

UNIVERSITY OF KWAZULU-NATAL

**INVESTOR OVERCONFIDENCE IN THE SOUTH AFRICAN
EXCHANGE TRADED FUND MARKET**

by

Damien Kunjal

215057603

**A thesis submitted in fulfilment of the requirements for the degree of
Master of Commerce**

**School of Accounting, Economics and Finance
College of Law and Management Studies
Westville Campus**


Supervisor: Ms. Faezah Peerbhai

2020

DECLARATION

I**Damien Kunjal** declare that

- (i) The research reported in this thesis, except where otherwise indicated, is my original research.
- (ii) This thesis has not been submitted for any degree or examination at any other university.
- (iii) This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
- (iv) This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
 - a) their words have been re-written but the general information attributed to them has been referenced;
 - b) where their exact words have been used, their writing has been placed inside quotation marks, and referenced.
- (v) Where I have reproduced a publication of which I am author, co-author or editor, I have indicated in detail which part of the publication was actually written by myself alone and have fully referenced such publications.
- (vi) This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Student's Signature:.....

Date:03 August 2020....

ACKNOWLEDGEMENTS

“If you’re trying to achieve, there will be roadblocks. I’ve had them; everybody has had them. But obstacles don’t have to stop you. If you run into a wall, don’t turn around and give up. Figure out how to climb it, go through it, or work around it.” ~ Michael Jordan

Despite this thesis being an individual work, I wish to express my deepest gratitude to all the people whose support and guidance has been monumental towards the success of this study.

First and foremost, I am deeply indebted to my supervisor, Ms. Faezah Peerbhai, who has mentored me throughout this process, for which my mere expression of gratitude does not suffice. I thank you for using your wealth of knowledge and dedicating countless hours towards providing constructive feedback which always aimed at reaching greater heights and depths of my thesis. Thank you for continually and convincingly motivating me to give it my best shot when the road got difficult and allowing me the freedom to reach the threshold of my mind. Without your continuous support, guidance, advice, and sincerity, this thesis would not have been possible. I wish you the very best with completing your PhD.

I would like to pay special regards to all my friends and family whose love and support are with me in whatever I pursue. A special thank you goes out to my mother for persistently supporting and encouraging me throughout my career. My gratitude is extended to my friend and colleague, Delane Naidu, for listening to all my complaints and motivating me during my perplexities. Thank you to all my colleagues in the Discipline of Finance and fellow Academic Development Officers in the College of Law and Management Studies who have contributed to my journey in various ways.

Financial assistance provided by the National Research Foundation (NRF) in respect to the costs of this study is hereby acknowledged. This work is based on the research supported wholly by the National Research Foundation of South Africa (Grant Number: 122160).

Most importantly, praises and thanks to God, the Almighty, for showering me with strength, wisdom, knowledge, and understanding along this journey. Blessed are they who find wisdom and those who gain understanding and without His generous grace, I would have not been able to realise my goals.

GLOSSARY OF ACRONYMS

ADF - Augmented Dickey-Fuller
ADR - American Depository Receipt
AIC - Akaike Information Criterion
AMEX - American Stock Exchange
ARCH - Autoregressive Conditional Heteroskedasticity
BAPM - Behavioural Asset Pricing Model
BLUE - Best Linear Unbiased Estimator
BPT - Behavioural Portfolio Theory
CAPM - Capital Asset Pricing Model
CARBS - Canada, Australia, Russia, Brazil, and South Africa
COVID-19 - Coronavirus
CRSP - Center for Research in Security Prices
DF - Degrees of Freedom
EGARCH - Exponential Generalised Autoregressive Conditional Heteroskedasticity
EMH - Efficient Market Hypothesis
ETF - Exchange Traded Fund
EUT - Expected Utility Theory
GARCH - Generalised Autoregressive Conditional Heteroskedasticity
GED - Generalised Error Distribution
HQ - Hannan and Quinn Information Criterion
JB - Jarque-Bera
JSE - Johannesburg Stock Exchange
LM - Lagrange Multiplier
LR - Likelihood Ratio
MAD - Mean Absolute Deviation
MPT - Markowitz Portfolio Theory
NAV - Net Asset Value
NYSE - New York Stock Exchange
OLS - Ordinary Least Squares
PP - Phillips-Perron
REIT - Real Estate Investment Trust
RII - Repeated-Imputation Inference
SC - Schwarz Information Criterion

SCF - Surveys of Consumer Finance

SUR - Seemingly Unrelated Regression

U.S. - United States

VAR – Vector Autoregression

VECM - Vector Error Correction Model

ZIB - Zero-Inflated Beta

ABSTRACT

In recent years, Exchange Traded Funds (ETFs) have transformed the investment management landscape. Despite the soaring popularity of ETFs, ETF traders may not always be rational. Mispricing of securities, excess trading volume, and excess return volatility present in financial markets can be attributed to the influence of the overconfidence bias. Several existing studies have explored the overconfidence bias in stocks markets, however, studies on investor overconfidence in ETF markets remain scanty. Therefore, the objective of this study is to investigate the presence of investor overconfidence in the South African ETF market.

Vector Autoregressive (VAR) models are employed to examine the lead-lag relationship between market turnover and market return for the market of South African ETFs tracking domestic benchmarks and for the market of South African ETFs tracking international benchmarks from the inception of the first ETF till August 2019. Consistent with the overconfidence hypothesis, a positive and significant relationship between current market turnover and lagged market returns is found for both markets, even after controlling for market volatility and cross-sectional return dispersion. This relationship holds for both market and individual ETF turnover indicating that the overconfidence bias also influences the trading activities of individual ETFs in both markets.

Additionally, using Exponential Generalised Autoregressive Conditional Heteroskedasticity (EGARCH) models, this study reports that overconfident trading exhibits a significant positive effect on the volatility of market return over the full sample periods. Notably, the sub-period analysis reveals that, there is a significant positive relationship between overconfident trading and market return volatility before and during the 2008 global financial crisis only in the market of ETFs tracking domestic benchmarks. However, for the post-crisis subsample, the positive effect of overconfident trading on market volatility is only significant for the market of ETFs tracking international benchmarks. These findings have important implications for ETF investors and traders who trade in the South African ETF market; investment management companies that guide investment decisions; as well as policymakers and regulators who are responsible for promoting the efficiency of the South African ETF market.

TABLE OF CONTENTS

DECLARATION	ii
ACKNOWLEDGEMENTS	iii
GLOSSARY OF ACRONYMS	iv
ABSTRACT	vi
CHAPTER 1: INTRODUCTION	1
1.1. Background and Problem Definition	1
1.1.1. Market Efficiency and Behavioural Finance.....	1
1.1.2. The Popularity of Exchange Traded Funds (ETFs)	2
1.1.3. Problem Statement	5
1.1.4. Research Problem	8
1.2. Research Objectives.....	8
1.3. Scope and Method of This Study	8
1.3.1. Scope of This Study	8
1.3.2. Method of This Study	9
1.4. Structure of This Thesis	9
CHAPTER 2: THEORETICAL FOUNDATIONS	11
2.1. Overview.....	11
2.2. Theories Relating to Investor Behaviour	11
2.2.1. Traditional Finance Approach to Investor Behaviour.....	11
2.2.2. Behavioural Finance Approach to Investor Behaviour	13
2.3. Theoretical Foundations of the Overconfidence Bias	16
2.3.1. Categories of Overconfidence.....	16
2.3.2. Overconfidence and Trading Volume	18
2.3.3. Overconfidence and Market Return	18

2.3.4. Overconfidence and the Disposition Effect	19
2.3.5. Overconfidence and Market Efficiency	20
2.3.6. Overconfidence and Market Volatility.....	21
2.3.7. Overconfidence and Investor Wealth.....	22
2.3.8. Factors Affecting Overconfidence	22
2.4. Theoretical Aspects of ETFs.....	23
2.4.1. ETF Characteristics.....	24
2.4.1.1. Creation and Redemption Process	24
2.4.1.2. Diversification.....	26
2.4.1.3. Low Cost.....	26
2.4.1.4. Trading Flexibility	26
2.4.1.5. ETF Market Impacts	27
2.4.2. The Popularity of ETFs.....	27
2.5. Summary	28
CHAPTER 3: REVIEW OF EMPIRICAL STUDIES.....	29
3.1. Overview.....	29
3.2. Overconfidence in Other Asset Classes	29
3.2.1. Statman, <i>et al.</i> (2006).....	29
3.2.2. Griffin, <i>et al.</i> (2007).....	31
3.2.3. Kyrychenko and Shum (2009)	32
3.2.4. Yung and Liu (2009).....	32
3.2.5. Zaiane and Abaoub (2009).....	33
3.2.6. Lin, Rahman and Yung (2010)	33
3.2.7. Bailey, <i>et al.</i> (2011).....	35
3.2.8. Chuang and Susmel (2011)	35
3.2.9. Cohen and Shin (2013)	35

3.2.10. Metwally and Darwish (2015)	36
3.2.11. Pak and Chatterjee (2016).....	37
3.2.12. Zia, Sindhu and Hashmi (2017)	37
3.2.13. Aharon and Qadan (2018).....	38
3.2.14. Gupta, Goyal, Kalakbandi and Basu (2018)	38
3.2.15. Chen and Sabherwal (2019).....	40
3.2.16. Alsabban and Alarfaj (2020).....	41
3.3. Overconfidence and Market Volatility	42
3.3.1. Chuang and Lee (2006).....	42
3.3.2. Sheikh and Riaz (2012).....	43
3.3.3. Abbas (2013).....	44
3.3.4. Jlassi, <i>et al.</i> (2014)	45
3.3.5. Mushinada and Veluri (2018)	45
3.4. Overconfidence in ETF Markets.....	46
3.4.1. Da Dalt, <i>et al.</i> (2019)	46
3.5. Evidence From South Africa.....	46
3.5.1. Charteris, Chau, Gavriilidis and Kallinterakis (2014).....	47
3.5.2. Willows and West (2015)	47
3.5.3. Dowie and Willows (2016).....	47
3.5.4. Charteris and Musadziruma (2017).....	48
3.5.5. Charteris and Rupande (2017)	48
3.6. Summary	48
CHAPTER 4: DATA AND METHODOLOGY	51
4.1. Overview.....	51
4.2. Dataset.....	51
4.2.1. Sample Period	51

4.2.2. Data Frequency	52
4.2.3. Survivorship Bias.....	53
4.2.4. Sample of ETFs.....	54
4.2.5. Types of Data and Sources of Data.....	56
4.2.6. Computation of Main Variables.....	56
4.2.7. Preliminary Data Analysis	60
4.3. Method Used to Detect Investor Overconfidence.....	66
4.3.1. Market VAR Models.....	66
4.4. Method Used to Determine Whether Investor Overconfidence Influences the Trading Activities of Individual ETFs	71
4.4.1. Individual Security VAR Model.....	71
4.4.2. Individual Security VAR Model with Panel Data Approach.....	73
4.5. Method Used to Examine the Effect of Investor Overconfidence on Market Volatility	75
4.6. Sub-Period Analysis of the Effect of Investor Overconfidence on the Volatility of the South African ETF Market.....	79
4.7. Validity and Reliability.....	79
4.8. Summary	81
CHAPTER 5: DATA ANALYSIS AND RESULTS.....	83
5.1. Overview.....	83
5.2. Preliminary Data Analysis	83
5.2.1. Number of ETFs Included in This Study	83
5.2.2. Stationarity of Variables	84
5.2.3. Descriptive Statistics.....	88
5.2.4. Correlation Analysis	90
5.2.5. Model Diagnostics Tests.....	92
5.3. Empirical Results of Market-Wide Investor Overconfidence.....	96
5.3.1. Optimal Lag Length Selection	96

5.3.2. Market Vector Autoregression (VAR) Models.....	99
5.3.3. Granger Causality Tests.....	106
5.3.4. Market Impulse Response Functions.....	108
5.4 Empirical Results for the Influence of Investor Overconfidence on Individual ETFs.....	113
5.4.1. Individual Security Panel Vector Autoregression (VAR) Models.....	114
5.4.2. Granger Causality Tests.....	118
5.4.3. Impulse Response Functions.....	120
5.5. Empirical Results of the Effect of Investor Overconfidence on Market Volatility.....	124
5.6. Empirical Results for the Sub-Period Analysis of the Effect of Overconfidence on Volatility.....	131
5.7 Summary.....	136
CHAPTER 6: CONCLUSIONS AND RECOMMENDATIONS.....	138
6.1. Review of Objectives.....	138
6.2. Summary of Findings.....	139
6.2.1. Research Question One - Are Investors Overconfident When Trading in the South African ETF Market?.....	139
6.2.2. Research Question Two – Does Investor Overconfidence Influence the Trading Activity of Individual ETFs?.....	139
6.2.3. Research Question Three – Does Investor Overconfidence Exhibit a Positive or Negative Effect on the Volatility of the Returns From the South African ETF Market?.....	140
6.2.4. Research Question Four – Does the Volatility of the South African ETF Market Respond Positively or Negatively to Trading Volume Induced by Investor Overconfidence Before, During, and After the 2008 Global Financial Crisis?.....	140
6.3. Implications of Findings.....	141
6.4. Recommendations for Future Studies.....	142
6.5. Conclusion.....	143

APPENDIX.....	144
Appendix 1: Summary of JSE-listed ETFs as at 30 August 2019	144
Appendix 2: VAR Model Estimates for the Market of ETFs Tracking Domestic Benchmarks.....	147
Appendix 3: VAR Model Estimates for the Market of ETFs Tracking International Benchmarks.....	148
Appendix 4: Impulse Response Table for the VAR Model Estimated for the Market of ETFs with Domestic Benchmarks	148
Appendix 5: Impulse Response Table for the VAR Model Estimated for the Market of ETFs with International Benchmarks	150
Appendix 6: Panel VAR Model Estimates for the Market of ETFs Tracking Domestic Benchmarks .	151
Appendix 7: Panel VAR Model Estimates for the Market of ETFs Tracking International Benchmarks	152
Appendix 8: Impulse Response Functions for Individual ETFs in the Market of ETFs Tracking Domestic Benchmarks.....	153
Appendix 9: Impulse Response Functions for Individual ETFs in the Market of ETFs Tracking International Benchmarks	153
Appendix 10: Impulse Response Table for Individual ETFs in the Market of ETFs Tracking Domestic Benchmarks.....	154
Appendix 11: Impulse Response Functions for Individual ETFs in the Market of ETFs Tracking International Benchmarks	154
Appendix 12: Results of the Estimated Regressions Used to Distinguish Between Overconfident Trading and Non-Overconfident Trading for the Market of ETFs Tracking Domestic Benchmarks	155
Appendix 13: Results of the Estimated Regressions Used to Distinguish Between Overconfident Trading and Non-Overconfident Trading for the Market of ETFs Tracking International Benchmarks	156
Appendix 14: Ethical Clearance Letter	157
Appendix 15: Amended Ethical Clearance Letter for Change in Title	158
 BIBLIOGRAPHY	 159

LIST OF TABLES

Table 3.1: Summary of Studies Reviewed in This Chapter	49
Table 4.1: Summary of Subsamples	52
Table 4.2: Sample of ETFs	55
Table 4.3: Summary of Empirical Methods.....	82
Table 5.1: Results from the PP Unit Root Test.....	86
Table 5.2: Results from the Choi Panel Unit Root Test.....	87
Table 5.3: Results from the PP Unit Root Test.....	87
Table 5.4: Descriptive Statistics of Each Monthly Market Variable	88
Table 5.5: Descriptive Statistics of Each Monthly Security Variable.....	88
Table 5.6: Descriptive Statistics of Each Daily Market Variable	89
Table 5.7: Correlation Coefficients Between Each Portfolio’s Market Variables	90
Table 5.8: Correlation Coefficients Between Variables in Panel VAR Model.....	90
Table 5.9: Correlation Coefficients Between Variables in the EGARCH Model.....	91
Table 5.10: LM Test Results for the Market VAR Model Estimated for the Market of ETFs With Domestic Benchmarks.....	93
Table 5.11: LM Test Results for the Market VAR Model Estimated for the Market of ETFs With International Benchmarks	94
Table 5.12: Heteroskedasticity Test Results for the Market VAR Model Estimated for the Market of ETFs With Domestic Benchmarks	94
Table 5.13: Heteroskedasticity Test Results for the Market VAR Model Estimated for the Market of ETFs With International Benchmarks	95
Table 5.14: ARCH-LM Diagnostics Test Results for the EGARCH (1,1) Model Estimated for the Market of ETFs With Domestic Benchmarks	95
Table 5.15: ARCH-LM Diagnostics Test Results for the EGARCH (1,1) Model Estimated for the Market of ETFs With International Benchmarks	95
Table 5.16: Lag Order Selection Criteria for the Endogenous Variables in the VAR Model Estimated for the Market of ETFs Tracking Domestic Benchmarks.....	97
Table 5.17: Lag Order Selection Criteria for the Endogenous Variables in the VAR Model Estimated for the Market of ETFs Tracking International Benchmarks	97
Table 5.18: Lag Order Selection Criteria for the Exogenous Variables in the VAR Model Estimated for the Market of ETFs Tracking Domestic Benchmarks	98

