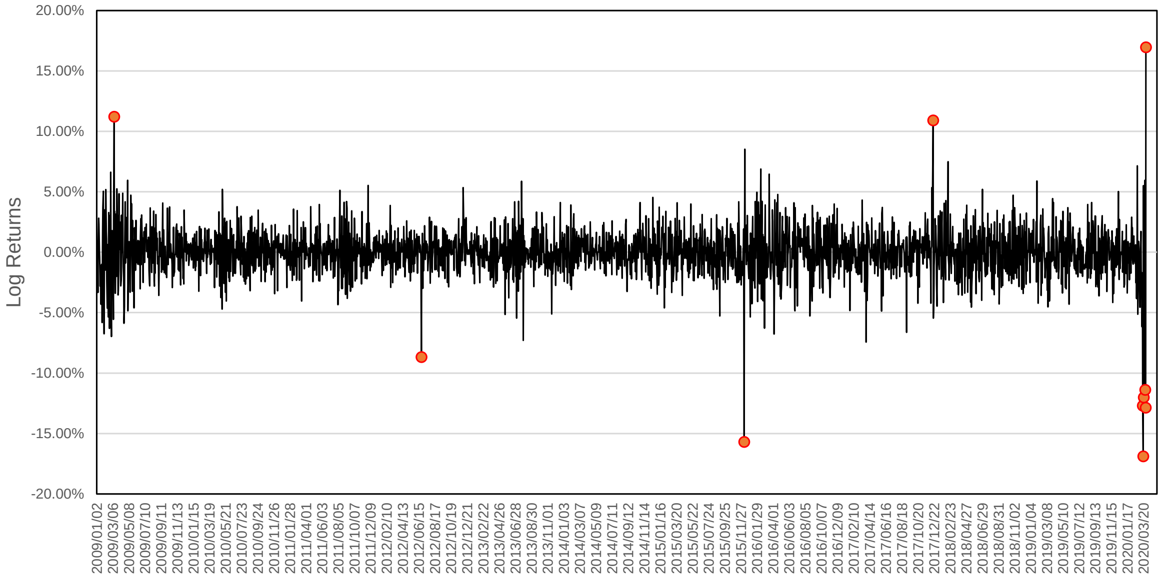
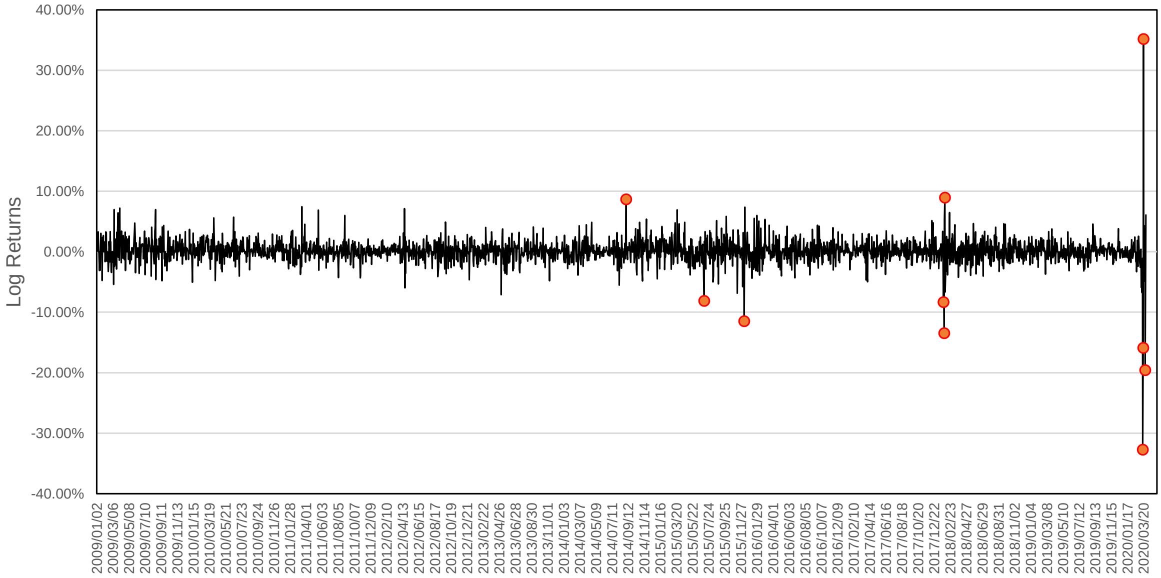


Figure 14: Nominal Daily Log Returns of ABG (Source: Reuters)

Nominal Daily Log Returns of CPI

Average nominal daily return over the sample period 0.099%

Red Circles represent the largest absolute price variations over the sample period



Source: Nominal Daily Log Returns of CPI (Source: Reuters)

Nominal Daily Log Returns of FSR

Average nominal daily return over the sample period 0.025%

Red Circles represent the largest absolute price variations over the sample period

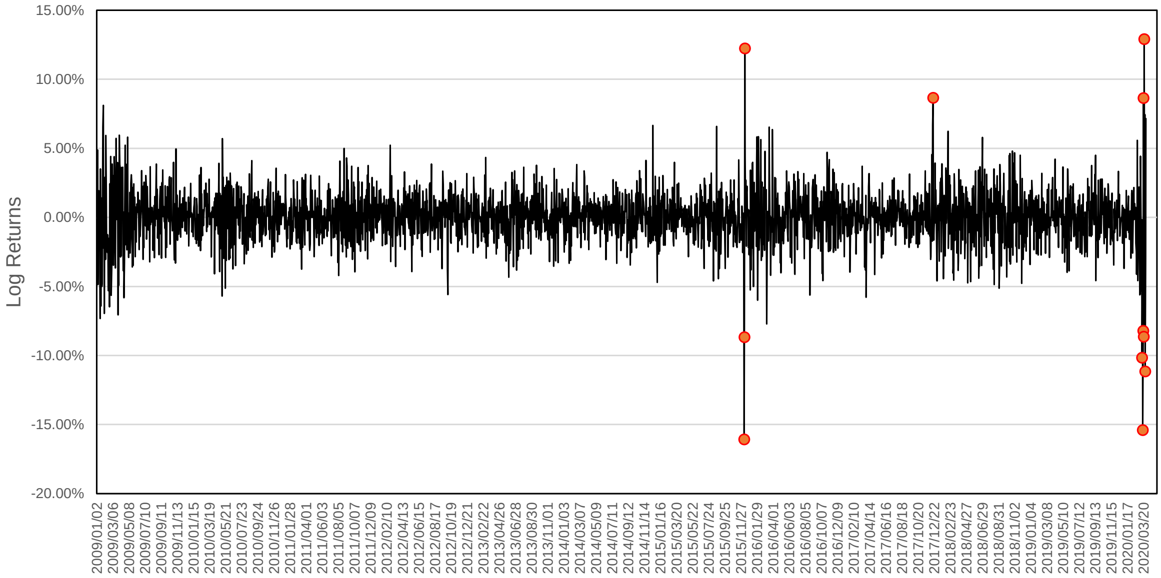


Figure 15: Nominal Daily Log Returns of FSR (Source: Reuters)

Nominal Daily Log Returns of INL

Average nominal daily return over the sample period -0.013%

Red Circles represent the largest absolute price variations over the sample period

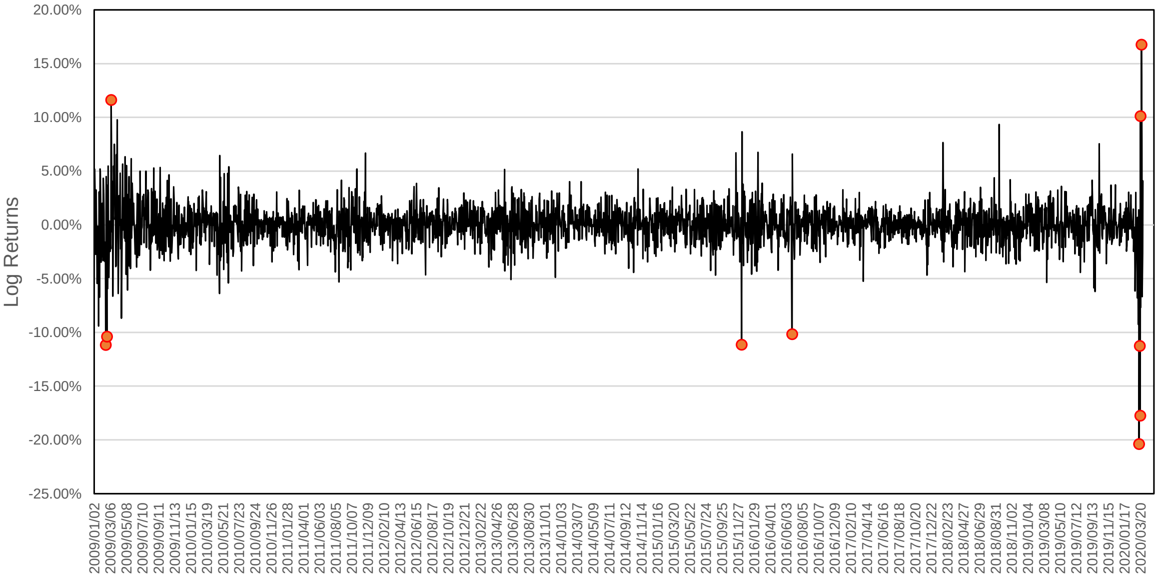


Figure 16: Nominal Daily Log Returns of INL (Source: Reuters)

Nominal Daily Log Returns of NED

Average nominal daily return over the sample period -0.031%

Red Circles represent the largest absolute price variations over the sample period

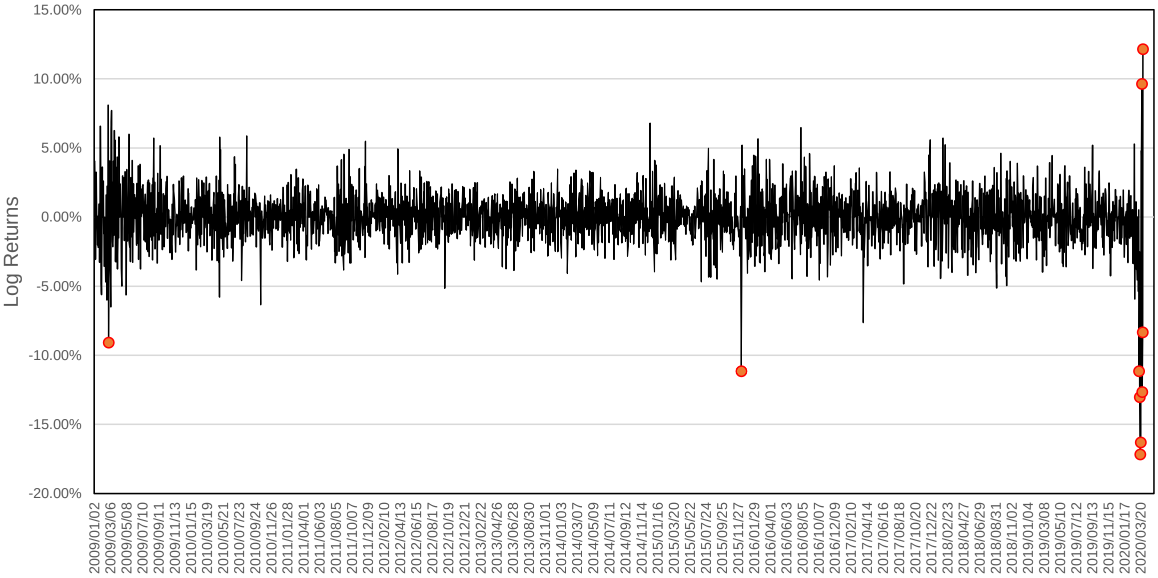


Figure 17: Nominal Daily Log Returns of NED (Source: Reuters)

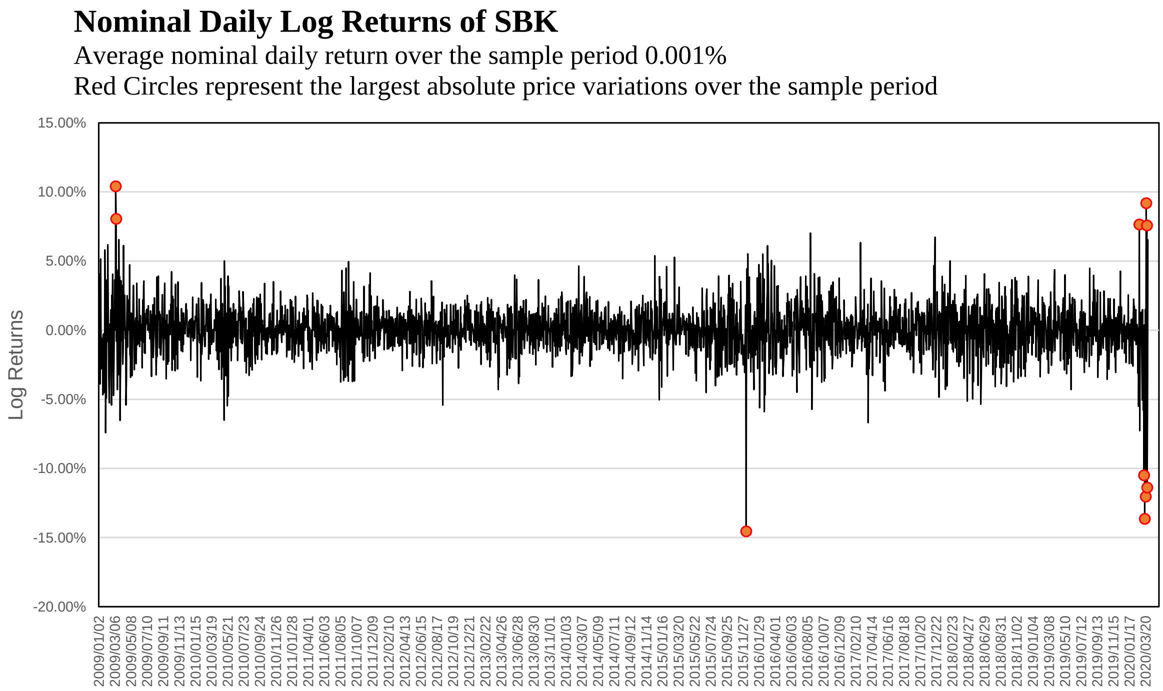


Figure 18: Nominal Daily Log Returns of SBK (Source: Reuters)

4.3 Building volatility models

4.3.1 Testing for ARCH effect / Volatility in the returns' series

As described in Chapter 3, this research used the Lagrange Multiplier test to confirm the presence of ARCH effects in the dataset. This test needed to be carried out before estimating a GARCH model to prevent the misapplication of GARCH models to the data. A Lagrange Multiplier test was used with the null hypothesis being that the dataset has no ARCH effects and then looked for a p-value close to zero, rejecting the null hypothesis. Table 2 presented the results of the ARCH LM test for lags ranging from 1 to 5; per Ramos (2021) running the LM test for up to 5 lags is sufficient "to capture volatility memory in daily returns" for the daily returns of each share included in the scope of this study (ABG, CPI, FSR, INL, NED, SBK). The ARCH LM test was used to determine whether there was an Autoregressive Conditional Heteroskedasticity (ARCH) effect (Engle, 1982), signifying varying volatility in the time series data.

Table 2: Results of ARCH LM test for all bank data

ABG			SBK			CPI		
Lag	LM Statistic	p-value	Lag	LM Statistic	p-value	Lag	LM Statistic	p-value
1	431	< 0.0001	1	142	< 0.0001	1	66	< 0.0001
2	601	< 0.0001	2	271	< 0.0001	2	422	< 0.0001
3	606	< 0.0001	3	400	< 0.0001	3	439	< 0.0001
4	606	< 0.0001	4	450	< 0.0001	4	465	< 0.0001
5	621	< 0.0001	5	457	< 0.0001	5	575	< 0.0001
FSR			INL			NED		
Lag	LM Statistic	p-value	Lag	LM Statistic	p-value	Lag	LM Statistic	p-value
1	299	< 0.0001	1	186	< 0.0001	1	419	< 0.0001
2	629	< 0.0001	2	331	< 0.0001	2	746	< 0.0001
3	640	< 0.0001	3	481	< 0.0001	3	882	< 0.0001
4	658	< 0.0001	4	585	< 0.0001	4	888	< 0.0001
5	659	< 0.0001	5	612	< 0.0001	5	911	< 0.0001

The results in Table 2 reported the following information:

Lag: The "Lag" column indicated the number of lags considered in the ARCH LM test. Each row corresponds to a different lag value, which assessed ARCH effects at different levels of time dependence.

LM Statistic: The "LM Statistic" column showed the test statistic obtained for each lag. This statistic was used to evaluate the presence of ARCH effects. A larger LM statistic suggests more substantial evidence of volatility in the data.

p-value: The "p-value" column provided the probability associated with the LM statistic under the null hypothesis that no ARCH effects exist in the data. A small p-value (typically less than the significance level of 0.05) indicated strong evidence against the null hypothesis. In Table 2, all p-values are reported as "< 0.0001," which means they are very close to zero.

As indicated by the results, for each bank (ABG, CPI, FSR, INL, NED, SBK) and at each lag, the p-values were extremely small (less than 0.0001), indicating that the null hypothesis of no ARCH effect was rejected. These small p-values suggested a significant presence of ARCH effects on all the daily returns time series datasets of the shares included in the study scope. In addition to the results of the null hypothesis test, the LM statistics (the Ljung-Box-McLeod statistic used to test whether a time series model possesses autocorrelation) were relatively high for most lags. A high LM statistic supported the assertion that a model possesses autocorrelation, which helps determine if the selected model is appropriate in capturing all information included in a dataset. This further supported the evidence of ARCH effects in the shares' daily returns time series datasets.

In summary, the results of the ARCH LM test provided strong evidence that the daily returns time series datasets of the shares are likely to possess ARCH effects and, in turn, justify the use of conditional volatility models like the GARCH family of models to estimate volatility.

4.3.2 Determining the Optimal Models for Fitting Major South African Banks' Data

As Ramos (2021) described, to estimate a GARCH model, it is essential to identify an appropriate model that fits the dataset. As a part of this process, the researcher defined the number of lags, the variance equation and distribution parameters to include in the model. For each shares daily returns time series dataset, the most appropriate GARCH model was found by comparing the most used measures of goodness of fit, namely the Akaike Information Criteria (AIC) and Bayesian Information Criterion (BIC). The method of interpreting the values of the AIC and BIC measures was that lower the value, the better the fit.

Moreover, as Ramos (2021) noted, the results leveraged the computing power of modern statistical software. In the present case, using R, instead of manually choosing a GARCH type of model, its lags and distribution assumption, the software analyzed the data and found the optimal model and parameters. Allowing the software to search for parameters automatically avoided the potential biases, swaying the selection away from an optimal model and parameter to suit the preferences.

Table 3 below presented the estimation results of GARCH models for shares included in this study: ABG, SBK, CPI, FSR, INL, and NED. The figures in Appendix 2 provided the graphical output from R for the GARCH model selection for each Share. Each figure indicated which GARCH family of models was most appropriate, and as noted above, where the AIC and BIC values were smallest. These points were noted on each figure by a blue asterisk. Each model identified as the most appropriate for each share was used to analyse the conditional variance (volatility) dynamics of each share's daily returns. The interpretation of the results was noted below:

Table 3: Estimation results of GARCH models that fit each bank data

Parameter	ABG	CPI	FSR	INL	NED	SBK
Mu	0.000152	0.001012*	0.000650*	0.000449*	0.000253*	0.000286
ar1	0.574620*	N/A	0.911651*	N/A	0.633258*	0.715091*
ma1	-0.655215*	N/A	-0.945991*	-0.056756*	-0.703133*	-0.781915*
Omega	-0.154396*	-0.203696*	0.000006	-0.187156*	0.000007*	0.000004
alpha1	-0.061643*	-0.002007	0.020095*	-0.060498*	0.039880*	0.022815
alpha2	N/A	-0.053189	N/A	N/A	N/A	N/A
beta1	0.980903*	0.974636*	0.928985*	0.977211*	0.912116*	0.935695*
gamma1	0.134477*	0.367753*	0.068858*	0.153920*	0.053287*	0.059963*
gamma2	N/A	-0.201996*	N/A	N/A	N/A	N/A
Shape	6.514796*	3.511752*	8.409472*	6.058313*	7.775713*	7.834395*
Variance model	eGARCH(1,1)	eGARCH(2,1)	gjrGARCH(1,1)	eGARCH(1,1)	gjrGARCH(1,1)	gjrGARCH(1,1)

Notes: "*" – Significant at 5% level (p-value < 0.05), N/A: Not Applicable indicates that the specific parameter is not present in that model. This is due to the data characteristics, which do not require that term for a good fit.

Mu (Mean): Mu represented the estimated mean return associated with each share's daily returns. Significant values at the 5% level (indicated by *) suggested a statistically significant mean return for CPI, FSR, INL, and NED banks. In these cases, the mean return was different from zero, indicating that these banks have positive mean returns. For ABG and SBK banks, mu values were also statistically significant at the 5% level but were relatively smaller. This implied that while these banks had statistically significant mean returns, those returns were closer to zero.