Table 5.19: Lag Order Selection Criteria for the Exogenous Variables in the VAR Model Estimated for the Market of ETFs Tracking International Benchmarks	99
Table 5.20: VAR Model Estimates for the Market of ETFs Tracking Domestic Benchmarks	100
Table 5.21: VAR Model Estimates for the Market of ETFs Tracking International Benchmarks	101
Table 5.22: Granger Causality Test Results for the Market of ETFs Tracking Domestic Benchmarks ...	107
Table 5.23: Granger Causality Test Results for the Market of ETFs Tracking International Benchmarks	107
Table 5.24: Panel VAR Model Estimates for ETFs Tracking Domestic Benchmarks	115
Table 5.25: Panel VAR Model Estimates for ETFs Tracking International Benchmarks	116
Table 5.26: Granger Causality Test Results for Individual ETFs in the Market of ETFs Tracking Domestic Benchmarks.....	118
Table 5.27: Granger Causality Test Results for Individual ETFs in the Market of ETFs Tracking International Benchmarks	119
Table 5.28: Lag Order Selection Criteria for Market Return in the Market of ETFs Tracking Domestic Benchmarks.....	125
Table 5.29: Lag Order Selection Criteria for Market Return in the Market of ETFs Tracking International Benchmarks.....	126
Table 5.30: PP Unit Test Results	126
Table 5.31: Correlation Coefficients Computed Between the Respective EGARCH Variables	127
Table 5.32: Results of the Estimated EGARCH Models	128
Table 5.33: Results of the EGARCH Models Estimated for the Market of ETFs Tracking Domestic Benchmarks During Different Subsamples.....	132
Table 5.34: Results of the EGARCH Models Estimated for the Market of ETFs Tracking International Benchmarks During Different Subsamples.....	133
Table 5.35: Summary of Results.....	137

LIST OF FIGURES

Figure 1.1: Market Capitalisation of the South African ETF Market	3
Figure 2.1: The Value Function	14
Figure 2.2: Creation and Redemption Process of an ETF.....	25
Figure 5.1: Number of ETFs Included in the Market-Value Weighted Portfolios	84
Figure 5.2: Graphical Representation of the Monthly Market Variables for the Market of ETFs With Domestic Benchmarks	85
Figure 5.3: Graphical Representation of the Monthly Market Variables for the Market of	85
Figure 5.4: Impulse Response Functions for the VAR Model Estimated for the Market of ETFs with Domestic Benchmarks	109
Figure 5.5: Impulse Response Functions for the VAR Model Estimated for the Market of ETFs with International Benchmarks	109
Figure 5.6: Impulse Response Functions for ETFs Tracking Domestic Benchmarks	121
Figure 5.7: Impulse Response Functions for ETFs Tracking International Benchmarks	121

CHAPTER 1: INTRODUCTION

1.1. Background and Problem Definition

1.1.1. Market Efficiency and Behavioural Finance

Traditional finance theories are based on two assumptions, namely; rationality and market efficiency. However, these two assumptions are not independent because the concept of rationality underpins the principal of market efficiency. Rational Choice Theory claims that market participants use logical procedures to solve problems and formulate strategies (Mushinada, 2020). Accordingly, the notion of market efficiency assumes that all investors are rational and make well-informed decisions (Dickason and Ferreira, 2018). The Efficient Market Hypothesis (henceforth, EMH) developed by Fama (1970) defines an efficient market as a market where all relevant information is fully reflected in the prices at which assets trade, such that, any change in an asset's price is due to the arrival of new information. Therefore, the EMH asserts that the amount an investor pays for a security is an efficient price that reflects the underlying fundamental value of the respective security (Mak and Ip, 2017). According to Fama (1970), the conditions for capital market efficiency include the following: all information is freely available, homogeneous beliefs about the implications of current information, and no transaction costs. However, these conditions are not descriptive of markets met in practice. On the contrary, investor psychology creates heterogeneous beliefs which drive the market price of a security away from its fundamental value (Daniel, Hirshleifer and Subrahmanyam, 1998). Hence, the EMH was and is still continuously criticised by researchers who argue that markets are not efficient. The criticism of the EMH has led to the field of behavioural finance.

Behavioural finance theories argue that investors are irrational because their investment decisions are based on their emotions, beliefs, and state of minds, contrary to the EMH which asserts that emotions do not influence rational decision-making processes (Virigineni and Rao, 2017). According to Joo and Durri (2015), behavioural finance theories examine how psychological and cognitive biases lead to irrational investment decisions, thus resulting in suboptimal investment choices and the existence of market anomalies. For instance, Economou, Hassapis and Philippas (2018) claim that individual investment decisions may be influenced by the perceptions and patterns of other market participants causing investors to herd toward a particular security. Furthermore, investors may trade overconfidently by overestimating their abilities, degree of control, or probability of success (Chen and Sabherwal, 2019). Moreover, market participants may overreact to market information by overweighting recent information and undervaluing fundamental information (Lerskullawat and Ungphakorn, 2018). Lowies, Hall, and Cloete (2016) find that heuristic-driven biases influence decisions made by fund managers in South Africa. Similarly, Dowie and Willows (2016) report that South African investors are influenced by behavioural biases when investing in

unit trusts. In a more recent study, Dickason and Ferreira (2018) prove that, in the South African market, individuals' investment decisions are influenced by several psychological and cognitive biases, including loss aversion, mental accounting, and self-control.

Zaiane and Abaoub (2009), Jlassi, Naoui and Mansour (2014), Yu (2014), and Baker, Kumar, Goyal and Gaur (2019) note that the overconfidence bias is the most prominent behavioural bias and is one of the most robust behavioural findings. According to Odean (1998), the overconfidence bias influences the decisions made by investors when these investors overestimate their own knowledge and ability to identify undervalued securities. Consequently, traders who are influenced by the overconfidence bias overreact to less relevant information but underreact to highly relevant news and information, causing security prices to deviate from their fundamental values (Odean, 1998). The overconfidence hypothesis proposed by Gervais and Odean (2001) asserts that traders who are overconfident associate market returns with their own abilities, and thus, trade more aggressively in periods after market gains. As a result, overconfident trading generates an increase in trading volume, market volatility and price distortions (Mushinada, 2020).

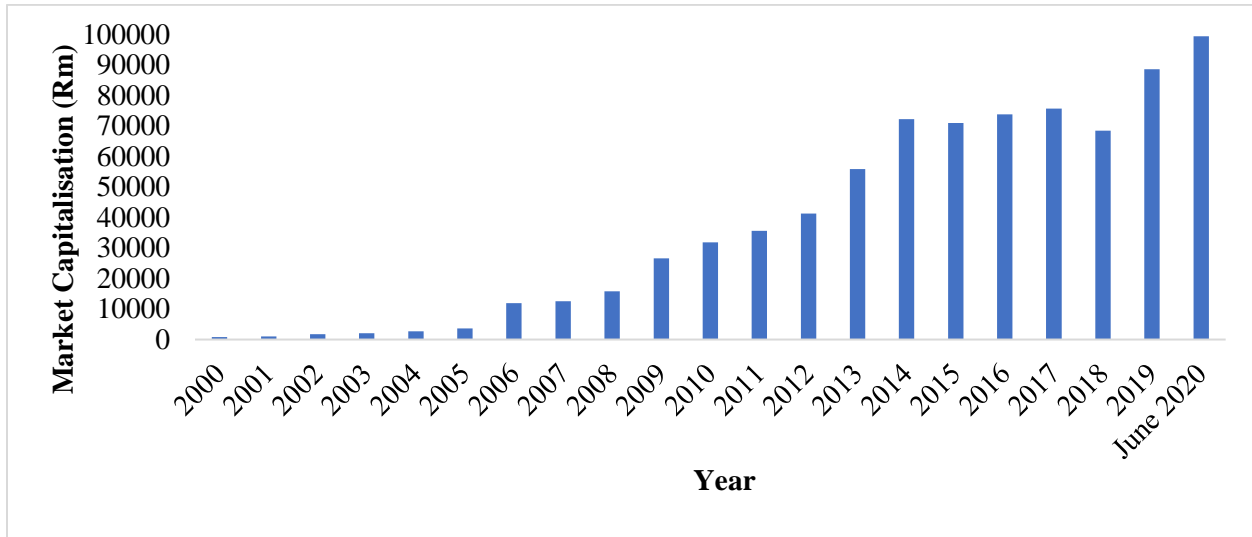
Studies by Statman, Thorley and Vorkink (2006), Abbes (2013), Metwally and Darwish (2015), and Alsabban and Alarfaj (2020) all found that investors are overconfident when trading in stock markets. However, alternative investment classes exhibit contradictory results. For instance, Bailey, Kumar and Ng (2011) report that mutual fund investors in the United States display overconfident trading behaviour, while Dowie and Willows (2016) discover that South African unit trust investors are underconfident as opposed to overconfident. To the knowledge of the author of this paper, there have not been any attempts to examine the existence of overconfident trading in South Africa's Exchange Traded Fund market; hence, this lack of research provides a basis for further analysis.

1.1.2. The Popularity of Exchange Traded Funds (ETFs)

An ETF is a pooled investment vehicle that consists of a combination of securities which are selected in an attempt to replicate the performance of a specific benchmark (Charteris, 2013). As such, ETFs give investors exposure to an array of financial instruments or assets via a single entry point (Petajisto, 2017). The first ETF was introduced in Canada in 1989 (Broms and Gastineau, 2007). In recent years, ETFs have changed the investment management landscape. By the end of June 2020, the assets invested in the global ETF industry reached approximately \$6 trillion, with more than 7 000 ETFs listed globally (Etfgi, 2020). With regards to the South African ETF market, the first ETF was launched in November 2000 (Charteris, 2013). While the initial growth of the South African ETF industry was sluggish, the popularity of ETFs has recently grown. At the end of June 2020, the South African ETF market realised a combined

market capitalisation of approximately R99 billion with 74 ETFs listed on the Johannesburg Stock Exchange (JSE henceforth) (Brown, 2020). This growth in the South African ETF market is illustrated in Figure 1.1 below.

Figure 1.1: Market Capitalisation of the South African ETF Market



Source: Author's own construction

Generally, ETFs represent an easily accessible, low-cost transformation of an index fund. These advantages have contributed to the global shift of investors away from actively managed funds to passively managed funds, like ETFs, since the majority of ETFs are passive funds (Neves, Fernandes and Martins, 2019; Steyn, 2019). According to Divine (2019) in Tokic (2019), the assets under the management of active funds trading in the United States (U.S.) was triple that of passive funds in 2009, whereas, in 2019, passive funds overtook active funds by market share. The main factors driving the shift to passive investment funds (like ETFs) is the high management costs associated with actively managed funds as well as the underperformance of active managers (Blitz and Huij, 2012; Elton, Gruber and de Souza, 2019; Neves, *et al.*, 2019; Steyn, 2019). The shift to passive investment funds have benefitted ETFs since ETFs have substantially lower expense ratios (Elton, *et al.*, 2019; Steyn, 2019).

The shift towards passive funds, like ETFs, is also the case in South Africa, where the ETF market has grown both in size and number, whilst active managers struggle to compete (Steyn, 2019). However, in South Africa, the size of the actively managed funds industry is still greater than that of the passively managed funds industry (Brown, 2019). Notably, according to Flood (2020), fees will no longer be the key point of competition between fund managers because of the recent turmoil in financial markets due to the

outbreak of the coronavirus (COVID-19). On the contrary, Dave Nadig, chief investment officer and director of research at ETF trends, argues that the manner in which ETFs have handled the stock market's outsized volatility (which resulted from the COVID-19 outbreak) is testament to the industry's strength, and therefore, is a huge victory for ETFs as a structure (Gurdus, 2020). Moreover, Adèle Hattingh, the manager of business development and Exchange Traded Products (ETPs) at the JSE, argues that, in today's volatile economic environment due to COVID-19, ETFs are attractive investment instruments because of their simplicity, diversification, liquidity, and cost advantages (IOL, 2020). Brown (2020) notes that, despite the market turmoil, the market capitalisation of the South African ETF industry increased by approximately R11 billion in the first half of 2020. A key contributor to the growth in the size of the South African ETF industry is the diversification benefits of ETFs (Brown, 2020). For instance, ETFs with direct exposure to global markets protect South African investors from losses in local equity markets (Brown, 2020).

Studies such as Madura and Richie (2004), Broman (2016), Ben-David, Franzoni and Moussawi (2018), and Da and Shive (2018) have found evidence of the ETF market attracting short-term (noise) traders. This result may be an indication that, despite the soaring popularity of ETFs, investment decisions made by ETF traders may not always be rational. For instance, Chen, Ho, Lai and Morales-Camargo (2011) and Bahadar, Mahmood and Zaman (2019) find that ETF investors display herd behaviour. Additionally, Ma, Ho, Yang, and Chu (2018) discover that ETF markets are influenced by investor overreaction. Irrational ETF investment choices have significant implications for the efficiency of ETFs and financial markets as a whole. This is because, the influence of behavioural biases on investment decision making may cause investors to trade too frequently, allow emotions to overrule logic, miscalculate probabilities, or trade at precisely the wrong times (Chen, Kim, Nofsinger and Rui, 2007). As a result, the efficiency of ETFs may be affected, and the trading volumes and returns generated by ETFs could become volatile. Therefore, there is a possibility that the exponential growth in the global and local ETF markets could pose a threat to broader financial markets by creating instability in these markets since ETFs have exposures to different asset classes, including; stocks, bonds, commodities, and currencies.

In 2019, several financial market experts questioned whether there is a passive investment bubble brewing which could subsequently lead to an ETF market crash (Tokic, 2019). For example, Michael Burry, a hedge fund manager at Scion Asset Management, who is renowned for correctly predicting the 2008 housing bubble and its subsequent financial crisis sees a bubble in passive investments (Kim and Cho, 2019; Li, 2019; Tokic, 2019). Michael Burry argues that as capital flows into ETFs and alternative index-tracking products, passively managed funds are inflating stock and bond prices in a similar way that collateralised

debt obligations did for subprime mortgages during the subprime financial crisis (Kim and Cho, 2019; Li, 2019). Similarly, Ben Volkwyn (2019), the head of private clients at Cannon Asset Managers, mentions that as investors continue to purchase ETFs without regard for fundamentals, share prices become decoupled, subsequently, increasing the potential effect of a market correction and leaving investors more susceptible to the impacts of a market crash.

The prediction of a passive investment bubble is supported by the theoretical arguments that active managers promote market efficiency by conducting research to beat the market, whereas passive managers contribute to market inefficiency by being price takers (Tokic, 2019). Therefore, the recent shift from actively managed funds to passively managed funds could increase the probability of price distortions and a passive investment bubble. Notably, the predicted passive investment bubble would have the most significant effect on the U.S. market, where the size of the passively managed funds industry is larger than that of the actively managed funds industry. However, Dovern and van Roye (2014) and Fink and Schüller (2015) provide evidence that systemic financial stress in the U.S is transmitted to emerging markets, like South Africa. Thus, a crash in the U.S. ETF market would generate systemic shocks that could affect the South African ETF market. Notably, if there is evidence of investor overconfidence in the South African ETF market, this could further exacerbate the impact of these systemic shocks. According to Abbes (2013), behavioural finance is an excellent way to identify the causes of the volatility experienced during the 2008 global financial crisis as well as to identify possible policy responses to mitigate the behaviours that lead to financial market instability. Thus, a further motivation for the current study is to determine whether overconfidence in the South African ETF market could increase the volatility experienced during an ETF market crash.

1.1.3. Problem Statement

The concept of market efficiency is centred on the assumption that, at any given time, prices of securities fully reflect all available information. Specifically, in an efficient market, market participants react instantaneously to any new information in an unbiased manner, and therefore, investors are unable to capitalise on any mispricing of securities (Dickason and Ferreira, 2018). The problem arises when investors make investment decisions based on their emotions, beliefs, and state of minds. This is because investment decisions based on heuristics and biases pose new challenges to the concept of market efficiency. The behaviour of market participants is an extensive topic that is entrenched in both standard financial theories as well as theories of investor psychology. However, the focus of this thesis is limited to the overconfidence bias because it is the most prominent bias according to Jlassi, *et al.* (2014) and Baker, *et al.* (2019).