Ar1 (Autoregressive Coefficient): ar1 represented the autoregressive coefficient in the mean equation. It indicated the persistence of returns (past returns or price movements tend to have a significant influence on future returns). Significant and positive values (indicated by *) for ABG, SBK, CPI, FSR, and NED suggested that returns exhibit positive autocorrelation, meaning that past returns have a positive impact on current returns. In these cases, returns tend to continue in the same direction. For INL bank, the "ar1" value was unavailable (N/A), suggesting that the autoregressive behaviour was not modelled for this share. The absence of an AR1 term or "N/A" in the context of a time series model suggested that based on the analysis and data characteristics, including an autoregressive relationship at lag one, it was not relevant. The parameter excluded was made to ensure that the model appropriately captured the underlying patterns and dependencies in the underlying data.

Ma1 (Moving Average Coefficient): ma1 represented the moving average coefficient in the mean equation, which also affected the persistence of returns. Significant and negative values (indicated by *) for ABG, SBK, CPI, FSR, and NED shares suggested that past shocks (good or bad news) to returns have a negative impact on current returns. In these cases, returns tend to revert to their mean value after past shocks. For INL bank, the "ma1" value was unavailable (N/A), suggesting that the moving average behaviour was not modelled for this bank. The presence of MA1 in a time series model signified that the model considered the immediate impact of past errors or shocks on the current observations. It can be a useful parameter when short-term dependencies were present in the data. However, its inclusion should be justified based on the data's characteristics and the objectives of the analysis.

Omega (Constant in Conditional Variance): Omega represented the constant term in the conditional variance equation, influencing the baseline level of conditional variance (volatility). Significant and negative values (indicated by *) for ABG, SBK, CPI, FSR, and NED shares suggested that the conditional variance is lower for these banks, indicating lower baseline volatility. For INL bank, the "omega" value was positive and statistically significant, indicating that this bank has a higher baseline volatility.

Alpha1 and Alpha2 (ARCH Parameters): alpha1 and alpha2 represented the coefficients of the ARCH terms in the conditional variance equation, measuring the persistence of volatility aftershocks to the market (good or bad news). alpha 1 was typically associated with short-term persistence, while alpha 2 captured longer-term persistence. The complexity and significance of these parameters depended on the data and model. Significant Alphas suggested that past volatility has a substantial impact on current volatility. alpha1 was significant for ABG, SBK, CPI, FSR, and NED shares, indicating that past volatility shocks have a persistent impact on current volatility. These shares exhibited volatility clustering, where high volatility periods tend to follow high volatility periods. alpha2 was only available for CPI banks and was significant, suggesting additional complexity in modelling volatility dynamics for CPI.

Beta1 and Gamma1 (GARCH Parameters): beta1 and gamma1 represented the coefficients of the GARCH terms in the conditional variance equation, capturing the impact of past squared returns and past squared shocks on current volatility, respectively. The parameters account for the differential responses of volatility. beta1 was significant for all shares, indicating that past squared returns have a significant impact on current volatility. gamma1 was significant for all shares, suggesting that past squared shocks have a significant impact on current volatility.

Variance Model: Table 3 specified the specific GARCH models used for each share, such as eGARCH(1,1), eGARCH(2,1) and gjrGARCH(1,1), indicating the lag order in the model. The volatility over the sample period of each share was included in the figures below. Actual Volatility patterns of each share daily returns data were shown in Panel 3 below.

Panel 3- Conditional Variance vs Absolute Daily Log Returns

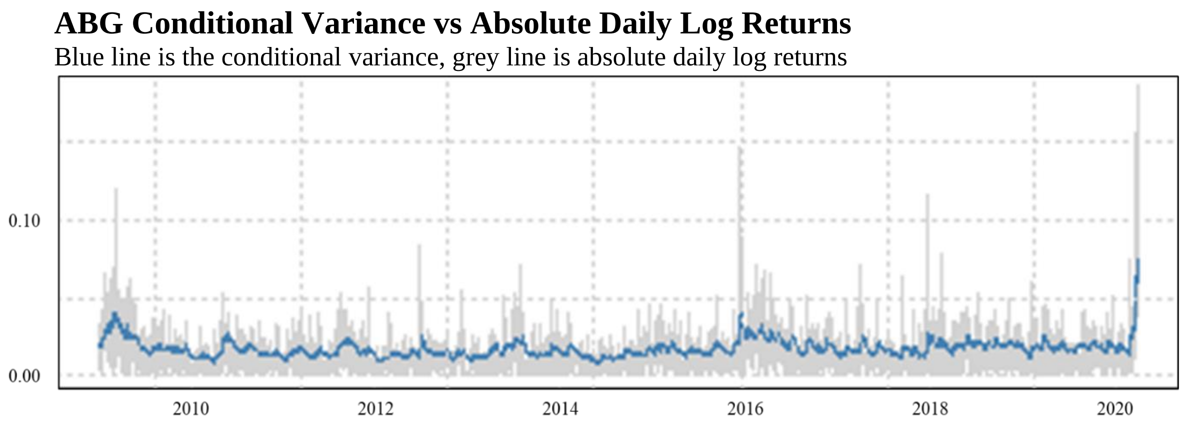


Figure 19: ABG Conditional Variance vs Absolute Daily Log Returns (Source: Reuters)

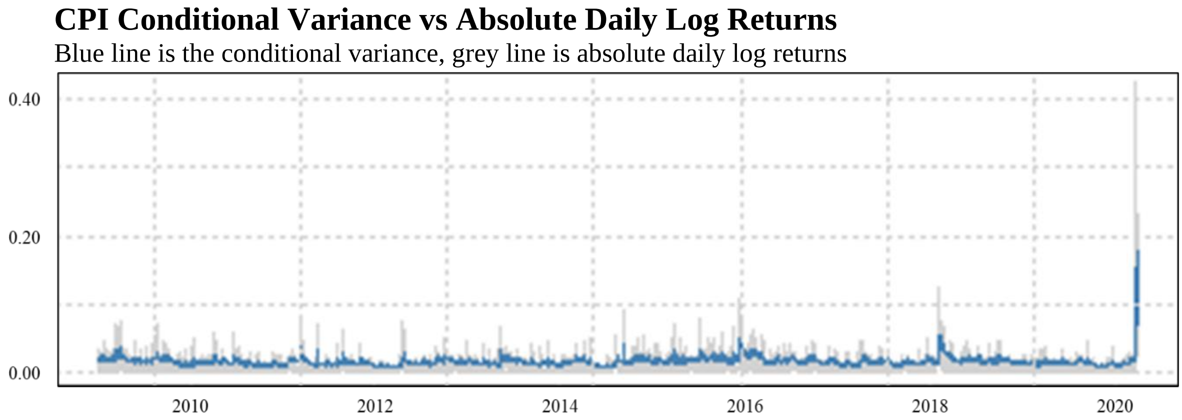


Figure 20: CPI Conditional Variance vs Absolute Daily Log Returns (Source: Reuters)

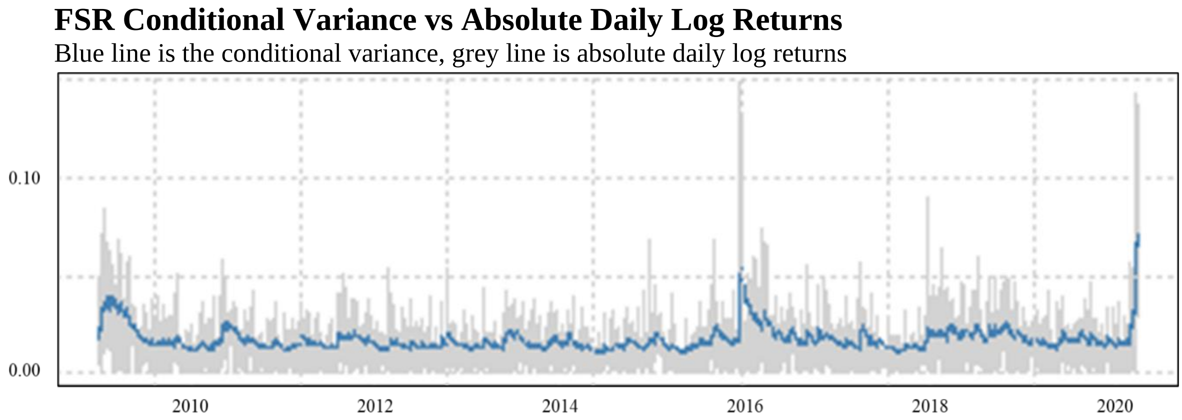


Figure 21: FSR Conditional Variance vs Absolute Daily Log Returns (Source: Reuters)

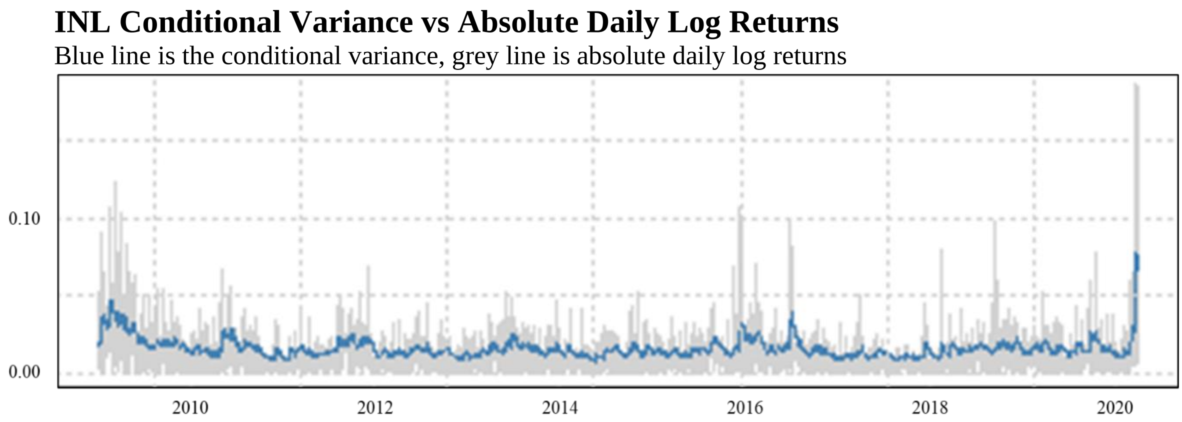


Figure 22: INL Conditional Variance vs Absolute Daily Log Returns (Source: Reuters)

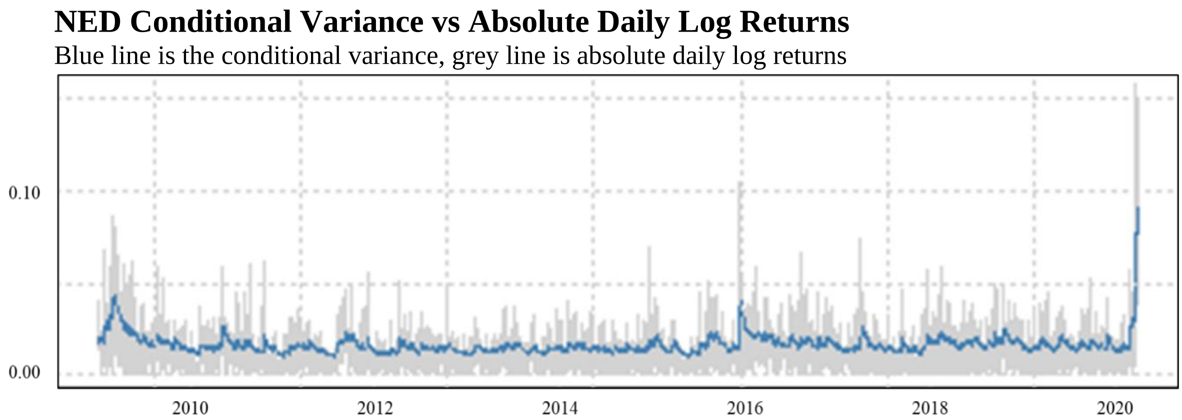


Figure 23: NED Conditional Variance vs Absolute Daily Log Returns (Source: Reuters)

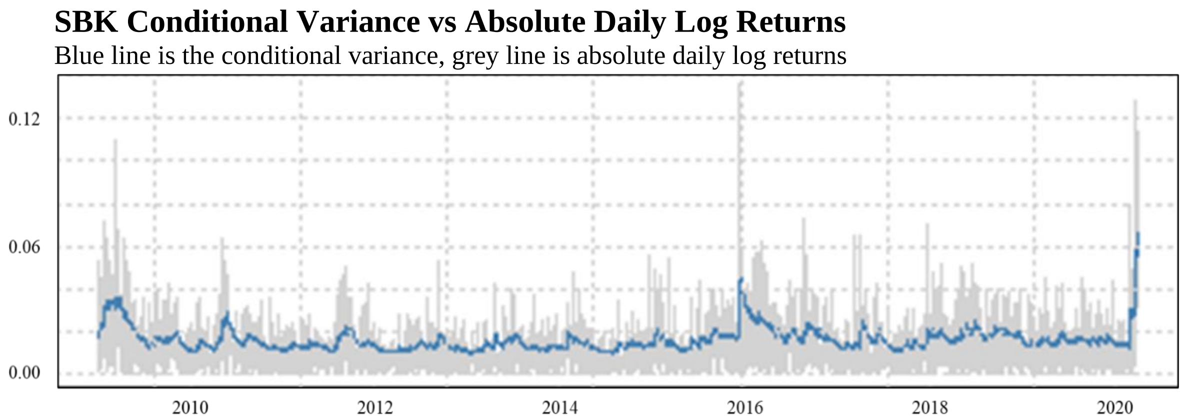


Figure 24: SBK Conditional Variance vs Absolute Daily Log Returns (Source: Reuters)

In summary, assessing the daily share returns of South African banks' time series data revealed varying natures of volatility models, ranging from different GARCH models to asymmetry in capturing volatility dynamics. Common themes emerged, including ARCH and GARCH effects, signifying volatility clustering and persistence. Additionally, the consistent observation of asymmetric responses to positive and negative shocks shed light on how these shares were impacted by the ever-changing landscape of financial market volatility.

These findings provided interesting insights into the behaviour of stock returns and volatility for each share, assisting in risk assessment and investment decision-making.

The research expands the analysis by introducing external regressors, namely the level of ETF ownership and the occurrence of ETF-Specific events, into the GARCH models described in Equations 4, 5, and 6 to examine the relationship between volatility and ETF ownership percentage, as well as the occurrence of ETF events. Subsequently, the parameters of these updated GARCH models are estimated using the available data for each share. An overview of the results from this analysis are laid out in the following sections.

4.4 Investigate the association between ETF-specific events and bank stock volatility

Having identified the suitable GARCH model for each share's daily returns time series dataset, the analysis was furthered by integrating the ETF Specific Event variable as an external regressor for each share. The analysis was extended by incorporating the ETF-Specific Event variable as an external regressor for each shares after identifying the appropriate GARCH models for each share’s daily returns time series dataset. Incorporating the external regressor into each of the share’s GARCH models helped to determine if ETF-Specific Events (the presence of an ETF-Specific Event occurring or not occurring) as a variable explains the share’s daily return volatility. Table 4 below summarized the results of the extended analysis and the impact of the introduction of external regressors on the model parameters. The interpretation of these results was noted below.

Mu (Mean): The mean "mu" parameter, representing the estimated mean return associated with each bank's stock returns during ETF events, demonstrated that an ETF-specific event has a statistically significant impact on the mean return for ABG and CPI banks, as indicated by the * symbol. The mean return significantly deviated from the norm for these shares during ETF-Specific Events. However, for SBK, FSR, INL, and NED banks, the mu values were not statistically significant, suggesting that ETF-Specific Events did not significantly explain the mean returns for these shares.

Ar1 (Autoregressive Coefficient): Regarding the autoregressive coefficient "ar1" coefficient, which measures the persistence of returns, negative and significant values for ABG, SBK, FSR and NED shares implied that returns exhibit mean-reverting behaviour after ETF-Specific

Events. In other words, when an ETF-Specific Event occurs, the shares; daily returns volatility moves closer to its mean. This was opposite to the results of the GARCH model before the inclusion of ETF-Specific Events as an external regressor. In contrast, CPI and INL shares did not display mean-reverting behaviour during these events, as their "ar1" values were unavailable (N/A). This was consistent with CPI and INL's GARCH model results before the inclusion of the ETF-Specific Event external regressor.

Mxreg1 (ETF Percent Ownership Coefficient): The "mxreg1" coefficient, representing the impact of ETF-Specific Events on conditional variance (volatility) of the shares daily returns, was not statistically significant for any of the shares, indicating that the occurrence or non-occurrence of ETF-Specific events did not significantly influence conditional variance (volatility).

Omega (Constant in Conditional Variance): The "omega" parameter, which affected the baseline level of conditional variance (volatility), represents the long-term average of the conditional variance. A positive Omega indicated that the model expects volatility to increase over time, while a negative value suggested a decrease. In the model Omega revealed that conditional variance was significantly lower for ABG, SBK, CPI, FSR, and NED shares during ETF-Specific Events. However, for INL bank, the "omega" value was positive and statistically significant, suggesting that the long-term average of the conditional variance was higher during ETF-Specific Events for this share.

Alpha1 and Alpha2 (ARCH Parameters): Examining the "alpha1" and "alpha2" parameters, which related to the persistence of volatility aftershocks (bad or good news), it was found that "alpha1" is significant for ABG, SBK, CPI, FSR, and NED banks, indicating that past volatility shocks persistently affect current volatility during ETF events. "Alpha2" was only available for the CPI bank and was not statistically significant, implying no significant persistence in volatility for this bank.

Beta1 and Gamma1 (GARCH Parameters): The "beta1" and "gamma1" coefficients, capturing the impact of past squared returns and past squared shocks on current volatility, were both significant for all shares, suggesting that past squared returns and shocks significantly affect current volatility of the share’s daily returns during ETF-Specific events.

Gamma2 (Additional GARCH Parameter): Furthermore, the "gamma2" parameter, an additional GARCH parameter, was significant only for the CPI share, indicating an extra level of complexity in modelling volatility dynamics for this particular share.

Shape (Distribution Parameter): The "shape" parameter, related to the distribution assumption of returns, showed that returns significantly deviated from a standard distribution during ETF-Specific Event for all shares, as indicated by the * symbol. This implied that returns exhibit non-normal behaviour during the ETF-Specific Event.

Variance Model: Lastly, the Table 4 specified the specific GARCH models used for each bank, such as eGARCH (1,1), eGARCH(2,1) and gjrGARCH(1,1), indicating the lag order in the model.

In summary, the results indicated that the occurrence or non-occurrence of ETF-Specific Events have varying impacts on different aspects of returns and volatility for different shares. Notably, ETF-Specific Events affect the mean return for ABG and CPI banks and have varying effects on conditional variance and distribution assumptions across all shares. However, the impact on conditional variance or volatility (mxreg1) was not statistically significant for any shares. The persistence of returns and volatility also varied between shares, as reflected in the autoregressive and ARCH/GARCH parameters.

Table 4: GARCH models with ETF Specific event as external regression

Parameter	ABG	CPI	FSR	INL	NED	SBK
Mu	0.000069*	0.001008*	0.000557	0.000499	0.000294	0.000121
ar1	-0.074803*	N/A	-0.043454*	N/A	-0.062349*	-0.061390*
mxreg1	-0.000046	0.000031	-0.000564	-0.000645	-0.000534	0.000261
Omega	-0.161499*	-0.203786*	0.000005	-0.202622*	0.000007*	-0.113698*
alpha1	-0.066362*	-0.001939	0.019371	-0.066712*	0.039984*	-0.057358*

alpha2	N/A	-0.053265	N/A	N/A	N/A	N/A
beta1	0.979983*	0.974625*	0.928286*	0.975291*	0.912237*	0.986069*
gamma1	0.138708*	0.367790*	0.075728*	0.158794*	0.054231*	0.121725*
gamma2	N/A	-0.202055*	N/A	N/A	N/A	N/A
Shape	6.486464*	3.511634*	8.471880*	6.066167*	7.745925*	7.721416*
Variance model	eGARCH(1,1	eGARCH(2,1)	gjrGARCH(1,1)	eGARCH(1,1)	gjrGARCH(1,1)	gjrGARCH(1,1)

*: Significant at 5% level.

4.5 Analysing the Association Between ETF Ownership Percentage and Constituent

Similar to the extended analysis relating to the ETF-Specific Events and the volatility of the share's daily returns, the analysis was extended by incorporating the level of ETF Ownership variable as an external regressor for each of the shares and the impact on the introduction of this variable on the models' parameters. Incorporating the external regressor into each of the share's GARCH model enabled to determine if the level of ETF ownership (the degree to which ETFs hold shares in constituent shares) as a variable explained the shares' daily return volatility. Table 5 below summarized the results of the extended analysis. The interpretation of these results was noted below.

Mu (Mean): Mu reflected the estimated mean return associated with each share's daily returns in response to changes in the degree of ETF ownership. Statistically significant values at the 5% level (indicated by *) signified a substantial mean return impact linked to fluctuations in ETF percent ownership for ABG, SBK, and CPI shares. This implied that shifts in ETF ownership percentage notably influenced the mean return. Conversely, for FSR, INL, and NED banks, mu values lacked statistical significance at the 5% level, indicating that changes in ETF percent ownership did not significantly affect mean returns for these shares.