Griffin, Nardari and Stulz (2007) discover that investor overconfidence is more pronounced in countries with high market volatility and in which there is a high level of corruption. According to Redl (2018), emerging markets, including South Africa, exhibit high levels of realised market volatility. Moreover, Salahuddin, Vink, Ralph and Gow (2019) mention that South Africa experiences several forms of corruption. Given that South Africa has both high market volatility and high levels of corruption, it is plausible to expect investors to be overconfident when trading in South African markets. The rationale for investigating whether investors exhibit overconfidence when trading in the South African ETF market stems from the soaring popularity of ETFs in recent years. Specifically, if investors recognise that ETFs are profitable, the correlation between overconfidence and the tendency to use ETFs may rise.

Several anomalies that are present in financial markets cannot be explained by traditional financial theories. However, some of these anomalies were successfully accounted for when the assumption of investor overconfidence was adopted. These anomalies include, but not limited to; high trading volumes, mispricing of securities, as well as excess volatility (Benos, 1998; Odean, 1998; Odean, 1999; Ko and Huang, 2007; Jlassi, *et al.*, 2014; Benigno and Karantounias 2019). According to Odean (1998), overconfident market participants overreact to subjective, less relevant information and underreact to abstract and highly relevant information. Consequently, investor overconfidence causes security prices to deviate from their fundamental values, resulting in excess price volatility around private signals. This may be due to market participants' incorrectly interpreting their private signals about the true values of assets (García, Sangiorgi and Urošević, 2007). The role of markets to efficiently allocate capital is threatened when the prices of securities depart from their intrinsic values, and this loss in efficiency can have a long-term impact (Bhattacharya and O'Hara, 2020). Notably, the loss of ETF market efficiency could have adverse effects on broader financial markets because ETFs have exposures to different asset classes.

A further problem is that overconfident investors trade more aggressively on winning stocks which, in turn, drives security prices up and leads to a persistent shift in their intrinsic values (Jlassi, *et al.*, 2014). Such behaviours generate an abnormal increase in market volatility and lead to price distortions that could trigger a market bubble (Chuang and Lee, 2006; De Grauwe, 2012; Abbes, 2013). Accordingly, Abbes (2013) and Jlassi, *et al.* (2014) find that the overconfidence bias contributed to the financial instability experienced during the 2008 global financial crisis. Thus, one could argue that overconfident trading by ETF investors could fuel the volatility experienced if there is an ETF market crash as predicted by market experts, thereby, necessitating the investigation of investor overconfidence in ETF markets. Notably, the study conducted by Da Dalt, Feldman, Garvey and Westerholm (2019) is the only study, known to the author, which covers overconfident trading in ETFs. Therefore, this lack of empirical studies that focus on overconfidence in

ETF markets provides a basis for further analysis. Hence, the current study is unique since it investigates investor overconfidence in an ETF market by applying the Statman, *et al.* (2006) model, which, to the knowledge of the author, has never been done before.

From the above discussion, it is evident that the stability of South African financial markets would be exposed to new threats if overconfident trading is present in South Africa's ETF market. Thus, the challenge for policymakers is to address those threats without restricting the freedom of markets to allocate resources efficiently. According to Bhattacharya and O'Hara (2020), these challenges may be addressed by analysing different segments of the asset management industry; however, this process would only be effective if it considers ETFs since these investment vehicles have modified the investment management industry in recent years. Therefore, this study is considered important for policymakers who are responsible for promoting the efficiency of the ETF market. Specifically, based on the results of this study, policymakers can implement policies that improve the efficiency of ETF markets by reducing the behavioural anomalies present in ETF markets. This is because examining financial data from the perspective of psychology may help identify the missing link between the prices at which ETFs trade and their fundamental values. Moreover, this study evaluates the effect of overconfidence on the volatility of the ETF market, and therefore, provides evidence of whether overconfident trading by ETF investors could increase the volatility of South Africa's ETF market. Therefore, based on the results of this study, policymakers can determine possible policy responses to reduce the behavioural biases that may fuel the volatility of the South African ETF market.

This study is also important for individual investors, traders, and other practitioners in the ETF industry. Traditionally, investors make investment decisions based on their personalities and risk tolerance levels, and therefore, investment companies only considered investors' personalities and risk tolerance levels when directing financial decisions. However, these same investors often depart from their initial investment decisions because cognitive and emotional biases begin driving their investment choices. As such, understanding what biases drive investment decisions will benefit both investment companies and individual investors by reducing their risk exposures and increasing their opportunities for generating higher returns since overconfident trading tends to reduce investor wealth (Trinugroho and Sembel, 2011). Moreover, since this study investigates whether the overconfidence bias influences investors trading in the ETF market, this study will help investors make rational investment decisions by reducing the effects of biased ETF investment choices. Overall, the present study fills a significant research gap by investigating whether the overconfidence bias influences ETF investment choices made by investors trading in South Africa's ETF market, which has not been examined before, to the knowledge of the author.

1.1.4. Research Problem

Based on the preceding problem statement, this study's primary focus is to answer the following question:

Are investors overconfident when trading in the South African ETF market?

1.2. Research Objectives

The objectives of this study are as follows:

- To investigate whether investors are overconfident when trading in the South African ETF market.
- To determine whether investor overconfidence influences the trading activities of individual ETFs.
- To examine the effect of investor overconfidence on the volatility of the returns from the South African ETF market.
- To assess the effect of overconfident trading on the volatility of the returns from the South African ETF market before, during, and after the 2008 global financial crisis.

The objectives of this study are achieved by answering the following questions:

- Are investors overconfident when trading in the South African ETF market?
- Does investor overconfidence influence the trading activity of individual ETFs?
- Does investor overconfidence exhibit a positive or negative effect on the volatility of the returns from the South African ETF market?
- Does the volatility of the South African ETF market respond positively or negatively to trading volume induced by investor overconfidence before, during, and after the 2008 global financial crisis?

1.3. Scope and Method of This Study

1.3.1. Scope of This Study

Steyn (2019) discovers that JSE-listed ETFs tracking local benchmarks track their benchmarks more efficiently when compared to JSE-listed ETFs tracking international benchmarks. Therefore, investors trading in JSE-listed ETFs replicating local benchmarks may be more overconfident given the increased tracking ability of ETFs tracking domestic benchmarks. Hence, the current analysis of overconfidence is conducted on the market of South African ETFs tracking domestic benchmarks and the market of South African ETFs tracking international benchmarks. The sample period for the market of ETFs with domestic benchmarks starts in November 2000 while the sample period for the market of ETFs with international benchmarks starts in October 2005. For both markets, the sample periods end in August 2019. The chosen sample periods provide results that are more valid and more accurate because patterns in trading behaviour

can be detected more efficiently through the analysis of longer periods of time. Regarding the data frequency, monthly data observations are used to detect the presence of investor overconfidence because Odean (1999), Statman, *et al.* (2006), and Zaiane and Abaoub (2009) claim that fluctuations in investor overconfidence tend to be more evident over monthly horizons. However, daily data observations are employed to examine the effect of overconfident trading on market volatility since Ledoit and Wolf (2008) argue that GARCH-type models are more efficient in modelling volatility when daily data is used.

1.3.2. Method of This Study

For each market, the relationship between the current turnover of the market and the lagged market returns was investigated using market-wide vector autoregressive (VAR) models and their associated impulse response functions. The market-wide VAR models estimated controlled for alternative explanations of trading volume, viz. market volatility and dispersion. In addition, individual security-level VAR models were estimated with a panel approach to ascertain that the market-wide overconfidence effect is not a direct aggregation of the disposition bias. Specifically, individual security VAR models were estimated to determine whether investor overconfidence influences the trading activities of individual ETFs, thus, confirming or rejecting the market-wide investor overconfidence results.

This study also estimated Exponential Generalised Autoregressive Conditional Heteroskedasticity (EGARCH) models to examine how trading volume induced by investor overconfidence impacts the volatility of the South African ETF market. The 2008 global financial crisis was employed as a structural break and a sub-period analysis was conducted to investigate the effect of investor overconfidence on the volatility of the returns from the South African ETF market before, during, and after the 2008 global financial crisis. Prior to the estimation of the empirical models, preliminary data analysis (that is, analysis of unit root tests and descriptive statistics) was conducted to ensure that the time series are suitable for the estimation. Moreover, after each model was estimated, diagnostics tests (for serial correlation and heteroskedasticity) were performed to confirm the reliability of the results obtained.

1.4. Structure of This Thesis

This thesis includes six chapters which are structured as follows:

➤ **Chapter 1: Introduction**

The introduction chapter discusses the problem statement which motivates the current study. Moreover, the objectives, scope and method of this study are introduced in Chapter 1.

➤ **Chapter 2: Theoretical Foundations**

Chapter 2 builds a framework of concepts relating to behavioural finance, with a specific emphasis on the overconfidence bias. Additionally, a discussion of the various aspects of an ETF as an individual asset class is provided.

➤ **Chapter 3: Review of Empirical Studies**

Chapter 3 reviews existing empirical research conducted on the overconfidence bias. The chapter begins by reviewing existing empirical research on the presence of investor overconfidence in different asset classes (except ETFs). Thereafter, studies examining the effect of overconfidence on market volatility are assessed. This is followed by a review of studies investigating the existence of overconfident trading in ETF markets. Lastly, the chapter reviews evidence of overconfident trading in South Africa.

➤ **Chapter 4: Data and Methodology**

Chapter 4 outlines various aspects relating to the data used in this study, including; the sample of ETFs, sample period, and data type. This is followed by a detailed discussion of how each variable is computed. The latter part of Chapter 4 provides a thorough discussion of the empirical models used to determine the existence and effect of overconfidence in the South African ETF market.

➤ **Chapter 5: Data Analysis and Results**

Chapter 5 reviews the results from the preliminary data analysis and provides a detailed discussion of the findings obtained from the estimated models.

➤ **Chapter 6: Conclusions and Recommendations**

The concluding chapter summarises the results obtained from the analysis in an attempt to answer the research questions posed. Chapter 6 concludes by discussing the implications of this study's findings for investors and policymakers and provides recommendations for future empirical studies.

CHAPTER 2: THEORETICAL FOUNDATIONS

2.1. Overview

Traditional finance theories seek to explain investor behaviour based on the assumption that all investors are rational, and thus, capital markets are efficient. However, several anomalies that exist in financial markets cannot be explained by standard finance theories. Instead, since the advent of Prospect Theory by Kahneman and Tversky (1979), these anomalies have been successfully explained by behavioural finance theories. Behavioural finance theories attempt to explain investor behaviour based on the assumption that not all investors are rational because heuristics and biases influence investment decisions. One of the most prevalent behavioural biases that influence the decisions of investors is the overconfidence bias. The overconfidence bias refers to the tendency of an investor to overestimate his or her ability to perform successfully. Overconfidence can affect the trading decisions of investors trading in any asset class, including traders of ETFs.

Before attempting to understand the behavioural finance approach to investor behaviour, it is important to understand the standard finance approach to investor behaviour. Hence, Chapter 2 begins by reviewing the traditional finance approach to investor behaviour. Thereafter, various theories underpinning the behavioural finance approach to investor behaviour are outlined. This is followed by a detailed discussion of the theoretical foundations of the overconfidence bias, including, but not limited to, the different forms of overconfidence and the relationship between overconfidence and trading volume, market returns, market efficiency, and market volatility. Chapter 2 concludes by describing the theoretical aspects of ETFs since this study focuses on the ETF asset class.

2.2. Theories Relating to Investor Behaviour

2.2.1. Traditional Finance Approach to Investor Behaviour

Traditional finance theories are conceptualised on two assumptions, namely; rationality and market efficiency. However, it is important to note that rationality and market efficiency are not independent because the notion of market efficiency is underpinned by rationality. The Rational Choice Theory assumes that market participants solve problems and formulate strategies using logical processes which are based on the timing and nature of the problem (Mushinada, 2020). Accordingly, the notion of market efficiency assumes that all investors are rational and make well-informed decisions (Fama, 1965). The Efficient Market Hypothesis (EMH) developed by Fama (1970) claims that, in an efficient market, the prices at which securities trade fully incorporate information about all events, both current and expected events. This is because market participants continuously conduct research in an attempt to identify mispriced securities

(Fama, 1970). As a result, any new information is reflected instantaneously in the market prices of securities because of this competition between market participants. Therefore, the concept of market efficiency advocates that movements in share prices are completely random, and thus, forecasting movements in share prices is highly improbable (Tran and Leirvik, 2019). Consequently, in an efficient market, traders cannot exploit security prices to earn abnormal profits.

According to Fama (1970), there are three sufficient conditions to achieve capital market efficiency. Firstly, there should be no transaction costs when trading securities. Secondly, all market participants should have costless access to all available information. Lastly, all investors should have homogeneous beliefs about the implications of current information. Therefore, the presence of heterogeneous beliefs among market participants, transaction costs, and information that is not freely available to all market participants are potential sources of market inefficiency (Fama, 1970).

There exists three forms of market efficiency, namely; weak, semi-strong, and strong (Fama, 1970; Fama, 1991; Tran and Leirvik, 2019). The weak form of EMH proposes that, at any given time, current asset prices reflect all past financial information (Fama, 1991). Therefore, traders cannot generate abnormal returns by using technical analysis (Guney and Komba, 2016). The semi-strong form of EMH suggests that, at any given time, security prices incorporate all public information, including historical information (Fama, 1991). Thus, if the capital market is semi-strong form efficient, technical analysis and fundamental analysis will not be able to generate a portfolio allocation that is more profitable than the case of a random portfolio of assets (Alexakis, Patra and Poshakwale, 2010). However, insider trading can be used to earn abnormal returns when the capital market is semi-strong form efficient (Del Brio, Miguel and Perote, 2002). Finally, the strong form of EMH advocates that, at any given time, an asset's price reflects all available information about the respective financial asset (Vidal-Tomás and Ibañez, 2018). Specifically, the strong form of market efficiency assumes that the prices of securities incorporate all historical, public, and private information. Hence, abnormal returns are not possible in a capital market that is strong form efficient (Fama, 1970).

Based on the assumption of efficient capital markets, the Random Walk Theory put forward by Malkiel (1973) postulates that current asset prices are unrelated to their historical price changes. In other words, the Random Walk Theory asserts that an asset's past price behaviour cannot be used to predict the future market prices of that asset (Ngene, Tah and Darrat, 2017). This implies that, at any given time, the market price of an asset reflects the best estimate of the asset's intrinsic value, based on all available information. Hence, the Random Walk Theory maintains that the arrival of any new information in markets will cause deviations

in an asset's price that is completely random (Malkiel, 1973). Therefore, technical analysis and charting will not assist traders in generating abnormal returns (Masry, 2017). Traditional finance theories were formerly regarded as key explanations of investor and market behaviour. However, researchers have found that standard finance theories do not adequately explain actual market conditions (Joo and Durri, 2015; Kapoor and Prosad, 2017). This is because, standard finance theories are based on incomplete assumptions and are built on foundations of how market participants should behave and not of how they actually behave (Joo and Durri, 2015). As a result, behavioural finance theories have emerged in an attempt to account for investor biases and irrationalities.

2.2.2. Behavioural Finance Approach to Investor Behaviour

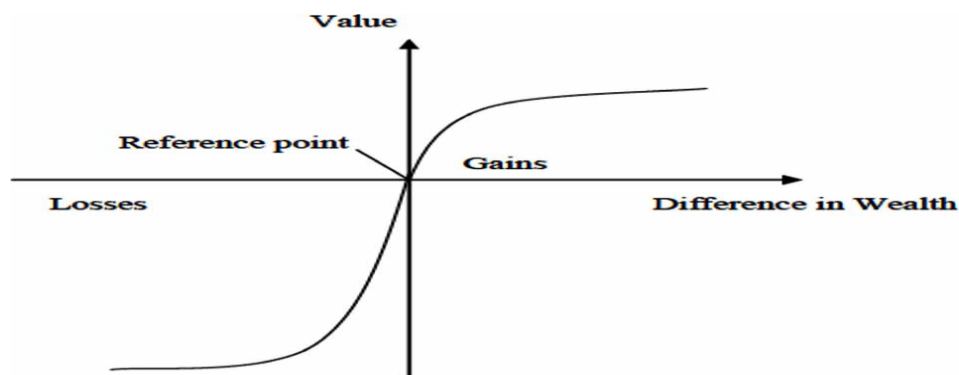
Although standard finance theories are sufficient to make sound financial decisions, they may not be able to explain certain anomalies and disturbances in the market (Kapoor and Prosad, 2017). This is because, in reality, a majority of investors are irrational, and their investment decisions are based on their emotions, beliefs, and state of minds. Hence, behavioural finance theories may be able to explain such disruptions. For instance, Naseer and Tariq (2015) mention that the existence of market underreaction or overreaction, momentum trading and reversal patterns is due to investors' cognitive and psychological behaviours, and thus, can be explained by behavioural decision theories. In broad financial terms, behavioural finance theories examine how the psychology of market participants influence their investment decisions, and as a result, the markets in which they trade (Dickason and Ferreira, 2018). More specifically, behavioural finance refers to the study of how different cognitive and emotional biases lead to irrational investment decisions that result in suboptimal investment choices or stock market anomalies (Raut, Das and Kumar, 2018).