Ar1 (Autoregressive Coefficient): ar1 represented the autoregressive coefficient in the mean equation, indicating the persistence of returns. Negative and significant values (indicated by *) for ABG, SBK, and NED shares suggested that returns exhibited mean-reverting behaviour following changes in ETF ownership. This meant that returns tend to revert to mean values after such changes. However, for FSR, CPI, and INL banks, Ar1 values were unavailable (N/A), suggesting that mean-reverting behaviour was not modelled for these shares data with the inclusion of the external regressor (Mxreg1).

Mxreg1 (ETF Percent Ownership Coefficient): mxreg1 referred to the coefficient associated with variations in the degree of ETF Ownership in the conditional variance (volatility) equation. It quantified the association of changes in ETF ownership on the conditional variance (volatility) of returns. Statistically significant negative values (indicated by *) for ABG and SBK shares suggested that increases in ETF ownership led to decreased conditional variance (volatility) for these shares. However, the impact was statistically insignificant for CPI, FSR, INL, and NED shares.

Omega (Constant in Conditional Variance): Omega represented the constant term in the conditional variance equation, influencing the baseline level of the long-term average of the conditional variance. Significant and negative values (indicated by *) for ABG, SBK, CPI, FSR, and NED shares implied that the conditional variance was lower during periods of changes in ETF percent ownership. However, for INL share, the positive and statistically significant "omega" value suggested that the conditional variance is higher during such periods.

Alpha1 and Alpha2 (ARCH Parameters): alpha1 and alpha2 denoted the coefficients of the ARCH terms in the conditional variance equation, gauging the persistence of volatility aftershocks (good or bad news). alpha1 was significant for ABG, SBK, CPI, FSR, and NED shares, signifying that past volatility shocks persistently affect current volatility during ETF percent ownership changes. alpha2 was unavailable (N/A) for most banks except for CPI, which was not statistically significant, indicating no significant persistence in volatility during such events.

Beta1 and Gamma1 (GARCH Parameters): beta1 and gamma1 coefficients captured the impact of past squared returns and past squared shocks on current volatility, as they represented the GARCH terms in the conditional variance equation. Statistically significant values for "beta1" and "gamma1" for all shares suggested that past squared returns and past squared shocks significantly affect current volatility during periods of changes in ETF ownership.

Gamma2 (Additional GARCH Parameter): gamma2 represented an additional GARCH parameter and was solely available for the CPI bank. Its statistical significance suggested an added level of complexity in modelling volatility dynamics during such events for this share.

Shape (Distribution Parameter): The "shape" parameter was related to the distribution assumption of returns (e.g., Student's t-distribution). Statistically significant values (indicated by *) for all shares indicate that the distribution of returns during periods of ETF ownership changes significantly deviated from a standard distribution, implying non-normal behaviour during these events.

Variance Model: The Table 5 specified the particular GARCH models used for each share, such as eGARCH(1,1) or gjrGARCH(1,1), indicating the lag order in the model.

Table 5: GARCH Models with ETF percent ownership as an external regressor

Parameter	ABG	CPI	FSR	INL	NED	SBK
Mu	0.000929*	0.001093*	0.00036	0.000401	0.0005	0.000687*
ar1	-0.073433*	N/A	-0.043331*	N/A	-0.062523*	-0.061528*
mxreg1	-0.474835*	-0.294643	0.044976	-0.010239	-0.180658	-0.199223*
Omega	-0.162970*	-0.203389*	0.000005	-0.203806*	0.000007*	-0.115812*
alpha1	-0.069201*	-0.001952	0.019013	-0.066022*	0.038308*	-0.058950*
alpha2	N/A	-0.053099	N/A	N/A	N/A	N/A
beta1	0.979818*	0.974670*	0.928357*	0.975155*	0.912610*	0.985818*
gamma1	0.135591*	0.367923*	0.075853*	0.158942*	0.056398*	0.120249*
gamma2	N/A	-0.201799*	N/A	N/A	N/A	N/A
Shape	6.505348*	3.511656*	8.482673*	6.100070*	7.774638*	7.753086*
Variance model	eGARCH(1,1)	eGARCH(2,1)	gjrGARCH(1,1)	eGARCH(1,1)	gjrGARCH(1,1)	gjrGARCH(1,1)

*: Significant at 5% level.

In summary, the results demonstrated that changes in ETF percent ownership exert varying impacts on different aspects of returns and volatility for different banks. Notably, fluctuations in ETF ownership percentage affected the mean return for ABG, SBK, and CPI banks and had diverse effects on conditional variance (volatility) and distribution assumptions across all banks. Nevertheless, the association between the level of ETF Ownership and the share's volatility was negative and statistically significant for ABSA and SBK while insignificant for the other banks.

4.6 Conclusion

Following the methodology of Ramos (2021) the study analysed the share's daily return time series data to identify the most appropriate GARCH model and parameters for each share to estimate the volatility of daily returns. Once the baseline GARCH model and parameters were estimated, the research extended the analysis to include external regressors, the occurrence and non-occurrence of ETF-Specific Events, and the degree of ETF Ownership of each share in the GARCH Models. The purpose of the inclusion of these external regressors was to determine if there was a relationship between the volatility of daily returns and each external regressor. The results of these analyses were mixed in that for some shares, there was an association between the external regressor and the volatility of the daily returns and, in some, there was not. There was evidence of a statistically significant negative association between the degree of ETF Ownership and the volatility of daily returns of ABG and SBK. In contrast, CPI, FSR, INL, and NED, the degree of ETF Ownership was not statistically significant in explaining the volatility of daily returns. On the other hand, it was also noted that there was no statistically significant association between ETF Specific Events and the volatility of daily returns for any of the shares. The results of the analysis were inconsistent with the results of the studies noted in the literature review.

With regards to the results of the analysis of ETF Specific Events, although this result was not in line with the results of Ben David 2018, where the increase in ETF ownership and the related trading activity resulted in an increase in underlying stock volatility, the result of no statistical significance in the analysis was no surprise due to the limitations and quality of data. As noted in Chapter 3, the quality of the variable as a proxy for ETF Specific Events (ETF Creations, Redemptions, and Rebalancing) was most likely too weak as it was based on SENS announcements made by the ETF sponsor around the period of the underlying trade. Ben-David et al. (2018) study utilized intraday trading, where he identified ETF-Specific Trades that impacted intraday volatility. The analysis of the volatility of daily returns did not allow to see the potential impact of ETF-specific trades.

In correspondence with the results of ETF-Specific Event, the results from the analysis of ETF Ownership and the volatility of daily returns were in contrary to the results of Ben-David et al. (2018), where in his study, he noted that there was a correlation between the degree of ETF Ownership and the volatility of the returns of the constituent shares. He found that due to the unique dynamics of ETFs, where there was an increase in the degree of ownership by ETFs of a share, the volatility of the share would increase. The increase in volatility, in his study, was attributable to the non-fundamental trading and flows in the constituent shares, which resulted in increased volatility. The analysis results were inconclusive for CPI, FSR, INL, and NED, with no statistically significant association. For ABG and SBK, a decrease in baseline volatility was noted with increased ETF ownership in shares. The most likely factor that could potentially reconcile the difference in results was the degree of ETF Ownership in the US equity market is significantly higher (7.05% in 2015 (Ben-David et al., 2018), this is closer to 21% as of 2022 (Gordon, 2022) than the levels noted in the study, leading to other factors not included in this study, that were better explanatory variables for the volatility of daily returns.

5 CONCLUSION

ETFs have seen a tremendous increase in popularity since their inception, with the number of ETFs and assets under management ("AUM") growing significantly over the last few years. ETF financial products have become more affordable and accessible to retail investors, which has seen the global AUM reach 6 trillion dollars by 2019. Due to their "in-kind" nature, the increase in their popularity has also attracted concerns raised by financial market participants. Market participants, as well as financial regulators, have raised concerns regarding ETFs and the potential impact they may have on the stability of financial markets. Various regulatory bodies have raised various concerns about ETFs and their mechanics and the potential contagion and spillover risk.

The current study set out to determine if there was an association between the volatility of daily returns of constituent shares (the research was limited to the six major South African banks, ABSA Bank Limited (ABG), Capitec Bank Holdings Ltd (CPI), FirstRand Ltd (FSR), Investec Ltd (INL), Nedbank Group Ltd (NED), and Standard Bank Group Ltd (SBK)) and on the one hand, the level of ETF ownership of a constituent stock, and on the other ETF Specific events.

The research methodology followed the approach laid out by Ramos (2021) where the researcher used statistical software to estimate which GARCH model would be best-suited for each of the shares to estimate the volatility of daily returns. Once the approach GARCH models were identified, then incorporated ETF-Specific Event and ETF Ownership variables as external regressors to assess whether our variables were statistically significant in explaining the volatility of daily returns of the shares and help understand the potential association between variables and the volatility of the daily returns.

The analysis results were mixed. In some instances, there was an association between the external regressor and the volatility of the daily returns and; in others, there was not. With ETF ownership, there was evidence of a statistically significant negative association between the degree of ETF Ownership and the volatility of daily returns of ABG and SBK, while for CPI, FSR, INL, and NED, the degree of ETF Ownership was statistically insignificant. For ETF-Specific Events, it was noted that there was no statistically significant association with the volatility of daily returns for any of the shares. The findings needed to be consistent with the results of the studies noted in the literature review.

The inconsistency with other research was likely attributable to using a poor proxy for ETF-Specific Events (trading activity), using daily pricing data rather than intraday trading data, and the overall lower degree of ETF ownership in the South African market.

Although the results of this study were inconsistent with other studies in more developed markets, it may provide some utility to market participants looking to gain a better understanding of the mechanics of ETFs and their impact on the South African equity market.

6 AREAS FOR FURTHER RESEARCH

Several areas of further study that can be explored and extend the scope of analysis of this study which would be of potential interest and address some of the limitations of this study.

The scope of this study is limited to the potential impact of South African ETFs on the volatility of daily returns of the Major South African banks. As a first potential extension of this study, the same data and methodology can be applied to other sectors, an example of which could be an analysis that replicates the work of Boney-Dutra et al. (2013), where the analysis is limited to the potential impact on REITs, and shares trading on the JSE. Another alternative could be a broader analysis of all shares included on the JSE. Secondly, several foreign ETFs that have exposure to the South African equity market, the size of which was much greater than ETFs in South Africa, which would be subject to the same dynamics as the South African ETFs. An extension of this study would be to include the holdings and trading of these foreign ETFs to determine if the changes in the degree of ownership and occurrence of ETF-specific events have an impact on the results of this study's analysis.

As the purpose of this study is to determine if there was an association between the degree of ETF ownership, ETF-specific events and the daily return volatility of underlying shares included in the ETFs, the methodology utilized in this study limited external regressors to the degree of ETF ownership and ETF-specific events variables. An opportunity to refine and extend the work of this study could be the introductions of other regressors that could explain the volatility of daily returns, such as the flows of unit trusts, which account for a material proportion of equity holdings.