Initial studies on the role of human psychology on decision making can be traced back to the 1700s. The Theory of Moral Sentiments introduced by Adam Smith in 1759 posits that the morality of an individual influences his economic, social, and financial decisions. Moreover, Smith (1759) notes that sentiments, such as; self-interest, sympathy, justice, and virtue, influence the decisions that humans make. This influence of psychology on human decision making implies that individuals may not always be rational. The Theory of Bounded Rationality proposed by Simon (1955) advocates that the rationality of individuals is constrained by their cognitive limitations. Specifically, the theory maintains that the rationality of an individual is constrained by the information that they have at their disposal as well as their computational capacity. Hence, the notion of bounded rationality relaxes the assumption of perfect rationality which, in reality, is often not viable. In 1956, Festinger, Riecken and Schachter put forward the Theory of Cognitive Dissonance. Festinger, *et al.* (1956) point out that individuals seek to maintain cognitions that are in

harmony because individuals begin to feel a mental discomfort when they hold conflicting cognitions at the same time. As a result, these individuals have to alter their beliefs, attitudes and behaviours in order to reduce this feeling of unpleasantness and to restore a balance.

Several behavioural finance theories are developed based on the Prospect Theory introduced by Kahneman and Tversky in 1979. A focal point of the Prospect Theory centres around the concept of loss aversion. Loss aversion refers to a condition in which the pain of a loss is felt with a greater intensity than the happiness of a gain of the same magnitude (Barberis and Huang, 2001). In simple terms, losses are more powerful than equivalent gains. Three key aspects are proposed by Prospect Theory (Kahneman and Tversky, 1979). Firstly, individuals do not have a uniform attitude towards risk (Tversky and Kahneman, 1991). The second aspect of the Prospect Theory is that individuals use a reference point to estimate the value of each prospect (Abdellaoui, Bleichrodt and l'Haridon, 2008). Specifically, an individual's reference point is their status quo or current wealth level which determines the gains or losses generated from each prospect (Kahneman and Tversky, 1979). The third aspect of the Prospect Theory is that individuals are loss averse, such that, individuals seek to avoid losses more than they seek gains (Schmidt and Zank, 2012). The three properties proposed by the Prospect Theory lead to the development of an asymmetric 'S' shaped value function that is concave for gains and convex for losses and is steeper in the domain of losses (List, 2004). Figure 2.1 illustrates this 'S' shaped value function which provides an approximation of the value that individuals assign to their gains or losses. Notably, the value function of the Prospect Theory replaces the utility function¹ of the standard Expected Utility Theory (EUT henceforth).

Figure 2.1: The Value Function



Source: Kahneman and Tversky (1979)

¹ The EUT suggests that there is a symmetrical response to losses and gains, whereas the Prospect Theory asserts that losses loom larger than gains (Tsaur, 2013).

It is evident from the aforementioned literature that the modification of traditional finance theories to account for behavioural aspects began quite early. The Prospect Theory, for example, provides an alternative to the conventional EUT. Further examples of the incorporation of behavioural aspects into standard finance theories include the Behavioural Asset Pricing Model (henceforth BAPM) and the Behavioural Portfolio Theory (BPT henceforth). Introduced by Shefrin and Statman (1994), the BAPM serves as an alternative to the traditional Capital Asset Pricing Model (hereafter CAPM). Shefrin and Statman (1994) derive the BAPM based on a market in which informational traders interact with noise traders. Informational traders are rational traders who are free of cognitive errors, while noise traders are traders who commit cognitive errors (Ramiah and Davidson, 2007). As a result, noise traders cause the prices of securities to deviate from their efficient levels. Therefore, the BAPM proposes that a security's equilibrium rate of return should be computed by taking into account the effect of noise traders (Shefrin and Statman, 1994).

Shefrin and Statman (2000) also formulated the BPT which is regarded as an alternative to the traditional Markowitz (1952) Portfolio Theory (henceforth MPT). According to Kapoor and Prosad (2017), the MPT prescribes that investors construct an optimal portfolio by analysing the covariances between securities as well as by optimising their risk-return trade-off, while the BPT suggests that investors construct portfolios as layered pyramids of assets in which each layer is related to their risk attitude and specific goals. Moreover, the BPT maintains that investors build portfolios by taking into account their expected wealth, desire for security, as well as the probabilities of obtaining their goals, and therefore, may overlook the covariances between the securities in the portfolio (Shefrin and Statman, 2000).

These theories of behavioural finance attempt to capture the influence of behavioural biases on the decisions made by investors. Shefrin (2000) categorises these biases into two groups, namely; frame-dependent biases and heuristic-driven biases. The manner in which a scenario is presented will influence the final decision made by an individual, and is, therefore, referred to as frame-dependent biases (Shefrin, 2000). Heuristic-driven biases relate to the rules of thumb that individuals use to ease their decision-making processes (Frankfurter, McGoun and Allen, 2004). According to Rajdev and Ranninga (2016), overconfidence is a heuristic-driven bias. The literature on the influence of behavioural biases on investment decision-making processes is extensive and includes a significant number of different biases. However, this study is limited to only the overconfidence bias since it is one of the most common traits among investors (Zaiane and Abaoub, 2009; Jlassi, *et al.*, 2014; Yu, 2014; Baker, *et al.*, 2019; Trejos, van Deemen, Rodríguez and Gómez, 2019). The next section discusses the theoretical foundations of the overconfidence bias.

2.3. Theoretical Foundations of the Overconfidence Bias

The broad use of the term overconfidence is associated with a cognitive bias that occurs when the level of subjective certainty is greater than the level of objective accuracy (Deaves, Lei and Schröder, 2019). However, overconfidence is generally classified into three categories. These three categories are discussed in Section 2.3.1, after which the impact of overconfidence on market characteristics is discussed in Sections 2.3.2 – 2.3.7.

2.3.1. Categories of Overconfidence

According to Moore and Healy (2008), overconfidence can be categorised into three different forms, namely; overestimation, overplacement, and overprecision. These three forms of overconfidence are discussed in Sections 2.3.1.1 to 2.3.1.3 that follow.

2.3.1.1. Overestimation

Moore and Schatz (2017) use the term overestimation to refer to the tendency of an individual to overestimate his or her own abilities, degree of control, or probability of success. Overestimation may be caused by an illusion of control or by the planning fallacy. Langer (1975) proposed the theory of illusion of control. According to Langer (1975), the illusion of control occurs when individuals have unwarranted beliefs that they have a higher chance of personal success. Burger (1991) maintains that people find it difficult to distinguish between skill-determined (controllable) and chance (non-controllable) events. The extent to which individuals perceive an event as controllable depends on factors, such as; choice, involvement, familiarity, and competition (Langer, 1975; Fellner, 2009). Such factors may be independent of the actual contingency, and therefore, individuals may believe that they can control events governed by chance (Blanco, Matute and Vadillo, 2011). This is because people gain satisfaction from being able to control a seemingly uncontrollable event (Sloof and von Siemens, 2017). Moreover, control allows individuals to feel competent and allows them to avoid the negative consequences associated with being perceived as having no control.

Buehler, Griffin and Ross (1994) posit that the planning fallacy occurs when individuals overestimate how quickly they can complete tasks, especially long and complicated tasks. In other words, individuals are too optimistic about their completion times because they underestimate the time that it will take to complete tasks. Pezzo, Pezzo and Stone (2006) argue that the planning fallacy is caused by individuals disregarding information about previous task lengths, and instead, focusing on plan-based scenarios.

2.3.1.2. Overplacement

The term overplacement relates to the propensity of individuals to believe that they are better than others (Pikulina, Renneboog and Tobler, 2017). Weinstein (1980) introduced the notion of unrealistic optimism to refer to the tendency of individuals to assume that they are not vulnerable. Moreover, these individuals assume that they are not victims of misfortune, but instead, other individuals are victims of misfortune (Shepperd, Klein, Waters and Weinstein, 2013). These assumptions imply that individuals possess an error in judgement by expecting that their future is more favourable than the future of other individuals, and therefore, this outlook can be regarded as unrealistic optimism (Weinstein, 1980).

The ‘better than average effect’ was proposed by Alicke (1985). According to Alicke (1985), individuals appraise themselves more favourably than they appraise other individuals. This is because people feel good about themselves when they perceive themselves as above average. Hence, the better than average effect may be driven by an individual’s self-enhancement needs (Wojcik and Ditto, 2014). However, the better than average effect may also be motivated by egocentrism, informational differences, focalism and naive realism (Zell and Alicke, 2011). One possible implication of unrealistic optimism, overestimated perceptions of control, and overstated positive self-appraisals is that these positive illusions help to foster mental health (Taylor and Brown, 1988). In other words, these positive illusions enhance an individual’s ability to be productive, happy, and content, as well as their ability to care about themselves and others.

2.3.1.3. Overprecision

Overprecision is defined by Wolfe, Reyna and Smith (2018) as the excessive certainty that the individual knows the truth or correct answer. The precision of decision making is measured using confidence intervals (Moore, Carter and Yang, 2015). Alpert and Raiffa (1982) argue that individuals tend to overstate the level of accuracy of their knowledge because their subjective probability distributions are too narrow, and therefore, these individuals omit several other possibilities from the event space. This being so, the narrow confidence intervals imply that individuals think that their knowledge is more accurate than it actually is (Alpert and Raiffa, 1982). The next section discusses how overconfidence, which could take any of the three forms, impacts the different aspects of market quality².

² Market quality refers to the ability of markets to achieve its objectives of price discovery and liquidity (O’Hara and Ye, 2011).

2.3.2. Overconfidence and Trading Volume

One of the earliest studies examining the effect of overconfidence on financial markets was undertaken by Odean in 1998. Odean (1998) proposes that trading volume rises when price takers, insiders, or market makers are overconfident. Trading volume increases as overconfidence increases when traders are price takers because these traders calculate their posterior beliefs³ by overweighting their private signals and undervaluing public signals or the signals of other traders (Barber and Odean, 2001). This leads to posterior beliefs that are more dispersed; subsequently, leading to an increase in trading activities. Similarly, when the trader is an insider, an increase in overconfidence leads to an increase in trading volume. An overconfident insider assumes that his private signals are more reliable than it actually is (Odean, 1998). Accordingly, the overconfident insider calculates his posterior belief of the final asset value by overweighting his private signals which results in a belief that is far from what it should be (Gervais and Odean, 2001). Based on this posterior expectation, the overconfident insider trades more frequently than is optimal, resulting in an increase in the trading volume.

Kamesaka, Nofsinger and Kawakita (2003) suggest that overconfidence may be linked to positive feedback trading (that is, trading in the direction of past patterns). This is because, if the credibility of an investor's trading patterns is confirmed by certain events, then that investor will feel overtly proud that his chosen strategy was the correct one (Odean, 1999). In addition, if the trend in prices continues, the investor may believe that he was able to predict this trajectory (Kamesaka, *et al.*, 2003). These biases can elicit aggressive trading, which can reinforce existing positive feedback trading and overconfidence. Overall, Odean (1998), Benos (1998), Kyle and Wang (1997), Statman, *et al.* (2006), Glaser and Weber (2007) and Alsabban and Alarfaj (2020) acknowledge that the higher the degree of an individual's overconfidence, the higher his or her trading volume.

2.3.3. Overconfidence and Market Return

According to Gervais and Odean (2001), a trader learns about their own abilities by observing their past successes and failures. However, the trader takes too much credit for his successes when evaluating his abilities. As such, when a trader is successful, the trader's belief about their abilities moves upward because the trader attributes too much of his success to his own abilities and disregards external factors (Gervais and Odean, 2001; Glaser and Weber, 2009). However, the same individuals tend to believe that their failures are caused by external factors and not their personal inabilities. This propensity of individuals to associate

³ A posterior belief refers to a prior belief that has been updated based on new information (Cook and Lewandowsky, 2016).

their successes with their own abilities is commonly referred to as the self-attribution bias (Glaser and Weber, 2009).

A trader's confidence in their own abilities changes with their successes and failures (Kim and Nofsinger, 2007). This being so, an investor is more overconfident when their successes are higher than usual (Chou and Wang, 2011). Gervais and Odean (2001) posit that a trader's measure of success is their returns. Therefore, the overconfidence hypothesis proposed by Gervais and Odean (2001) asserts that traders who associate market returns with their own abilities become more overconfident after periods of market gains, and as a result, trade more frequently in periods after market gains. As such, periods of market gains tend to be followed by periods of increased market trading. Thus, Gervais and Odean (2001), Statman, *et al.* (2006) and Chen and Sabherwal (2019) argue that the presence of a positive lead-lag relationship between trading volume and market return is evidence of investors' overconfidence. Given that most stocks generate positive returns during bull markets, Chuang and Susmel (2011) argue that investor overconfidence is more pronounced during bull markets in comparison to bear markets in which overconfidence levels tend to decrease.

2.3.4. Overconfidence and the Disposition Effect

A positive lead-lag relationship between current trading volume and historical returns may also be associated with the disposition effect. Barber and Odean (1999) mention that, in addition to the excessive trading by overconfident investors, a common mistake by investors is to sell winning investments and hold onto losing investments. This tendency to hold onto losing positions whilst selling winning positions is referred to as the disposition effect (Shefrin and Statman, 1985).

Statman, *et al.* (2006) note that there are two key differences between the overconfidence bias and the disposition effect. Firstly, the disposition effect provides an explanation for only the sell-side of the trade since the disposition effect only materialises when traders want to sell a security (Chou and Wang, 2011). Specifically, when a stock is experiencing gains, investors will sell the stock to realise a gain on paper, whereas, during periods of losses, investors will be resistant to sell the security. Contrarily, the overconfidence bias affects both the buy-side and the sell-side of a given trade (Statman, *et al.*, 2006; Chou and Wang, 2011). Specifically, two investors with differing opinions and inappropriately tight errors bounds will be able to conduct a trade without requiring other rational information traders.

The second difference between the two biases is that the disposition bias relates to the attitude of an investor towards a particular security or securities in his portfolio (Prosad, Kapoor, Sengupta and Roychoudhary,

2017). In contrast, the overconfidence bias is indicative of an investor's attitude about the general market rather than a specific security or securities (Statman, *et al.*, 2006). This is because, if traders form too high estimates of their personal ability to raise their wealth by actively trading, these traders are likely to maintain this overestimation about securities in general instead of just a specific security in their portfolio (Prosad, *et al.*, 2017). Therefore, to ensure that the overconfidence effect observed is not an aggregation of the disposition effect, Statman, *et al.* (2006) investigate how the current trading activity of individual securities respond to historical market returns. Notably, a positive lead-lag relationship between individual security trading volume and market returns indicates the existence of overconfidence at the individual security level (Statman, *et al.*, 2006).

2.3.5. Overconfidence and Market Efficiency

Overconfidence can either improve or deteriorate market efficiency, however, this depends on the manner in which information is disseminated to markets. On the one hand, when information is distributed publicly and then interpreted differently by traders, overconfident traders overweight the aggregate signal (Odean, 1998). These heterogeneous biased expectations increase the deviation of an asset's price from that of its fundamental value, subsequently, worsening the quality of prices (Odean, 1998). On the other hand, when information is held only by an insider, the quality of prices improves (Yeh and Yang, 2014). This is because, through aggressive trading, an overconfident insider reveals more of his private information than he otherwise would, thus, enabling market makers to set prices that are closer to the asset's fundamental value (Odean, 1998). However, this gain in efficiency is short-lived if the insider's information becomes public soon after he trades. Moreover, García, *et al.* (2007) and Ko and Huang (2007) argue that, under certain conditions, overconfident trading could enhance price quality. In particular, overconfident traders believe that they can generate higher than average returns, and thus, invest resources in obtaining information about financial securities. As a result, overconfident investors can introduce information to markets, and this information could have a greater impact on price quality than the mispricing caused by overconfidence (Shah, Ahmad and Mahmood, 2018).

Another adverse consequence of overconfident trading is that it may lead to return predictability. This is because, overconfident traders overreact to subjective and less relevant information but underreact to abstract and highly relevant information (Daniel and Hirshleifer, 2015). Odean (1998) asserts that price changes display negative serial correlations when traders overvalue new information, whilst price changes display positive serial correlations when traders undervalue new information. This being so, undervaluing new information leads to price trends, whereas overvaluing new information leads to price reversals (Daniel and Hirshleifer, 2015). Nevertheless, return predictability is inconsistent with the weak form EMH as they

reflect historical information which should be priced in a market that is weak form efficient (Robinson, 2005).

Whilst there are several drawbacks of overconfident trading, overconfident trading may bring about some benefits to financial markets. For instance, a key benefit of overconfident trading is that it increases market depth by forcing market makers to allow for additional trading in response to the aggressive trading by overconfident traders (Benos, 1998). This is further supported by Liu (2015) who argues that excessive trading caused by overconfidence improves the liquidity of financial markets.

2.3.6. Overconfidence and Market Volatility

Overconfident trading generates excess volatility in financial markets. The reason for this is that by overweighting their own private signals relative to the general consensus, overconfident traders drive asset prices further from its intrinsic values (Yeh and Yang, 2011). Therefore, overconfident trading generates excess volatility by distorting asset prices. Similarly, Benos (1998) argues that the presence of overconfident traders in markets leads to higher volatility and larger transaction volumes. Contrarily, rational traders decrease trading volume and volatility and increase the efficiency of prices (Odean, 1998).

Daniel, *et al.* (1998) argue that, due to the self-attribution bias, traders become overconfident when their private signals coincide with public signals. On the contrary, traders ignore public signals when public signals contradict their private signals, and therefore, the price remains unchanged. This being so, a sequence of public signals that confirm private signals will cause the price to move in the direction of the private signals (Daniel, *et al.*, 1998). Despite a large adjustment in the price, an immediate response could still represent an underreaction whilst later adjustments in the price could represent a delayed overreaction (Daniel, *et al.*, 1998). However, these price changes are not warranted by fundamentals, and therefore, represent a bubble (Michailova and Schmidt, 2016). An asset price bubble is caused when asset prices increase far above their fundamental values (Gürkaynak, 2008; Janssen, Füllbrunn and Weitzel, 2019). Such deviations are not justified by asset fundamentals, and when investors are not willing to purchase the asset at elevated prices, a massive sell-off occurs, causing the bubble to burst (Watanapalachaikul and Islam, 2007).

Scheinkman and Xiong (2003) also argue that overconfidence leads to price bubbles. In particular, when information is disseminated to markets, traders interpret that information differently, and overconfident traders overestimate the informativeness of the different signals (Scheinkman and Xiong, 2003). This overestimation creates heterogeneous expectations among traders regarding an asset's true value, and as a

result, some traders are relatively more optimistic than others. Investors with heterogeneous beliefs engage in speculative trading, causing the price of an asset to deviate from its fundamental value, leading to a price bubble (Bidian, 2015). Hence, Scheinkman and Xiong (2003) claim that a price bubble is a consequence of heterogeneous beliefs among traders. Similarly, Jlassi, *et al.* (2014) assert that the overconfidence bias creates an autocorrelation in investors' biased expectations about the intrinsic value of assets. As a consequence, investors' overconfidence generates price distortions which could lead to the formation of price bubbles.

2.3.7. Overconfidence and Investor Wealth

Daniel, *et al.* (1998) acknowledge that when investors overweight their own abilities, they underestimate their forecast errors. Moreover, investors tend to be more overconfident about the information that they personally generate, which implies that they overestimate the accuracy of their own private information relative to publicly available information. Accordingly, overconfident investors are less likely to follow the direction of the market, and instead, may choose to follow contrarian strategies (Daniel, *et al.*, 1998). As a result, overconfident investors may trade even when the gains generated through trading are less than their trading costs (Odean, 1999). Therefore, the excessive trading by overconfident individual investors may be hazardous to investors' wealth (Kleine, Wagner, and Weller, 2016).

Chuang and Lee (2006) argue that overconfident investors trade more in riskier securities because these overconfident investors underreact to risks associated with their investments, subsequently, resulting in lower expected utilities from their investments. In fact, since overconfident investors underestimate their downside risk, overconfident trading may reduce the wealth of overconfident traders (Barber and Odean, 1999; Odean, 1999; Barber and Odean, 2000) However, Odean (1998) notes that a risk-averse trader who is overconfident may opt for a riskier than normal portfolio, and thus, may be compensated for bearing higher risk with higher expected returns. Therefore, Odean (1998) argues that the profits from bearing higher risk may overweight the losses associated with biased judgements. Moreover, Benos (1998) posits that the aggressive trading of overconfident investors may create first-mover advantages. Therefore, it is possible for overconfident investors to generate higher returns than rational traders.

2.3.8. Factors Affecting Overconfidence

Trehan and Sinha (2011) acknowledge that the level of overconfidence displayed by an investor is influenced by demographic factors, like gender, education, and income, as well as trading behaviour, such as trading experience, trading success, and trading frequency. Barber and Odean (2001) argue that overconfidence is more pronounced in males than females, especially in male-dominated areas, such as

finance. Consequently, men tend to trade more excessively than women. However, since men are more overconfident than women, men are more likely to generate lower returns than women (Barber and Odean, 2001). Gervais and Odean (2001) further argue that a trader's level of experience influences the degree of overconfidence. Specifically, Gervais and Odean (2001) maintain that overconfidence decreases as traders become more experienced. According to Gervais and Odean (2001), traders gain experience by participating in financial markets, and thus, the amount of experience is dependent on the time spent participating in the financial market as well as the intensity of participation. Overall, the framework suggests that overconfidence is more prominent among inexperienced traders relative to experienced traders because traders develop better self-assessments with more experience (Gervais and Odean, 2001).

Müller and Weber (2010) posit that overconfident investors are less likely to invest in ETFs or passive index funds because these unsophisticated (overconfident) investors rely on financial advisors who have no interest to recommend passive funds. Moreover, investors with high levels of better-than-average thinking in terms of their security-picking abilities believe that they have the skills to identify funds that can outperform their benchmark (Müller and Weber, 2010). Accordingly, investors who overestimate their security-picking skills are more likely to invest in funds that are actively managed. Prior to investigating the presence of investor overconfidence in ETF markets, it is crucial to understand the dynamics of ETFs. Hence, the next section elaborates on the dynamics of ETFs.

2.4. Theoretical Aspects of ETFs.

An ETF is a pooled investment vehicle that consists of a combination of securities which are selected in an attempt to track the performance of a specific benchmark index (Charteris, 2013; Da and Shive, 2018). As highlighted by Naumenko and Chystiakova (2015), ETFs can follow one of the two replication schemes; physical or synthetic. On the one hand, physical ETFs track the underlying index by owning the index constituents (Naumenko and Chystiakova, 2015). On the other hand, synthetic ETFs attempt to replicate the benchmark index by using derivatives as opposed to holding the physical assets contained in the index (Naumenko and Chystiakova, 2015). Additionally, ETFs are either actively or passively managed (Rompotis, 2013). Passively managed ETFs differ from actively managed ETFs in that the former is structured to replicate a specific benchmark while the latter seeks to outperform the market (Rompotis, 2013). This being so, actively managed ETFs are typically designed to mirror an accomplished investment manager's stock choices or to pursue a specific investment target in order to generate abnormal returns for investors. Remarkably, a majority of ETFs aim to track their benchmarks and, as such, are passively managed.

Similar to stocks, ETFs can be purchased or sold through a broker (Naumenko and Chystiakova, 2015). Furthermore, ETFs can be sold short or purchased with a margin, just like stocks (Mohamad, Jaafar and Goddard, 2016). In addition, there are several similarities between ETFs and index funds. ETFs and mutual funds are similar in structure as their investment strategy involves attempting to replicate the risk and return of a specific benchmark index (Strydom, Charteris and McCullough, 2015). Furthermore, in South Africa, both ETFs and index funds operate under the Collective Investment Schemes Control Act, number 45 of 2002 (Colegrave, 2008).

The key distinction between ETFs and index funds is that, while ETFs can be traded throughout the day, index funds can only be traded once per day (Farinella and Kubicki, 2018). Another important difference is that index funds trade at their closing net asset values (NAV hereafter) whereas ETFs trade on exchanges at prices determined by the market (Venkataraman and Venkatesan, 2016). Notably, ETFs have the ability to create more ETF units as their demand increases, and therefore, ETFs tend to trade at prices close to their NAV (Venkataraman and Venkatesan, 2016). However, the degree to which ETF prices are efficient is subject to any limits to arbitrage, such as transaction costs (Petajisto, 2017). Collectively, Deev and Linnertová (2014), Aldridge (2016), and Da and Shive (2018) argue that ETFs signify an increasingly popular saving and investment mechanism for investors since ETFs represent a cheaper and more convenient asset class. Hence, the ensuing section describes some of the characteristics of ETFs.

2.4.1. ETF Characteristics

2.4.1.1. Creation and Redemption Process

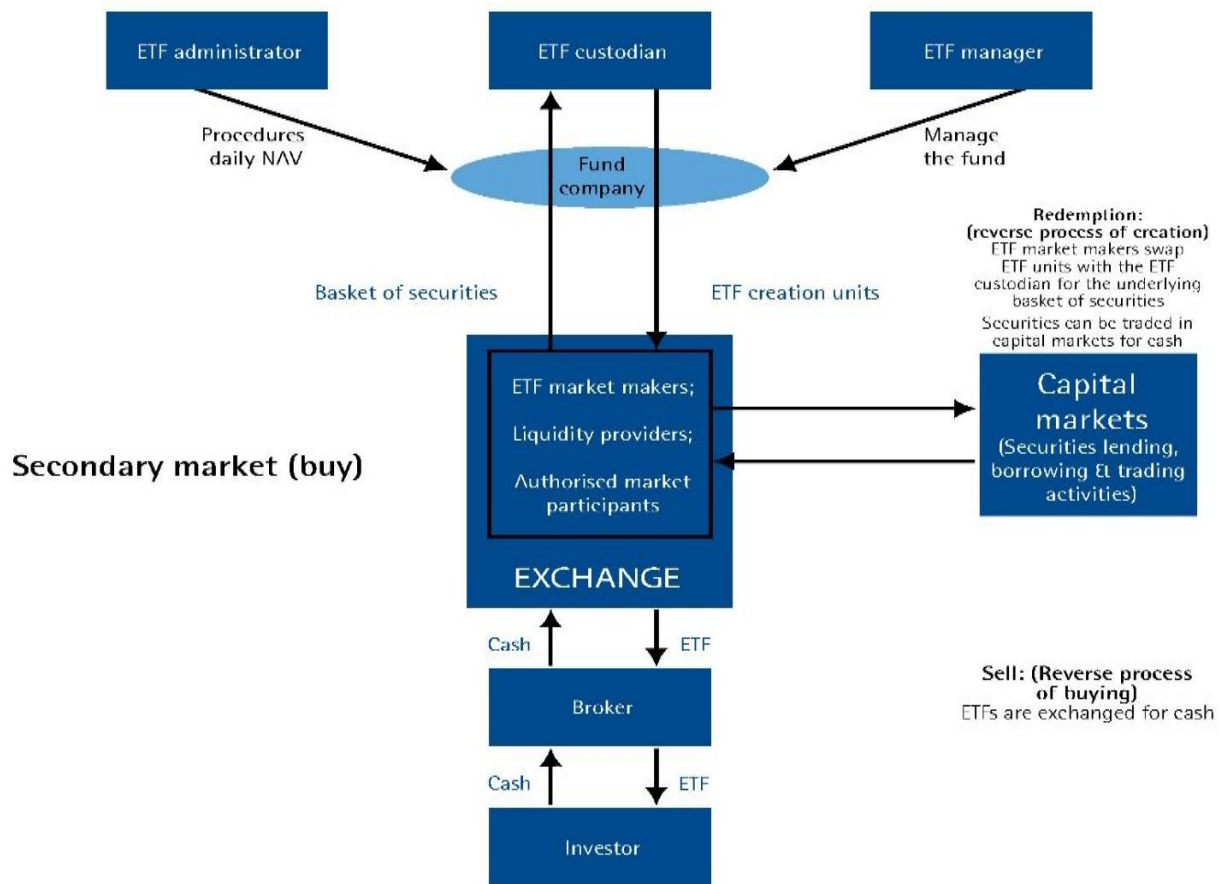
ETFs are able to closely track their benchmark index due to their unique creation and redemption processes. Specifically, an ETF is supported by a management company that manages the fund, a custodian that holds the ETF's assets, and an administrator who produces the daily NAV (Amenc, Goltz and Le Sourd, 2017). In comparison to conventional mutual funds, ETF shares are created by authorised market participants, who purchase blocks of securities, and then, deposit the blocks of securities with the ETF (Staer, 2017). These blocks of securities are securities constituted in the benchmark index and are purchased from the exchange market. The authorised market participants then receive a specified amount of ETF shares for their deposit. Some or all of created ETF units are then sold on the exchange (Petajisto, 2017). Thereafter, investors can buy or sell ETF shares through an intermediary. ETFs can be sold short, and therefore, traders may short sell ETFs to hedge their risk exposures (Ben-David, *et al.*, 2018).

The market price of ETFs is determined by their demand and supply mechanisms since ETF shares are traded throughout the day on exchanges (Lettau and Madhavan, 2018). Thus, the motivation to create or

redeem ETF units stems from whether ETFs are trading at a discount or premium to the NAV of their underlying securities. This is because, if ETF prices deviate from its NAV, traders can capitalise on these differences and generate arbitrage profits (Ben-David, *et al.*, 2018). If ETFs are overvalued, arbitrageurs purchase the underlying securities, exchange them for creation units, and then sell the created ETF units in the secondary market (Ben-David, *et al.*, 2018). Consequently, ETF prices face downward pressure. On the other hand, if ETFs are undervalued, arbitrageurs purchase ETF units and then exchange them for the underlying securities at the custodial bank (Ben-David, *et al.*, 2018). As a result, ETF prices face upward pressure. Therefore, any mispricing of the ETF tends to be short-lived, and ETF prices tend to remain close to their NAV. This creation and redemption process of an ETF is illustrated in Figure 2.2. It is important to note that this creation and redemption process applies only for passive ETFs and not active ETFs.

Figure 2.2: Creation and Redemption Process of an ETF

Primary market (creation)



Source: Amenc, Goltz, and Le Sourd (2017)

2.4.1.2. Diversification

Each ETF attempts to track a different category, including, but not limited to, a specific asset class, sector, fundamental characteristic, or emerging market (Madura and Ngo, 2008; Guedj and McCann, 2011; Blitz and Huij, 2012). Through the use of ETFs, traders can gain cost-efficient access to various asset classes, like commodities and corporate bonds that were previously accessible to only institutional investors or investors with a high net worth (Padungsaksawasdi and Daigler, 2014). Additionally, retail investors can now utilise ETFs to gain exposure to different markets, such as real estate markets and emerging market assets (Rompotis, 2014). Huang and Lin (2011) found that internationally diversified portfolios can be established by indirectly investing in foreign markets through ETFs. Therefore, it is evident that the introduction of various types of ETFs facilitate a low-cost method for investors to diversify their portfolios.

2.4.1.3. Low Cost

One of the key advantages of investing in ETFs is that they provide the same benefits offered by index funds but at a lower cost. A majority of ETFs are passive investment vehicles, and therefore, incur lower management costs (Blitz and Huij, 2012; Neves, *et al.*, 2019). Furthermore, Gastineau (2001) notes that the cost advantage of ETFs is derived from economies of scale since they tend to be very large funds. In addition, ETFs do not incur transfer agency costs as with mutual funds (Gastineau, 2001). Another source of cost reduction stems from pooling securities together, and therefore, ETFs are able to reduce transaction costs and lower information asymmetry (Madhavan, 2014). ETFs are also more tax-efficient than mutual funds due to the flexibility of the ETF creation and redemption process (Bernstein, 2002; Hill, 2016). However, according to Bernstein (2002), the cost advantages of ETFs are weakened when investors trade too frequently.