As noted in Chapter 3, there were several limitations of the source data used in this study. A notable shortcoming of source data identified related to ETF-specific event data where the study was limited to using information obtained from SENS announcements, where data across authorised participants and event types were inconsistent from a quantitative data perspective and only provided the effective trade date, rather than actual underlying trading data, which was used in the Ben-David et al. (2018) study, as well as daily closing prices being the shortest cadence of share price data available.

As a result of these limitations, this study is limited to estimating the volatility of daily returns rather than the volatility of intraday returns of the bank’s shares. By analysing the daily volatility rather than intraday volatility, a full day’s worth of trading was included in the daily return data, which, in the context of the small size of the ETF market in South Africa, potentially hides the potential impact of trading related to ETF-Specific Events. A potential extension of this study would be to obtain intraday share price data as well as the actual ETF trade data to focus the analysis on the volatility of the specific ETF trades.

As noted above, due to the inconsistency of the SENS announcement data, the ETF-Specific Event variable is an indicator variable, 0 where no event occurs and one where an event occurs. By using an indicator variable, it oversimplified the ETF-Specific Event data by treating all events as equal; an ETF creation event with a value of R1,000,000 is the same as an ETF creation event with a value of R1,000,000,000, where the magnitude of a trade would vary the potential non-fundamental pricing pressure of the share being traded. By including the magnitude of the ETF-specific event, one could refine the estimation of the ETF-Specific Event as an external regressor, enhancing the explanatory power of ETF-specific events of daily return volatility.

As noted in 4.2 the inconsistent methodology used for treatment of public holidays as trading days, is not consistent with the standard methodology used in other studies. The re-running of the tests performed in this paper provides an opportunity for further research to determine if the findings differ from this dissertation.

The sample period of this study is limited to 31 March 2020; by extending the analysis by updating the sample period to include current market data, one could determine if the results of this study's analysis were consistent with a more extended sample period or a more consistent result among the shares were found.

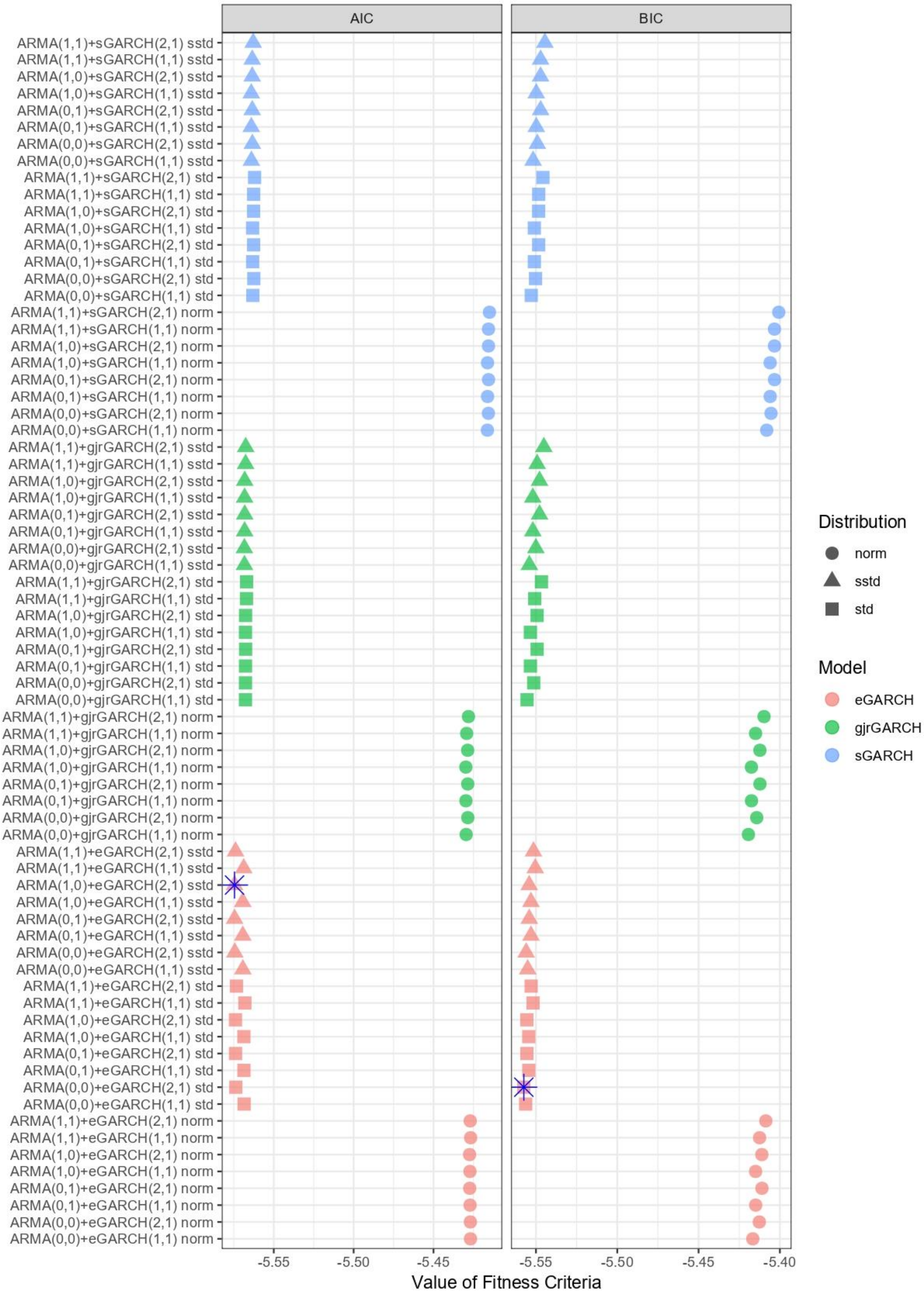
Appendix 1 - Applicable South African Equity ETFs

Fund Name	JSE Ticker	Inception Date
Ashburton Top 40	ASHT40	16 October 2008
1nvest SA Property	ETFSAP	13 February 2013
1nvest SWIX 40	ETFSWX	18 October 2010
1nvest Top 40	ETFT40	18 October 2010
ABSA NewFunds Equity Momentum	NFEMOM	26 January 2012
ABSA NewFunds Low Volatility Equity	NFEVOL	26 March 2018
ABSA NewFunds MAPPS Growth	MAPPSG	25 May 2011
ABSA NewFunds MAPPS Protect	MAPPSP	25 May 2011
ABSA NewFunds S&P GIVI SA Top 50	GIVISA	23 June 2008
ABSA NewFunds Shariah Top 40	NFSH40	06 April 2009
ABSA NewFunds Value Equity	NFEVAL	26 March 2018
ABSA NewFunds Volatility Managed High Growth Equity	NFEHGE	25 February 2019
ABSA NewFunds Volatility Managed Moderate Equity	NFEMOD	25 February 2019
ABSA NewFunds Volatility Managed Defensive Equity	NFEDEF	25 February 2019
Ashburton MidCap	ASHMID	15 August 2012
CoreShares DivTrax	DIVTRX	14 April 2014
CoreShares PrefTrax	PREFTRAX	28 March 2012
CoreShares SA Property Income	CSPROP	30 October 2019
CoreShares SciBeta M-FI	SMART	10 July 2019
CoreShares Top 50	CTOP50	13 May 2015
Satrix Top 40	STX40	27 November 2000
Satrix DIVI	STXDIV	29 August 2007
Satrix FINI	STXFIN	08 February 2002
Satrix INDI	STXIND	08 February 2002
Satrix Momentum	STXMMT	17 November 2018
Satrix Property	STXPRO	24 February 2017
Satrix Quality SA	STXQUA	26 September 2017
Satrix RAFI	STXRAF	16 October 2008
Satrix RESI	STXRES	10 April 2006
Satrix SWIX Top 40	STXSWX	10 April 2006
Sygnia Itrix SWIX Top 40	SYGSW4	30 October 2017
Sygnia Itrix Top 40	SYGT40	30 October 2017

Appendix 2 - GARCH Model Selection Results

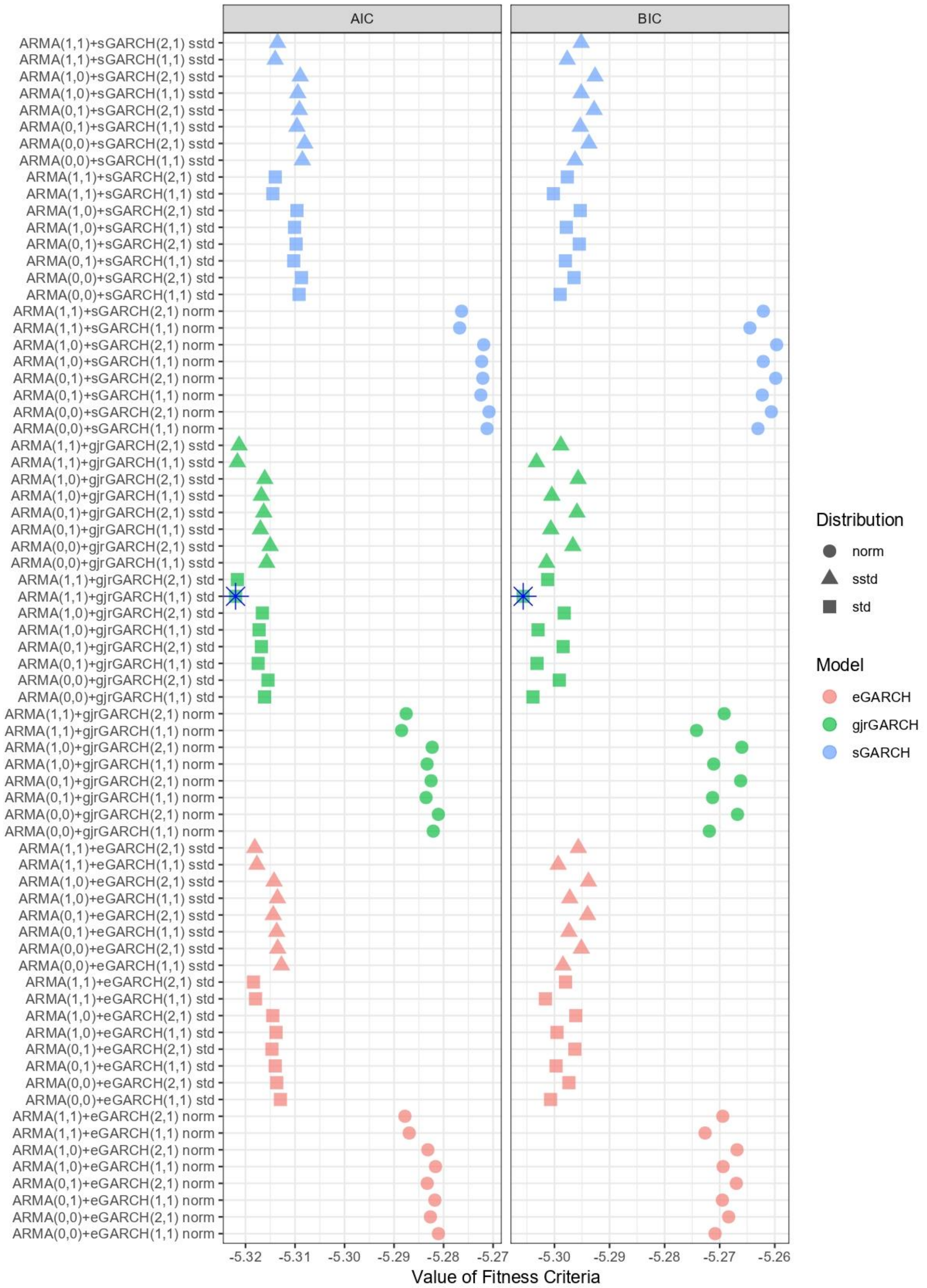
GARCH Models by Fitness Criteria for ABG

eGARCH(1,1); ARFIMA(1,0,1)