2.4.1.4. Trading Flexibility

Unlike mutual funds, ETFs are traded on stock exchanges. As such, ETFs offer significant trading flexibility because ETFs can be traded throughout the day whereas mutual funds can only be traded at the end of the day (Lettau and Madhavan, 2018). Moreover, ETFs can be bought or sold with a market order, limit order, stop-loss order, or on a margin (Rompotis, 2009). The intraday trading of ETFs allows traders to get into or out of a position at any time throughout the day (Rompotis, 2009). Therefore, ETFs may become a vehicle for noise traders, and thus, overconfident investors.

2.4.1.5. ETF Market Impacts

Hamm (2014) argues that uninformed investors tend to migrate away from investing in individual stocks, and instead, invest in diversified ETFs. Consequently, securities that form part of diversified ETFs experience a decrease in liquidity. This phenomenon is further supported by Israeli, Lee and Sridharan (2017). Israeli, *et al.* (2017) mention that the unique creation and redemption process of ETFs allow them to be much more liquid than their constituent securities, and therefore, ETFs are attractive instruments for traders and speculators. On the contrary, Glosten, Nallareddy, and Zou (2016) argue that securities constituted in ETFs experience an increase in information efficiency because securities that are part of ETFs reflect information more quickly. Similarly, Madhavan and Sobczyk (2016) document that ETFs improve financial market information. However, Ben-David, *et al.* (2018) argue that the formation of ETFs leads to an increase in the volatility of the underlying securities because of the arbitrage activities between ETFs and their constituent securities. Moreover, according to Ben-David, *et al.* (2018), liquidity shocks in ETF markets are transmitted to underlying securities via arbitrage trades, thus, increasing the non-fundamental volatility of constituent securities. While there are a few drawbacks associated with the formation of ETFs, their popularity as an individual asset class has soared in recent years.

2.4.2. The Popularity of ETFs

Merton (1987) was one of the earliest researchers to investigate the popularity of basket securities. The Investors Recognition Theory put forward by Merton (1987) suggests that investors can reduce their fixed set-up costs by pooling their assets together under one portfolio manager. Additionally, investors could enjoy reduced transaction and information costs and better investment strategies by obtaining a highly-skilled portfolio manager (Merton, 1987). Overall, the Investors Recognition Theory suggests that ETFs provide a cheaper and more convenient trading instrument and requires minimal expertise in comparison to investing in individual securities (Merton, 1987). Furthermore, according to Fremault (1991), which is supported by Li and Zhu (2018), the introduction of basket instruments in imperfectly integrated markets will decrease the information asymmetry across markets as well as the costs associated with arbitrage activities. This is because basket securities weaken the constraints that prevent arbitrageurs from constructing profitable portfolios. Consequently, the increase in arbitrage activity leads to an increase in the price efficiency and liquidity of basket instruments.

Based on the assumption of perfectly integrated markets, the behaviour of investors who can opt to invest in basket instruments or their individual constituents is modelled by Subrahmanyam (1991) and Gorton and Pennacchi (1993). Subrahmanyam (1991) report that adverse selection costs are diversified away in basket instruments, and therefore, basket securities serve as the lowest-cost market for discretionary liquidity

traders. Likewise, Gorton and Pennacchi (1993) claim that investors are migrating away from individual securities towards investing in basket securities since basket instruments enable liquidity traders to establish portfolios at the lowest cost. Moreover, a majority of ETFs follow a passive investment strategy, and therefore, incur lower management costs (Blitz and Huij, 2012). Hence, the lower costs associated with passive investment funds have contributed to the global shift away from actively managed funds to passively managed funds (Steyn, 2019).

2.5. Summary

The current chapter outlined various aspects underpinning behavioural finance, with a specific emphasis on the overconfidence bias. This was followed by a discussion of the various aspects of ETFs as an individual asset class. In summary, whilst traditional finance theories support the notion that investors are always rational, behavioural finance theories argue that investors are not always rational since they make decisions based on their emotions, beliefs, and biases. A prominent behavioural bias is the overconfidence bias which occurs when an individual's subjective confidence level is greater than the reasonable objective level, for any given decision. Investor overconfidence may be caused by overestimation, overplacement, and overprecision. Nevertheless, overconfidence may lead to excessive trading volumes, market volatility and price distortions. Overconfident investors, who rely on financial advisors, are less likely to invest in ETFs that are passively managed. However, ETFs may be passively or actively managed and may follow a physical or synthetic replication scheme. Nonetheless, the unique creation and redemption process of ETFs allows them to trade throughout the day and with minimal costs, and as a result, their popularity has soared in recent times.

CHAPTER 3: REVIEW OF EMPIRICAL STUDIES

3.1. Overview

Whilst the previous chapter discussed the theoretical foundations of the overconfidence bias, the current chapter reviews existing empirical research conducted on the overconfidence bias. However, existing empirical research evaluating the presence of overconfident trading in ETF markets is quite limited. To overcome this limitation, Chapter 3 begins by reviewing empirical evidence in different asset classes. Notably, the ensuing discussion is fourfold. Section 3.2 reviews existing empirical research that investigates the presence of overconfident trading in different asset markets (except ETFs). For each study under review, the methodologies used in the respective study are outlined in an attempt to develop an appropriate empirical method for research questions one and two.

Section 3.3 provides a discussion of studies examining the effect of overconfident trading on market return volatility. The empirical methods used by these studies are outlined to assist with the development of an appropriate method for research questions three and four. Additionally, Section 3.4 reviews empirical evidence of the prevalence of overconfident trading in ETF markets, while Section 3.5 reviews existing empirical research that can be used to provide an indication of overconfident trading by South African investors.

3.2. Overconfidence in Other Asset Classes

The following section reviews empirical evidence of the presence of overconfident trading in different asset classes, except ETFs. It is important to note that the majority of overconfidence research has been conducted using data from stock markets, and thus, there is limited evidence relating to influence of the overconfidence bias in other asset classes, such as; commodities, mutual funds, and Real Estate Investment Trusts (REITs).

3.2.1. Statman, *et al.* (2006)

One of the most prominent empirical methodologies employed in studies investigating the presence of investor overconfidence is the Statman, *et al.* (2006) model⁴. Statman, *et al.* (2006) test Odean's (1998) proposition that the high trading volume present in equity markets is caused by overconfident investors. Specifically, the study investigates whether the degree of overconfidence, and therefore, trading volume, fluctuates with lagged returns. The test for overconfidence is conducted on all common stocks trading on the New York Stock Exchange (NYSE), excluding closed-end funds, REITs and American Depository Receipts (ADRs) and the observation period ranges from August 1962 until December 2002.

⁴ This study uses the Statman, *et al.* (2006) model to answer research questions one and two.

To investigate the magnitude and perseverance of the association between current trading volume and lagged returns, Statman, *et al.* (2006) make use of VAR models and their associated impulse response functions. The authors note that, apart from the overconfidence explanation, there may be alternative explanations for the lead-lag relationship between volume and returns. For instance, large changes in asset prices may encourage investors to rebalance their portfolios, which, as a consequence, induces trading activities. Similarly, when market participants interpret informational events differently, a correlation between concurrent trading volume and return volatility may develop. Therefore, the VAR model employed by Statman, *et al.* (2006) includes the dispersion and volatility of returns as exogenous variables in order to account other potential explanations of trading activity, such as heterogeneous beliefs and portfolio rebalancing.

Accordingly, the interaction between the trading volume and returns of the market is examined by including market volatility and dispersion as exogenous variables in their VAR model. Market volatility (*msig*) is included to account for the contemporaneous volume-volatility relationship, while dispersion (*disp*) is included to account for any trading activity that may be related to portfolio rebalancing. Statman, *et al.* (2006) compute these variables as follows:

$$msig_t^2 = \sum_{\tau=1}^T r_\tau^2 + 2 \sum_{\tau=1}^T r_\tau r_{\tau-1} \quad (3.1)$$

$$disp_t = \sqrt{\sum_{\tau=1}^T \left[\frac{(r_\tau - \mu)^2}{T} \right]} \quad (3.2)$$

$msig_t^2$ represents month t 's market volatility and $disp_t$ represents month t 's market dispersion. r_τ denotes the market return on day τ , μ signifies the average daily market return for month t and T denotes the number of trading days in month t .

The endogenous variables in the VAR model employed by Statman, *et al.* (2006) are market turnover⁵ (*mturn*) and market return (*mret*). Hence, the market version of the VAR model follows the following model specification:

$$\begin{bmatrix} mturn_t \\ mret_t \end{bmatrix} = \begin{bmatrix} \alpha_{mturn} \\ \alpha_{mret} \end{bmatrix} + \sum_{k=1}^K A_k \begin{bmatrix} mturn_{t-k} \\ mret_{t-k} \end{bmatrix} + \sum_{l=0}^L B_l \begin{bmatrix} msig_{t-l} \\ disp_{t-l} \end{bmatrix} + \begin{bmatrix} e_{mturn,t} \\ e_{mret,t} \end{bmatrix} \quad (3.3)$$

Additionally, Statman, *et al.* (2006) examine individual security volume in order to confirm that the market-wide overconfidence is not a direct aggregation of the disposition bias. The VAR model estimated for individual securities contains security turnover (*turn*), market return (*mret*) and security return (*ret*) as the endogenous variables while security volatility (*sig*) is the only exogenous variable. *sig* is computed in

⁵ Consistent with Lo and Wang (2000), market turnover is used as a measure of trading activity, and individual security turnover is aggregated to compute market turnover on a value-weighted basis.

the same manner as *msig*, however, r_t now denotes the return of the individual security on day t . Accordingly, the VAR model for individual securities employs the following specification:

$$\begin{bmatrix} turn_t \\ ret_t \\ mret_t \end{bmatrix} = \begin{bmatrix} \alpha_{turn} \\ \alpha_{ret} \\ \alpha_{mret} \end{bmatrix} + \sum_{k=1}^K A_k \begin{bmatrix} turn_{t-k} \\ ret_{t-k} \\ mret_{t-k} \end{bmatrix} + \sum_{l=0}^L B_l sig_{t-l} + \begin{bmatrix} e_{turn,t} \\ e_{ret,t} \\ e_{mret,t} \end{bmatrix} \quad (3.4)$$

In equation 3.3 and 3.4, K denotes the optimal number of lagged observations of the endogenous variables and L indicates the optimal number of lagged observations of the exogenous variables. The optimal lag length is selected by analysing the information criteria. Furthermore, for each variable, the mean coefficient between all individual securities is reported in order to achieve brevity in the VAR model for individual securities. However, the statistical significance of the average coefficient is difficult to ascertain. To address this problem, Statman, *et al.* (2006) employ a VAR bootstrap measure to generate standard errors for each VAR coefficient and to generate coefficients for the security impulse response functions. This is because, based on the VAR models, Statman, *et al.* (2006) analyse impulse responses to determine how endogenous variables respond to each other over time.

The regressed market VAR model (equation 3.3) illustrates a positive association between the turnover of the market (*mturn*) and past returns of the market (*mret_{t-k}*). Accordingly, Statman, *et al.* (2006) argue that this finding conforms to the overconfidence hypothesis. Additionally, the regressed VAR model for individual securities (equation 3.4) reveal that security turnover (*turn*) has a significant, positive relationship with lagged market returns (*mret_{t-k}*) and with lagged security returns (*ret_{t-k}*). Statman, *et al.* (2006) interpret the reliance of individual security turnover on historical market returns as consistent with the overconfidence hypothesis. The dependence of security turnover to its own past returns is construed as the disposition effect. Overall, Statman, *et al.* (2006) conclude that, even after controlling for the influence of the disposition effect, realized trading volume is positively influenced by market return-induced investor overconfidence.

3.2.2. Griffin, *et al.* (2007)

The objective of Griffin, *et al.* (2007) is to determine whether investors trade more frequently when stocks are generating favourable returns. The study provides evidence from 46 countries by analysing daily and weekly observations of the market index in each country for the period of January 1993 until June 2003. VAR models and their associated impulse response functions are employed to investigate the relationship between current trading activity and historical returns. However, the empirical analysis is twofold. First, the VARs are evaluated on a country-by-country basis. Thereafter, the VARs are estimated with the inclusion of global return and volatility variables. Market turnover and market returns are used as the

endogenous variables in the VAR models. Volatility is used as an exogenous variable in the VARs to account for alternative explanations of the lead-lag interaction between turnover and returns. Notably, Griffin, *et al.* (2007) use an EGARCH (1,1) specification to compute the respective volatilities.

The regressed VAR models demonstrate that, for a majority of the observed stock markets, a strong, positive relationship is present between the contemporaneous turnover and past returns of the respective market, thus, indicating the presence of investor overconfidence. These findings hold even after controlling for volatility, during different sample periods, alternative data frequencies, and are present with differing definitions of turnover. Additionally, Griffin, *et al.* (2007) report that the positive influence of past market returns on current turnover is more significant in countries with high market volatility, with short-sale constraints, and in which the level of corruption is high.

3.2.3. Kyrychenko and Shum (2009)

Kyrychenko and Shum (2009) examine the behavioural characteristics of investors in the United States (U.S.) to determine which factors influence their investments in foreign stocks and bonds. The study employs the Repeated-Imputation Inference (RII) technique to generate inferences from information contained in the U.S. Surveys of Consumer Finance (SCF) for the period 1992 till 2004. With respect to investor overconfidence, Kyrychenko and Shum (2009) use the frequency of trades as a proxy of overconfidence. The frequency of trades exhibits a significant, positive relationship with foreign asset ownership, and thus, indicates that overconfident investors are more likely to hold foreign assets (including equities and bonds). This is because overconfident investors overestimate their knowledge about foreign assets, and thus, overconfident investors are less averse to investing in international equities and bonds (Kyrychenko and Shum, 2009).

3.2.4. Yung and Liu (2009)

Yung and Liu (2009) test the persistence of overconfident trading in the U.S futures market. The study examines the prices and trading volumes of six commodity futures (crude oil, copper, unleaded gas, natural gas, silver, and gold) for the sample period of January 1995 to December 2006. The interaction between returns and turnover is investigated using a Vector Error Correction Model (VECM) which contains futures turnover and futures returns as endogenous variables and futures volatility as an exogenous variable. Yung and Liu (2009) find that, for all commodities, lagged futures returns do not positively influence futures turnover. Therefore, the study concludes that, over the sample period, overconfident trading is not present in the commodity futures market. Notably, these findings are not consistent with the findings of Statman, *et al.* (2006) and Griffin, *et al.* (2007) who find that overconfident trading is present in stock markets.

3.2.5. Zaiane and Abaoub (2009)

Zaiane and Abaoub (2009) examine the influence of the overconfidence bias on the decisions made by investors in an emerging market. The observed emerging market is the Tunisian stock market, for which the study analyses common stocks from January 2000 to December 2006. Consistent with Statman, *et al.*, (2006), the study employs VAR models and its related impulse response functions to examine the extent to which lagged market returns are correlated with current trading activity. The VAR model consists of market return and market turnover as endogenous variables, and market volatility and dispersion as exogenous variables. Zaiane and Abaoub (2009) find little evidence indicating that investors are overconfident in the Tunisian stock market when shares traded are used as a proxy of the trading volume.

3.2.6. Lin, Rahman and Yung (2010)

Lin, *et al.* (2010) note that portfolio managers regard REITs as an asset class that is distinct from common stocks. As such, the objective of Lin, *et al.* (2010) is to validate the overconfidence paradigm in the REITs market. To achieve this objective, the authors analyse monthly returns of all REITs that are found in The Center for Research in Security Prices (CRSP) database for the study period of January 1990 to December 2006. The study uses an empirical model that is similar to that of Statman, *et al.* (2006). Accordingly, the study's primary tools of analysis consist of VAR models and their related impulse response functions.

The study's overconfidence model without market (that is, the stock market) control variables has REIT turnover (RT) and REIT return ($REITret$) as the endogenous variables, while REIT volatility ($REITvolat$) and REIT dispersion ($REITdisp$) are the exogenous variables. This results in a model specification following equation 3.5:

$$\begin{bmatrix} RT_t \\ REITret_t \end{bmatrix} = \begin{bmatrix} \alpha_{RT} \\ \alpha_{REITret} \end{bmatrix} + \sum_{p=1}^P A_p \begin{bmatrix} RT_{t-p} \\ REITret_{t-p} \end{bmatrix} + \sum_{q=0}^Q B_q \begin{bmatrix} REITdisp_{t-q} \\ REITvolat_{t-q} \end{bmatrix} + \begin{bmatrix} \varepsilon_{RT,t} \\ \varepsilon_{REITret,t} \end{bmatrix} \quad (3.5)$$

In equation 3.5, P and Q denote the optimal lag lengths of the endogenous and exogenous variables and are selected based on the information criteria. $REITret$ denotes the monthly value-weighted average REIT return, while RT denotes the monthly value-weighted average of individual REIT turnover, computed as in equation 3.6.

$$RT = \sum_{i=1}^n \left[\left(\frac{V_i}{S_i} \right) \left(\frac{P_i S_i}{\sum_{i=1}^n P_i S_i} \right) \right] \quad (3.6)$$

In equation 3.6, REIT i 's trading volume, outstanding shares, and price are denoted as V_i , S_i , and P_i , while n indicates the number of REITs in the CRSP database. $REITvolat$ denotes the volatility of the value-weighted monthly returns of the REITs and is computed in a similar manner as equation 3.1. $REITdisp$ is

the cross-sectional standard deviation of the monthly returns of all the REITs and is computed in a similar manner as equation 3.2. The regressed VAR model (equation 3.5) demonstrates that, consistent with the overconfidence hypothesis, current REIT turnover is significantly, positively impacted by the lagged one month REIT return. This indicates that the trading activity of REIT investors is only influenced by the most recent (last month) REIT returns, and their trading decisions are not biased towards by longer-term REIT returns.

In addition, the VAR model estimated by Lin, *et al.* (2010) to control for stock market variables contains market turnover (MT), market return ($Mret$), market volatility ($Mvolat$) and market dispersion ($Mdisp$) as additional exogenous variables, as highlighted in equation 3.7. In equation 3.7, $Mret$ is based on the value-weighted average return of all common stocks (except REITs and ADRs) found in the CRSP database. MT represents the value-weighted average turnover of each stock. Moreover, $Mdisp$ is the cross-sectional standard deviation of the returns of all stocks, while $Mvolat$ represents the volatility of the monthly value-weighted average return of all stocks.

$$\begin{bmatrix} RT_t \\ REITret_t \end{bmatrix} = \begin{bmatrix} \alpha_{RT} \\ \alpha_{REITret} \end{bmatrix} + \sum_{p=1}^p A_p \begin{bmatrix} RT_{t-p} \\ REITret_{t-p} \end{bmatrix} + \sum_{q=0}^Q B_q \begin{bmatrix} MT_{t-q} \\ Mret_{t-q} \\ REITdisp_{t-q} \\ REITvolat_{t-q} \\ Mdisp_{t-q} \\ Mvolat_{t-q} \end{bmatrix} + \begin{bmatrix} \varepsilon_{RT,t} \\ \varepsilon_{REITret,t} \end{bmatrix} \quad (3.7)$$

The regressed VAR model (equation 3.7) demonstrates that, after controlling for the general market overconfidence, the current REIT turnover is significantly, positively influenced by the REIT return, and is, therefore, indicative of the presence of overconfidence by REIT investors. On the other hand, contemporaneous REIT turnover displays no significant association with the stock market return. Lin, *et al.* (2010) mention that these findings suggest that the overconfidence caused by stock market return does not spill-over into the REIT market. Consequently, these results indicate that the dynamics of the REIT market is distinct from that of the general stock market. Furthermore, the study finds that, subsequent to positive market returns, overconfident investors have a low tendency to trade in stocks with high levels of company-specific risks.

3.2.7. Bailey, *et al.* (2011)

Bailey, *et al.* (2011) explore how behavioural biases influence mutual fund choices. The study examines the panel of trades and portfolio positions of a sample of U.S. individual investors from January 1991 until November 1996. Descriptive statistics, correlation analysis, and factor analysis are used as the study's main tools of analysis. Bailey, *et al.* (2011) discover that investors who are more experienced and more informed tend to generate higher returns due to their effective use of mutual funds. Further analysis reveals that unsophisticated investors may be subject to behavioural biases or do not consider macro-economic or firm-specific news when purchasing mutual funds. Specifically, the factor analysis reports that, amongst other behavioural biases, investors display overconfidence when trading in mutual funds. Therefore, biased investors chose mutual funds for incorrect reasons (Bailey, *et al.*, 2011). These findings emphasize the importance of research question one because it is possible that investors display overconfidence when trading in ETFs since ETFs and mutual funds are similar in that they are both pooled investment funds.

3.2.8. Chuang and Susmel (2011)

Chuang and Susmel (2011) assess whether investor overconfidence is more pronounced in institutional or individual investors in the Taiwan Stock Exchange. The study analyses weekly stock returns, turnover, and market capitalisations of all stocks listed on the Taiwan Stock Exchange from January 1996 until December 2005. Moreover, fractional institutional ownership data is used to form portfolios based on the portion of shares held by institutional investors. To achieve their objective, Chuang and Susmel (2011) employ a Seemingly Unrelated Regression (SUR) model to examine the interaction between market returns and portfolio volumes. This relationship is further investigated across different market states using a multivariate SUR model.

The regressed multivariate SUR model shows that, subsequent to positive market returns, individual investors trade more frequently in bull markets relative to non-bull markets. However, subsequent to market gains, the trading behaviour of institutional investors displays no significant differences in bull and non-bull markets. Overall, Chuang and Susmel (2011) find that, conforming with Gervais and Odean's (2001), overconfidence is more pronounced in individual investors than institutional investors, in which case, individual investors are more overconfident in bull markets relative to non-bull markets.

3.2.9. Cohen and Shin (2013)

Cohen and Shin (2013) study the trade flows and price changes of U.S. treasury notes to detect the presence of positive feedback trading in the U.S. Treasury securities market. The interactions between prices changes and trades are analysed using VAR models for the sample period of January 1999 till December 2000.

Results from the regressed VAR models indicate that decreases in treasury note prices induce more sell orders while increases in treasury note prices induce more buy orders. Cohen and Shin (2013) conclude that the U.S. treasury market exhibits positive feedback trading, particularly during market stress periods. This positive feedback trading present in the U.S. treasury market could lead to overconfident trading by U.S. treasury note investors. This is because, if the trend in prices continues, investors may believe that they were able to predict this trajectory, and thus, overestimate their trading abilities.

3.2.10. Metwally and Darwish (2015)

Metwally and Darwish (2015) test the presence of investor overconfidence on the Egyptian Exchange. The study's sample covers all stocks listed on the Egyptian Exchange, however, the overconfidence bias is investigated on an aggregate market level by examining the association between the turnover of the market and its returns for the period 2002 till 2012. The sample period is divided into two tranquil upward-trending periods (2002 to 2004 and 2005 to 2007) and two volatile downward-trending periods (2008 to 2010 and 2011 to 2012). Following Statman, *et al.* (2006), the relation between market turnover and market returns is investigated by using VARs and its associated impulse response functions. Market turnover is used as the dependent variable in the VAR model, while market return is used as the independent variable. Additionally, the study employs Granger causality tests to determine how market turnover is related to market returns.

For the entire sample period, the regressed VAR model indicates that the first lagged market return has a significant positive relationship with current market turnover, and therefore, coincides with the overconfidence hypothesis. These findings are supported by the impulse response functions associated with the estimated VAR models. Moreover, the Granger causality tests support the notion that market return Granger causes market turnover. With respect to the sub-samples, Metwally and Darwish (2015) find that, in the Egyptian stock market, trading activity is significantly influenced by the state of the market. Specifically, the study reports that investors' overconfidence triggers trading activity when the Egyptian stock market is experiencing an uptrend. Both Egypt and South Africa are emerging African economics with financial markets in which investors may display similar trading patterns. Therefore, if investors are overconfident in Egypt, it is possible that South African investors are also influenced by the overconfidence bias. Hence, these findings further emphasize the need to investigate the presence of overconfident trading by South African investors.

3.2.11. Pak and Chatterjee (2016)

Pak and Chatterjee (2016) examine the extent to which overconfidence influences the riskiness of an investor's portfolio. Using data obtained from the Cognitive Economics Study (CogEcon), the study analyses the portfolio allocations of American investors from 2009 till 2013. The Zero-Inflated Beta (ZIB) model is employed to study the changes in investors' portfolio ownerships and allocations. Results from the regressed ZIB model indicate that overconfident investors are more likely to invest in risky assets, such as equities. These overconfident investors stay longer in equity markets even when they lack the cognitive capacity to handle investment information. Additionally, the results suggest that overconfident investors are less likely to hold less risky assets (such as bonds) and professionally managed funds (such as mutual funds). In contrast, Pak and Chatterjee (2016) find that, in order to reduce information costs, well-calibrated and financially sophisticated investors are more likely to hold less risky assets and professionally managed funds. Hence, Pak and Chatterjee (2016) argue that financial sophistication plays an important role in minimising the effect of the overconfidence bias.

3.2.12. Zia, Sindhu and Hashmi (2017)

Zia, *et al.* (2017) explore the existence of investor overconfidence in the Pakistani stock market. A random sampling procedure is used to select stocks trading on the Karachi Stock Exchange during 2005 to 2013. Unit root tests, VARs, impulse response functions, and Granger causality tests form part of the study's econometric techniques. Specifically, a security level panel VAR model is employed to investigate how current turnover is related to lagged returns. According to Zia, *et al.* (2017), panel VAR is used since the study analyses firm-specific data across time. For the panel VARs, stock returns and turnover are used as endogenous variables, while volatility is used as the exogenous variable. The panel VAR used by Zia, *et al.* (2017) takes the following form:

$$Y_t = \alpha + \sum_{k=1}^K A_k Y_{t-k} + \sum_{l=1}^L B_l X_{t-1} + e \quad (3.8)$$

In equation 3.8, Y_t represents a matrix of the endogenous variables (that is, stock turnover and stock return) and X_t represents stock volatility. Interestingly, Zia, *et al.* (2017) find that the current day's turnover is significantly influenced by the returns of preceding days, and therefore, this finding is indicative of the existence of investor overconfidence. Furthermore, impulse response functions illustrate that historical returns positively authenticate the current market turnover. The Granger causality test further supports the notion that past returns have an impact on current turnover. Overall, Zia, *et al.* (2017) conclude that, in the Pakistani stock market, traders are overconfident since current turnover depends significantly on historical returns. Notably, these findings are consistent with Metwally and Darwish (2015) who employ VAR

models, impulse response functions, and Granger causality tests, and document that investor overconfidence is present in the Egyptian stock market.

3.2.13. Aharon and Qadan (2018)

Aharon and Qadan (2018) investigate how U.S. investors' confidence levels affect their demand for information about commodities. Using a sample period of January 2004 till February 2017, the study utilises spot prices and Google search query data for platinum, gold, silver, palladium, copper, zinc and oil as its sample data. The response of investors' attention to changes in commodity prices is analysed using a linear regression. The results indicate that, for all commodities, except zinc and platinum, investors' demand for information relative to fluctuations in the prices of commodities is asymmetric. In other words, investors' demand for information increases when commodity prices decrease and investors' demand for information decreases when the prices of commodities increase (Aharon and Qadan, 2018). According to Aharon and Qadan (2018), these results conform to the overconfidence hypothesis, which suggests that overconfident investors trade asymmetrically between losses and gains.

3.2.14. Gupta, Goyal, Kalakbandi and Basu (2018)

Gupta, *et al.* (2018) explore the influence of the overconfidence bias in the emerging equity markets of Indian and China. The study examines the persistence of the overconfidence bias before, during, and after the 2008 subprime financial crisis, and thus, the study period of April 2001 to March 2016 is segregated into subsamples to account for pre-recession (April 2001 to June 2007), during-recession (July 2007 to December 2009), and post-recession (January 2010 to March 2016). All common stocks (excluding ADRs, Global Depository Receipts (GDRs), and closed-end funds) listed on the Bombay Stock Exchange and Shanghai Stock Exchange are used as the study's sample data. The Bombay Stock Exchange and Shanghai Stock Exchange represent the Indian and Chinese markets, respectively.

Following Statman, *et al.* (2006), the study uses VAR models to investigate the relationship between current turnover and lagged returns. However, Gupta, *et al.* (2018) employ market liquidity, *mliquidity*, as an additional endogenous variable since liquidity is considered as one of the drivers of investor overconfidence. Hence, the endogenous variables used in the market VAR model include; market turnover, market returns, and market liquidity. Market turnover, *mturnover*, and market returns, *mreturns*, follow the computation proposed by Statman, *et al.* (2006). Market volatility, *mvolatility*, is the only exogenous variable and is computed in the manner suggested by Statman, *et al.* (2006). Accordingly, the market version of the VAR model follows equation 3.9 (page 39). In equation 3.9, the optimal lag lengths K and L are selected based on the information criteria.

$$\begin{bmatrix} mreturns_t \\ mturnover_t \\ mliquidity_t \end{bmatrix} = \begin{bmatrix} \alpha_{mreturns} \\ \alpha_{mturnover} \\ \alpha_{mliquidity} \end{bmatrix} + \sum_{k=1}^K A_k \begin{bmatrix} mreturns_{t-k} \\ mturnover_{t-k} \\ mliquidity_{t-k} \end{bmatrix} + \sum_{l=0}^L B_l [mvolatility_{t-l}] + \begin{bmatrix} e_{mreturns,t} \\ e_{mturnover,t} \\ e_{mliquidity,t} \end{bmatrix} \quad (3.9)$$

The regressed VAR models for the entire sample indicate that, in each of the two markets, current market turnover depends on past market returns, and therefore, the findings support the overconfidence hypothesis. Additionally, Gupta, *et al.* (2018) find that investors in India are less overconfident than Chinese investors. Using the 2008 financial crisis as a structural break, the regressed VAR models for each subsample demonstrates that both Chinese and Indian investors are overconfident before the global recession. However, only Chinese investors are overconfident during the recession. Likewise, after the global recession, Chinese investors exhibit overconfident trading, while investors in India lose their confidence. The loss of confidence by Indian investors may be due to the severe impact of the global financial crisis on the Indian market which caused investors to trade less fearing the possibility of substantial losses (Gupta, *et al.*, 2018).

To determine whether or not the observed overconfidence is a summation of the disposition effect, Gupta, *et al.* (2018) estimate VARs at the individual security level. The liquidity factors are not included in the individual security VAR since they will not influence the overconfidence hypothesis (Gupta, *et al.*, 2018). Accordingly, the VAR model for each individual security becomes tri-variate:

$$\begin{bmatrix} sreturns_t \\ sturnover_t \\ mreturns_t \end{bmatrix} = \begin{bmatrix} \alpha_{sreturns} \\ \alpha_{sturnover} \\ \alpha_{mreturns} \end{bmatrix} + \sum_{k=1}^K A_k \begin{bmatrix} sreturns_{t-k} \\ sturnover_{t-k} \\ mreturns_{t-k} \end{bmatrix} + \sum_{l=0}^L B_l [svolatility_{t-l}] + \begin{bmatrix} e_{sreturns,t} \\ e_{sturnover,t} \\ e_{mreturns,t} \end{bmatrix} \quad (3.10)$$

In equation 3.10 above, *sreturns*, *sturnover*, and *svolatility* denote the individual stock's returns, turnover, and volatility, respectively. According to Gupta, *et al.* (2018), reporting the VAR model for each individual security is extensive. Instead, the study opts to estimate the VAR based on the average value of each variable for day *t*, and this model is labelled security-VAR. For each country, the regressed security-VAR model demonstrates that the turnover of individual securities is positively influenced by the past returns of the market, thereby, proving that the overconfidence observed in the respective market is not an aggregation of the disposition bias. Interestingly, the study finds that, in India, although overconfidence boosts trading volume at the market level, it does not increase trading volume at security level since the coefficient of *sturnover* in relation to lagged *mreturns* is not significant. However, Chinese investors

display overconfident trading at individual security level since *sturnover* significantly depends on past *mreturns*.

Additionally, the study employs a threshold VAR model to determine whether the overconfidence bias is more prominent in up or down markets. The threshold VAR uses a pre-defined computation process, and therefore, its algorithm is able to capture up and down regimes. Equation 3.11 below presents the general form of the threshold VAR:

$$X_t = A^0(L)X_{t-1} + (A^1(L)X_{t-1})I(c_{t-d} > r) + \varepsilon_t \quad (3.11)$$

In equation 3.11 above, X_t is a vector of variables while $A^0(L)$ and $A^1(L)$ are lag polynomials. c_{t-d} denotes the threshold variable which controls the regime that the system is in and r represents the critical threshold value. $I(c_{t-d} > r)$ is a dummy variable that takes the value of one when $c_{t-d} > r$ and zero otherwise. Before estimating the threshold VAR, a test for non-linearity is conducted. The study reports that, for both countries, contemporaneous turnover is reliant on past market returns. However, the study finds that, for both Indian and Chinese traders, overconfidence is more pronounced in upmarket regimes than in down market regimes. These findings are consistent with the findings of Chuang and Susmel (2011). Thus, investors display less confidence and pessimistic investment behaviour when the market is following a downtrend.