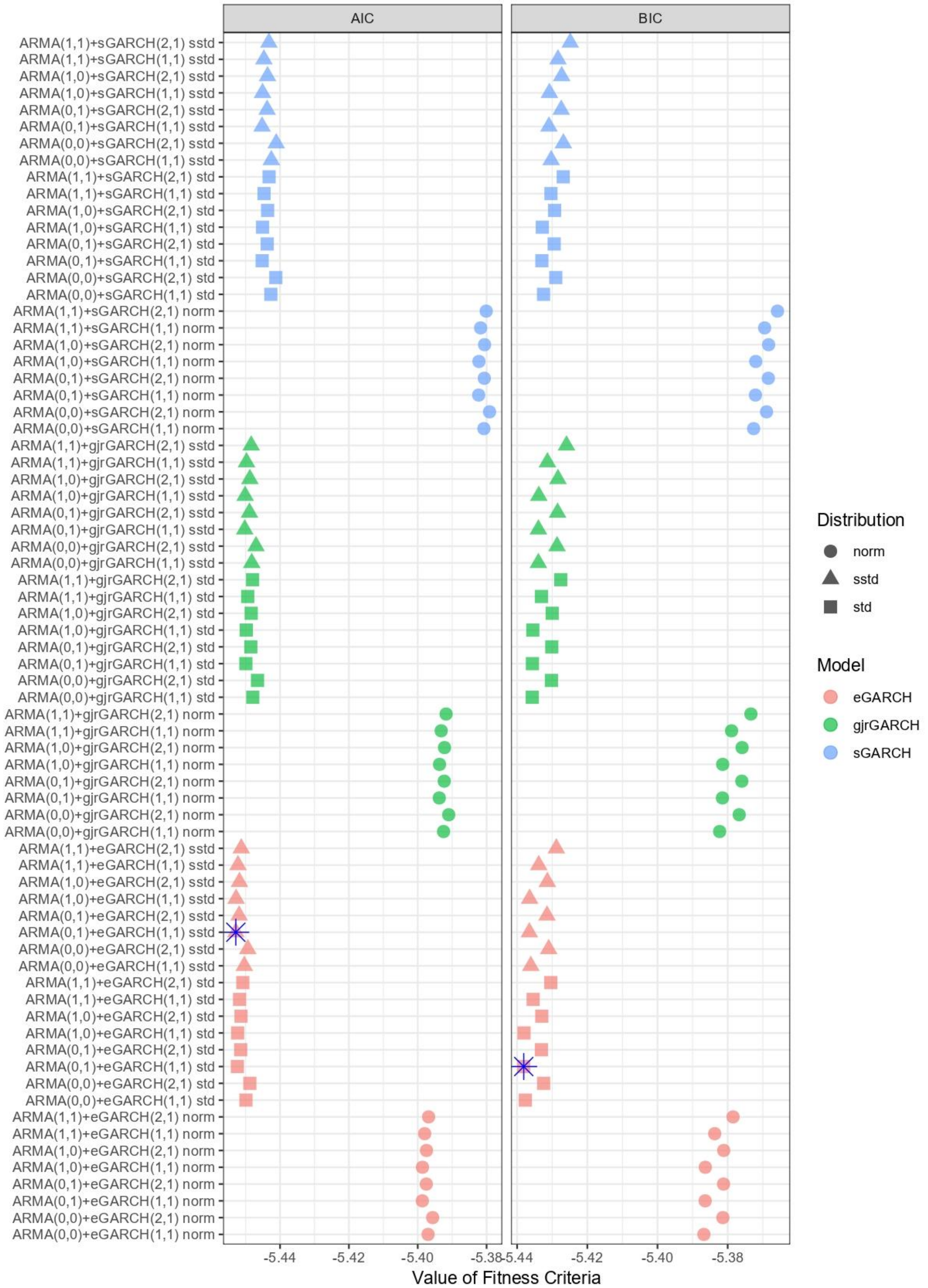


GARCH Models by Fitness Criteria for CPI
eGARCH(2,1); ARFIMA(1,0,0)

GARCH Models by Fitness Criteria for FSR
gjrGARCH(1,1); ARFIMA(1,0,1)

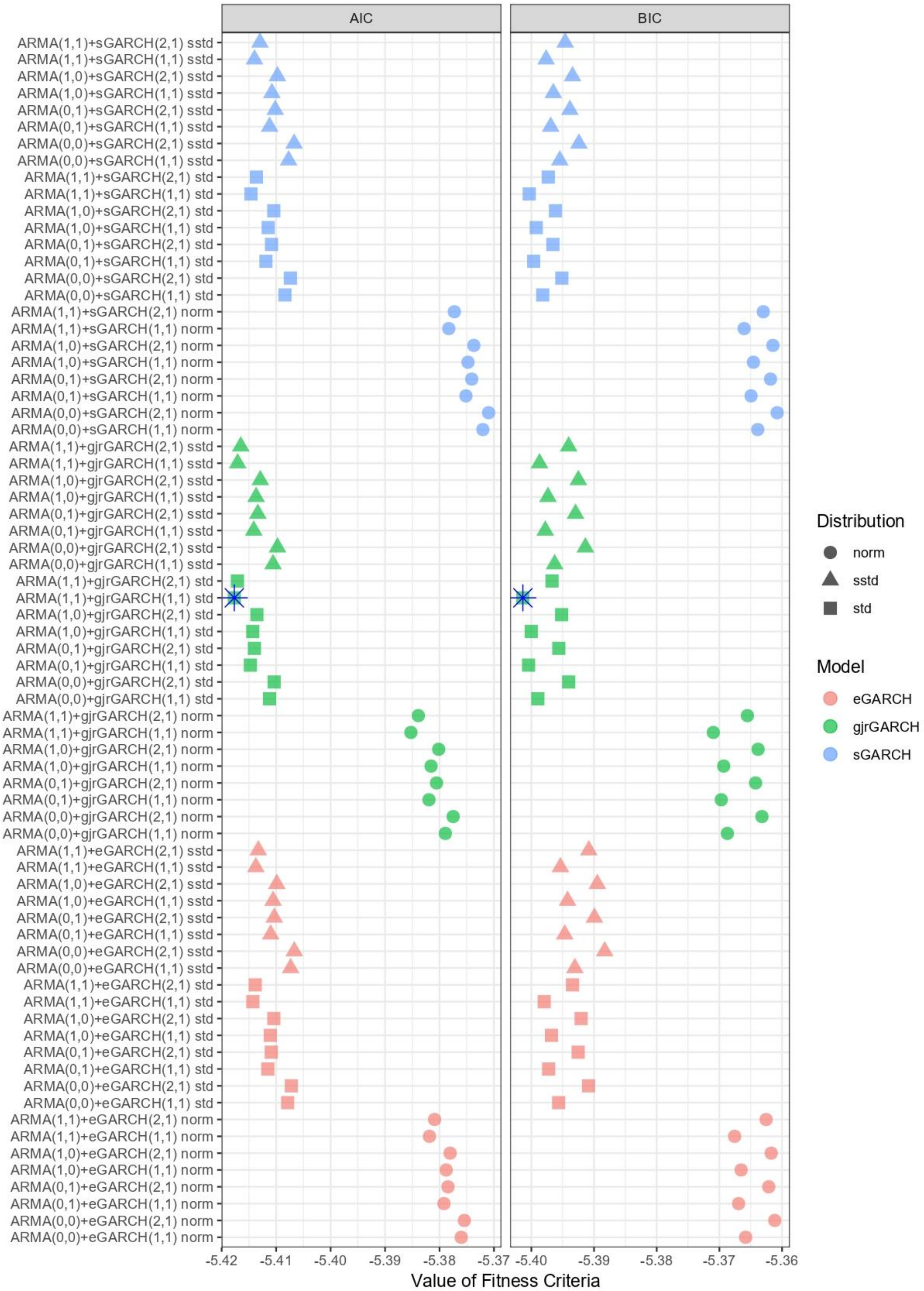


GARCH Models by Fitness Criteria for INL
eGARCH(1,1); ARFIMA(0,0,1)

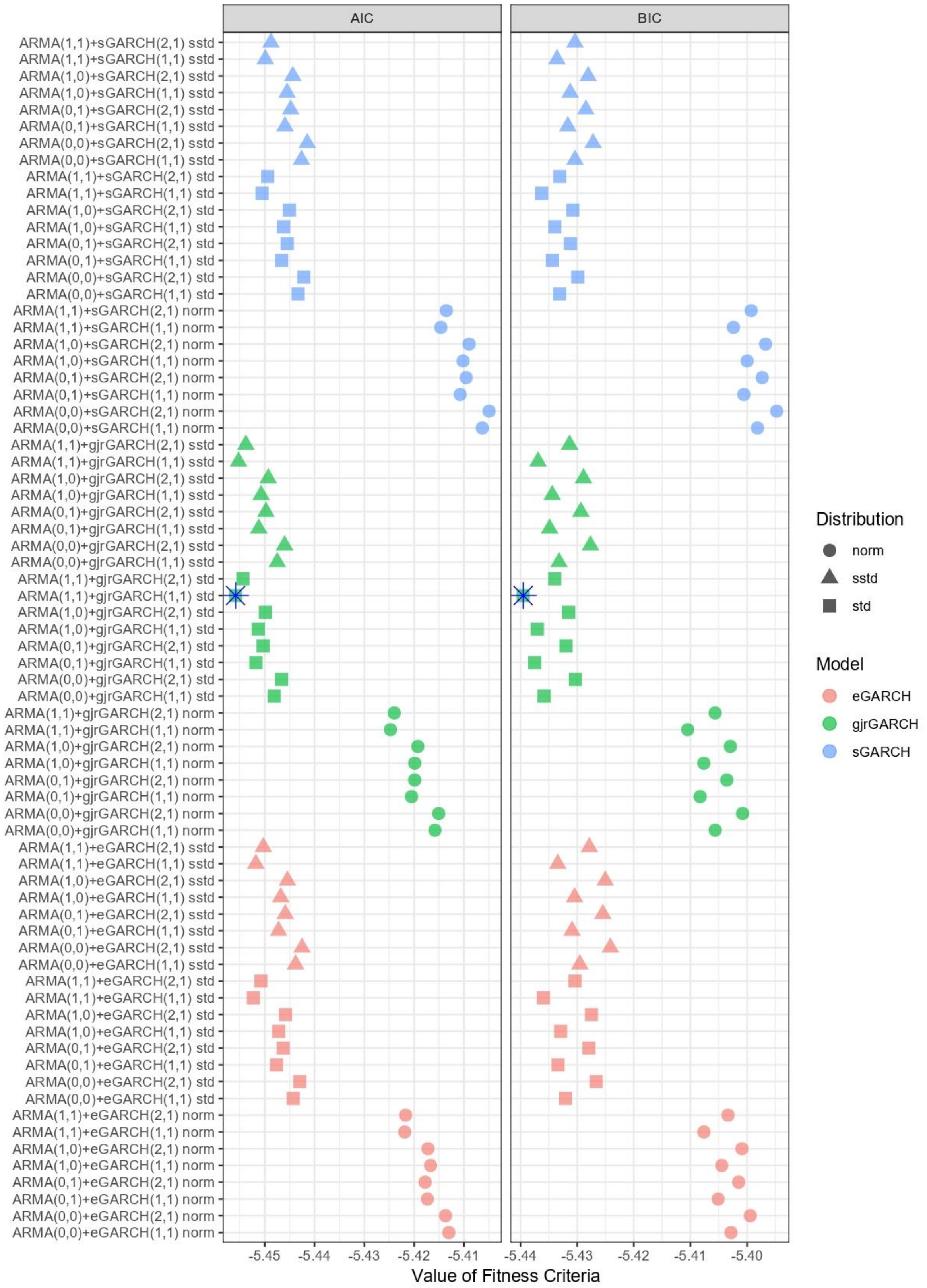


GARCH Models by Fitness Criteria for NED

gjrGARCH(1,1); ARFIMA(1,0,1)



GARCH Models by Fitness Criteria for SBK
gjrGARCH(1,1); ARFIMA(1,0,1)



7 REFERENCES

BAE, K.-H., WANG, J. & KANG, J.-K. 2012. The Costs of ETF Membership: Valuation Effect of ETFs on Underlying Firms. *SSRN Electronic Journal*.

BEN-DAVID, I., FRANZONI, F. & MOUSSAWI, R. 2018. Do ETFs Increase Volatility? *The Journal of Finance*, 73**,** 2471-2535.

BLACK, F. 1976. Studies of stock market volatility changes. *Proceedings of the American Statistical Association, Business & Economic Statistics Section, 1976*.

BLACKROCK 2018. Spot Trends: 5 go-to ETF statistics to watch in high-velocity markets.

BOJINOV, S. 2015. *A Brief History of ETFs* [Online]. ETF Database. Available: https://etfdb.com/etf-industry/a-brief-history-of-etfs/ [Accessed 30 January 2020 2020].

BOLLERSLEV, T. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31**,** 307-327.

BONEY-DUTRA, V., GUIRGUIS, H. & MUELLER, G. R. 2013. Did Intraday Trading by Leveraged and Inverse Leveraged ETFs Create Excess Price Volatility? A Look at REITs and the Broad Market. *The Journal of Real Estate Portfolio Management*, 19**,** 1-16.

CHENG, M. & MADHAVAN, A. 2009. The Dynamics of Leveraged and Inverse Exchange-Traded Funds. *Journal of Investment Management : JOIM***,** 7.

CONT, R. 2001. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1**,** 223.

CORBET, S. & TWOMEY, C. 2014. Have Exchange Traded Funds Influenced Commodity Market Volatility? *International Journal of Economics and Financial Issues*, 4**,** 323-335.

CURCIO, R. J., ANDERSON, R. I., GUIRGUIS, H. & BONEY, V. 2012. Have leveraged and traditional ETFs impacted the volatility of real estate stock prices? *Applied Financial Economics,* 22**,** 709-722.

ENGLE, R. 2001. GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics. *The Journal of economic perspectives,* 15**,** 157-168.

ENGLE, R. F. 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica,* 50**,** 987-1007.

FLORES, I. M. 2015. Do Leveraged ETFs Induce Volatility on their Underlying Indices? The German Case. *Working Paper, European Central Bank*

FRANCQ, C. & ZAKOIAN, J.-M. 2019. *GARCH models: structure, statistical inference and financial applications*, John Wiley & Sons.

GASTINEAU, G. L. 2010. *The Exchange Traded Funds Manual*, John Wiley & Sons.

GLOSTEN, L. R., JAGANNATHAN, R. & RUNKLE, D. E. 1993. ON THE RELATION BETWEEN THE EXPECTED VALUE AND THE VOLATILITY OF THE NOMINAL EXCESS RETURN ON STOCKS. *The Journal of finance (New York),* 48**,** 1779-1801.

GORDON, J. 2022. *Passive ownership of S&P 500 doubles in seven years* [Online]. ETF Stream. Available: https://www.etfstream.com/articles/passive-ownership-of-sp-500-doubles-in-seven-years [Accessed].

GREGORIOU, G. N. 2015. *Handbook of high frequency trading,* Academic Press.

HANCK, C., ARNOLD, M., GERBER, A. & SCHMELZER, M. 2023. Introduction to Econometrics with R.

HEGDE, S. P. & MCDERMOTT, J. B. 2004. The market liquidity of DIAMONDS, Q's, and their underlying stocks. *Journal of banking & finance,* 28**,** 1043-1067.

HILL, J., NADIG, D. & HOUGAN, M. 2015. *A Comprehensive Guide To Exchange Traded Funds*, CFA Institute Research Foundation.

KAPLAN, P. 2021. *In Long History of Market Crashes, Coronavirus Crash Was the Shortest* [Online]. Morningstar. Available: https://www.morningstar.com/portfolios/long-history-market-crashes-coronavirus-crash-was-shortest [Accessed 24 January 2024 2024].

LETTAU, M. & MADHAVAN, A. 2018. Exchange-Traded Funds 101 for Economists. *Journal of Economic Perspectives,* 32**,** 135-154.

LI, M. & ZHAO, X. 2014. Impact of leveraged ETF trading on the market quality of component stocks. *The North American Journal of Economics and Finance,* 28**,** 90-108.

LIN, C.-C. & CHIANG, M.-H. 2006. Volatility effect of ETFs on the constituents of the underlying Taiwan 50 Index. *Applied Financial Economics,* 15**,** 1315-1322.

MADURA, J. & NGO, T. 2008. Impact of ETF inception on the valuation and trading of component stocks. *Applied Financial Economics,* 18**,** 995-1007.

MAHLANGU, A. 2017. *JSE rallies to two-week high as investors cheer Ramaphosa victory* [Online]. Business Day. Available: https://www.businesslive.co.za/bd/markets/2017-12-19-jse-rallies-to-two-week-high-as-investors-cheer-ramaphosa-victory/ [Accessed 24 January 2024 2024].

MATARUTSE, J. 2014. Volatility characteristics of stocks underlying Exchange Traded Funds in South Africa. *Journal of Economics and Behavioral Studies,* 6**,** 829-839.

NELSON, D. B. 1991. Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica,* 59**,** 347-370.

PATRO, D. K. 2001. MARKET SEGMENTATION AND INTERNATIONAL ASSET PRICES: EVIDENCE FROM THE LISTING OF WORLD EQUITY BENCHMARK SHARES. *The Journal of financial research,* 24**,** 83-98.

RAMOS, M. S. P. M. M. D. F. V. H. P. 2021. A GARCH Tutorial with R. *Journal of Contemporary Administration*.

RICHIE, N. & MADURA, J. 2007. Impact of the QQQ on liquidity and risk of the underlying stocks. *The Quarterly Review of Economics and Finance,* 47**,** 411-421.

SHUM, P., HEJAZI, W., HARYANTO, E. & RODIER, A. 2016. Intraday Share Price Volatility and Leveraged ETF Rebalancing. *Review of Finance,* 20**,** 2379-2409.

TRAINOR, W. J. 2010. Do Leveraged ETFs Increase Volatility. *Technology and Investment,* 01**,** 215-220.

TREANOR, J. 2015. *South Africa finance minister's departure drives rand down to all-time low* [Online]. The Guardian. Available: https://www.theguardian.com/business/2015/dec/10/south-africa-finance-minister-nene-departure-drives-rand-down-to-all-time-low [Accessed 24 January 2024 2024].

TSAY, R. S. 2005. *Analysis of financial time series,* Hoboken, N.J, Wiley.

VLASTELICA, R. 2017. *Here's how much ETFs are dominating on the trading floor* [Online]. Market Watch. Available: https://www.marketwatch.com/story/heres-how-much-etfs-are-dominating-on-the-trading-floor-2017-01-10#:~:text=In%202016%2C%20ETFs%20accounted%20for,the%20first%20launched%20in%201993. [Accessed 6 August 2020 2020].

WANG, H. & XU, L. 2019. Do exchange-traded fund flows increase the volatility of the underlying index? Evidence from the emerging market in China. *Accounting & Finance,* 58**,** 1525-1548.

WERMERS, R. & XUE, J. 2015. Intraday ETF Trading and the Volatility of the Underlying. *A Research Paper Prepared for Lyxor Asset Management*.

WONGOUTONG, C. 2021. Imputation Methods in Time Series with a Trend and a Consecutive Missing Value Pattern. 866-879.

XU, L., GAO, H., SHI, Y. & ZHAO, Y. 2020. The heterogeneous volume-volatility relations in the exchange-traded fund market: Evidence from China. *Economic Modelling,* 85**,** 400-408.

XU, L. & YIN, X. 2017. Exchange Traded Funds and Stock Market Volatility. *International Review of Finance,* 17**,** 525-560.

XU, L., YIN, X. & ZHAO, J. 2018. Differently motivated exchange traded fund trading activities and the volatility of the underlying index. *Accounting & Finance,* 59**,** 859-886.

YATES, T. 2018. *Dissecting Leveraged ETF Returns* [Online]. Investopedia

Available: https://www.investopedia.com/articles/exchangetradedfunds/07/leveraged-etf.asp [Accessed 16 April 2020 2020].