3.2.15. Chen and Sabherwal (2019)

In a more recent study, Chen and Sabherwal (2019) investigate whether the high trading activity present in options markets is associated with investor overconfidence. The study observes equity options traded in the U.S. for the period 1996 to 2015. To reduce the noise that is usually associated with high-frequency data, Chen and Sabherwal (2019) use monthly observations in their analysis. The association between option volume and lagged stock market returns is examined using an Ordinary Least Squares (OLS) regression as opposed to the VAR model. This is because, in options markets, investor overconfidence may build gradually, and correspondingly, may not dissolve away easily (Chen and Sabherwal, 2019). Subsequently, Chen and Sabherwal (2019) contend that rolling cumulative lagged market returns are a more suitable proxy for investor overconfidence than individual monthly returns.

Accordingly, the OLS regression contains option trading volume as the dependent variable and cumulative stock market returns as the independent variable, which results in the following model specification:

$$Vol_t = \alpha + \beta_1 Ret_{t-k,t} + \sum \beta_j ControlVariable_{j,t} \quad (3.12)$$

In equation 3.12, Vol_t denotes the measure of trading volume while the monthly cumulative lagged stock market returns are denoted as $Ret_{t-k,t}$ for time period $t - k$ until t . Following Statman, *et al.* (2006), the

control variables in equation 3.12 are market volatility and market dispersion, computed as in equation 3.1 and 3.2. To examine the robustness of the model, Chen and Sabherwal (2019) model equation 3.12 with the mean absolute deviation (MAD) in place of market volatility. The MAD is computed as the value-weighted average of the deviations existing between the firm's return from the market return.

The results of the regressed OLS model (equation 3.12) demonstrates that historical stock market returns significantly positively impact the trading activities of options investors. This relationship holds even after controlling for the volatility in option and stock markets and after controlling for the idiosyncratic risks of individual firms. As a result, these findings are indicative of the presence of overconfidence in options markets. Specifically, the evidence from this study suggests that improved performance in the market increases investors' confidence levels regarding their trading and security valuation skills, causing them to trade more frequently. Additionally, the study reports that, for overconfident investors, options markets are more likely to be a trading vehicle than equity markets. Remarkably, these findings emphasize the need to investigate the presence of overconfident trading in ETF markets because ETFs and options have the same basic characteristic, that is, their value is derived from an underlying asset.

3.2.16. Alsabban and Alarfaj (2020)

More recently, Alsabban and Alarfaj (2020) examine the existence of overconfident trading by investors in the Saudi Stock Exchange (Tadawul) by analysing monthly data of 172 stocks for the period January 2007 to December 2018. In this study, overconfident behaviour is detected using the Statman, *et al.* (2006) model. More specifically, Alsabban and Alarfaj (2020) study the presence of overconfident trading by inspecting the interaction between market turnover and market return using VAR models and impulse response analysis. The market-wide VAR model estimated follows equation 3.3. Notably, Alsabban and Alarfaj (2020) compute the optimal lag length of the exogenous variables by selecting the optimal lag length of the endogenous variables and then estimating the VAR model with different lags (starting from 1 lag to 5 lags) of the exogenous variables, and the lag number with the smallest Akaike's (1973) Information Criterion (AIC) is selected as optimal⁶.

The results obtained by Alsabban and Alarfaj (2020) confirm the overconfidence hypothesis because there exists a positive and statistically significant association between the current turnover of the market and the historical returns of the market. Additionally, Granger causality tests indicate a unidirectional Granger causality from lagged market return to current market turnover and impulse response functions display

⁶ This method of selecting the optimal lag lengths is used in the present study.

significant, positive responses of market turnover to lagged market returns. Overall, Alsabban and Alarfaj (2020) report robust results that investors in the Saudi stock market are overconfident. These findings are similar to the findings of Metwally and Darwish (2015) and Zia, *et al.* (2017) who analyse the Egyptian and Pakistani stock markets, respectively.

3.3. Overconfidence and Market Volatility

The following section reviews empirical studies that examine the effect of trading volume induced by overconfidence on the volatility of the market. Additionally, the studies conducted by Abbes (2013) and Jlassi, *et al.* (2014) inspect how trading by overconfident investors impacts volatility during, before and after the 2008 financial crisis. Notably, the studies that are reviewed under this section survey asset classes that do not include ETFs because there are no existing studies that examine the effect of overconfident trading on ETF market volatility.

3.3.1. Chuang and Lee (2006)

The study conducted by Chuang and Lee (2006) examines the overconfidence hypothesis by employing aggregate data of all firms trading on the AMEX and NYSE from January 1963 till December 2001. Chuang and Lee (2006) report four key empirical findings. Firstly, Chuang and Lee (2006) employ a moving average model to observe how the prices of stocks respond to information (both public and private) and report that overconfident investors overreact to private information but underreact to public information. Secondly, Chuang and Lee (2006) make use of bivariate Granger causality tests to examine the causal relationship between returns and trading volume. The results of the Granger causality tests indicate that market return Granger causes trading volume, and thus, the second finding of Chuang and Lee (2006) is that positive market returns lead to excessive trading by overconfident traders in periods after positive market returns.

The third objective of Chuang and Lee (2006) is to investigate whether trading by overconfident investors contributes to excess volatility. Remarkably, one of the most common methods used to examine the effect of overconfident trading on market volatility is the Chuang and Lee (2006) model⁷. Chuang and Lee (2006) begin by decomposing the trading volume into two components, specifically, the component of trading volume that is related to investor overconfidence and the component of trading volume that is not associated with investor overconfidence. Equation 3.13 (page 43) is employed to decompose the trading volume into the two components.

⁷ This study employs the Chuang and Lee (2006) model to answer research questions three and four.

$$V_t = \alpha + \sum_{j=1}^J \beta_j R_{t-j} + \varepsilon_t = \left[\sum_{j=1}^J \beta_j R_{t-j} \right] + [\alpha + \varepsilon_t] = OVER_t + NONOVER_t \quad (3.13)$$

In equation 3.13, the trading volume, V_t , follows the definition of the turnover ratio, R_t represents the return for day t and R_{t-j} is the return at time $t - j$ where J is the optimal number of lags, β_j is the coefficient that captures the relationship between past market returns and current trading volume, α is the constant term, and ε_t denotes the residual term. Chuang and Lee (2006) compute the component of trading volume that is not related to overconfidence ($NONOVER_t$) as the summation of the constant and residual terms. Accordingly, the component of trading volume that is related to overconfidence is calculated as the trading volume (V_t) minus the sum of the constant and residual terms ($NONOVER_t$).

The two trading volume components are then incorporated into the conditional variance equation of an EGARCH (1,1) model as follows:

$$R_t = \mu_t + \eta_t \quad (3.14)$$

$$\ln(\sigma^2) = \omega + f_1 \frac{|\eta_{t-1}|}{\sqrt{\sigma_{t-1}^2}} + f_2 \frac{\eta_{t-1}}{\sqrt{\sigma_{t-1}^2}} + f_3 \ln(\sigma_{t-1}^2) + f_4 OVER_t + f_5 NONOVER_t \quad (3.15)$$

where equation 3.14 represents the mean equation of the EGARCH (1,1) model and equation 3.15 represents the conditional variance equation. In equation 3.14, μ_t is the average R_t based on the historical information and η_t denotes the residual term of the mean equation. In equation 3.15, f_1 , f_2 , and f_3 are parameters representing the Autoregressive Conditional Heteroskedasticity (ARCH) effect, the asymmetry effect, and the Generalised Autoregressive Conditional Heteroskedasticity (GARCH) effect, respectively. f_4 captures the effect of overconfident trading on volatility while f_5 captures the impact of non-overconfident trading on volatility. According to Chuang and Lee (2006), when trading volume induced by overconfidence contributes to volatility, $f_4 > f_5 > 0$. Overall, Chuang and Lee (2006) find that the results from the estimated EGARCH (1,1) models indicate that trading volume induced by investor overconfidence adds to the observed conditional volatility. Specifically, the third key empirical finding of this study is the overconfident trading exhibits a significant, positive relationship with volatility. Finally, the fourth empirical finding of this study is that investors who are overconfident underestimate risk, thus, trading more in riskier securities.

3.3.2. Sheikh and Riaz (2012)

Sheikh and Riaz (2012) investigate whether market gains (losses) are followed by high (low) trading activities and whether trading volume induced by overconfidence leads to excessive volatility. The study analyses data of companies listed on the Karachi Stock Exchange from November 1999 to October 2010. To examine how the current turnover of the market relates to the past returns of the market, Sheikh and

Riaz (2012) employ the Statman, *et al.* (2006) model and estimate a VAR model that follows equation 3.3. Consistent with the overconfidence hypothesis, the results of the estimated VAR model indicate that, even after controlling for dispersion and volatility, past market returns are significantly, positively related to current trading activity.

Following the Chuang and Lee (2006) model, the effect of investor overconfidence on volatility is examined by decomposing trading volume into an overconfidence component and a non-overconfidence component as specified in equation 3.13. Thereafter, the trading volume components are merged into the variance equation of an EGARCH model as in equation 3.15. The estimated EGARCH models indicate that trading volume induced by overconfidence exhibits a positive relationship with volatility, however, this relationship is insignificant. Hence, Sheikh and Riaz (2012) conclude that there is no significant evidence that trading associated with investor overconfidence leads to increased volatility. Interestingly, these findings are inconsistent with the findings of Chuang and Lee (2006) who report that investor overconfidence significantly increases market volatility.

3.3.3. Abbes (2013)

The primary objective of Abbes (2013) is to examine whether the high volatility experienced during the 2008 financial crisis was caused by investor overconfidence. Abbes (2013) analyses the market index from 15 different countries from January 1999 to December 2009. The study employs the Chuang and Lee (2006) model to investigate how overconfident trading impacts market return volatility. Using equation 3.13, the trading volume of each market index is separated into a component of trading volume that is due to investors' overconfidence and a component of trading volume that is not related to investors' overconfidence. The effect of overconfidence on volatility is then examined by including the two trading volume components into the conditional variance equation of the EGARCH (1,1) model following equation 3.15.

Abbes (2013) finds that, for the entire sample, the trading volume produced by overconfidence is positively related to conditional volatility. This finding is indicative of the strong presence of investor overconfidence in the emerging and developed stock markets observed. Therefore, Abbes (2013) argues that the overconfidence bias was a key contributor to the financial instability that erupted during the 2008 global financial crisis. However, during the crisis period, Abbes (2013) reports that there is no significant relationship between the overconfidence bias and volatility. This was because, during the crisis period, investors lost confidence in financial markets (Abbes, 2013).

3.3.4. Jlassi, *et al.* (2014)

Jlassi, *et al.* (2014) investigate the effect of investor overconfidence on the volatility of stock markets during the 2008 subprime financial crisis. The study's sample data comprises of price and transaction volume for market indices from 27 different countries for January 2000 till December 2012. The JSE All-Share Index (ALSI) is used to provide evidence from South Africa as an emerging market. Consistent with Chuang and Lee (2006) and Abbes (2013), the trading volume of each index is decomposed into an investor overconfidence component and a component unrelated to investor overconfidence using equation 3.13. The two trading volume components are then merged into the conditional variance equation of the EGARCH model, as illustrated in equation 3.15.

Jlassi, *et al.* (2014) report robust evidence of the positive effect of overconfident trading on volatility in global equity markets, except for Chile. Furthermore, the study finds that, consistent with the theoretical assumptions, the overconfidence bias is more pronounced in up markets and before the crisis. These findings imply that, during tranquil market conditions, investors ignore market warning signals and trade excessively, and thus, cause an increase in asymmetric stock volatility. Further analysis reveals that investor overconfidence can account for a large portion of the excess and asymmetric volatility that is present in global financial markets, and therefore, investor overconfidence represents a key incentive that triggered the 2008 subprime financial crisis. This is because, according to Jlassi, *et al.* (2014), overconfidence causes investors to trade excessively; this subsequently causes an increase in stock prices beyond their true values. Therefore, there is an abnormal increase in stock volatility because stock prices shift away from their fundamental values. Contrary to Abbes (2013), Jlassi, *et al.* (2014) find that overconfidence has a significant, positive effect on volatility during and after the 2008 global financial crisis.

3.3.5. Mushinada and Veluri (2018)

Mushinada and Veluri (2018) employ the Chuang and Lee (2006) model to investigate the effect of overconfident trading on market volatility in the Bombay Stock Exchange (BSE). The study observes 1 290 stocks trading on the BSE from April 2004 till March 2012. The results from the estimated EGARCH models indicate that the excessive trading by overconfident investors contributed to high market volatility during April 2004 to September 2008. However, during the post-crisis period from October 2008 to March 2012, Mushinada and Veluri (2018) find that market volatility is insignificantly impacted by overconfident trading. Notably, these findings are inconsistent with Jlassi, *et al.* (2014) who document that overconfident trading exhibits a significant, positive effect on market volatility after the 2008 global financial crisis.

3.4. Overconfidence in ETF Markets

This section provides a review of empirical research that studies the existence of investor overconfidence in ETF markets. However, there is limited research surrounding the presence of overconfident trading in ETF markets, thus, highlighting the need for empirical research on the subject.

3.4.1. Da Dalt, *et al.* (2019)

Da Dalt, *et al.* (2019) examine the impact of contrarianism on the investment behaviour of Finnish ETF investors trading on the Helsinki Stock Exchange (OMXH). Contrarianism refers to the tendency of investors to trade against the prevailing market trends (Da Dalt, *et al.*, 2019). According to Da Dalt, *et al.* (2019), overconfidence leads to contrarian behaviour if investors are overconfident in their stock-picking ability, and therefore, traders who are the most overconfident exhibit stronger contrarian behaviour. Conversely, the lack of confidence would lead to investors trading in the same direction as the general market. Da Dalt, *et al.* (2019) observe the trading activity of investors trading the OMXH25 ETF which tracks the OMXH 25 Index. The sample period ranges from February 2002 to December 2014.

Analysis of buy-ratios, which measure the degree of aggregate trading volume by households, and analysis of ETF trades conditional on trade direction forms part of the study's primary tools of analysis. Da Dalt, *et al.* (2019) find that the strength of contrarian behaviour is stronger when Finnish households purchase ETFs than when they sell them. Additionally, the results demonstrate that Finnish households display less contrarian behaviour when trading ETFs than when trading common stocks. According to Da Dalt, *et al.* (2019), these findings indicate that investors are less overconfident when trading ETFs than when trading common stocks. Overall, the lower contrarian behaviour of Finnish ETF investors implies that Finnish ETF investors are less overconfident in their ability to identify ETFs that are undervalued.

3.5. Evidence From South Africa

The current section reviews studies conducted on the South African market. Except for Dowie and Willows (2016), none of the studies specifically test for overconfidence in ETFs or other asset classes; however, these studies can be used to provide an indication of the influence of the overconfidence bias on the investment decisions made by South African investors. This lack of empirical research on the overconfidence bias in South Africa highlights the need for research on the subject in South Africa.

3.5.1. Charteris, Chau, Gavriilidis and Kallinterakis (2014)

Charteris, *et al.* (2014) examine the existence of feedback trading in ETFs listed in emerging markets. The study observes broad-index ETFs listed in South Africa, Brazil, India, and South Korea from the inception date of each ETF until 7 December 2012. Using the empirical model prescribed by Sentana and Wadhvani (1992), the study finds that ETFs listed in emerging markets exhibit significant feedback trading once its premiums and discounts are taken into account. With respect to the South African ETF, Charteris, *et al.* (2014) find that the ETF exhibits positive feedback trading during the pre-crisis period. This presence of positive feedback trading in South African ETFs could be attributed to overconfident trading such that if an investor experiences immediate gains after investing in the ETF, they could overestimate their ability to select ETFs and underestimate ancillary market conditions, subsequently, giving the investor the confidence to continue trading. Moreover, if the price trend continues, the investor may believe that he was able to predict this trajectory (Odean, 1999).

3.5.2. Willows and West (2015)

Willows and West (2015) assess whether behavioural biases manifest differently in different genders. Using a sample period of January 2007 to December 2011, the trading behaviour of individual investors from a South African investment house is analysed. Results from the Wilcoxon rank-sum test suggest that males trade significantly more than females. According to Willows and West (2015), this finding is consistent with the findings of Barber and Odean (2001) who report that overconfidence is more prevalent in men than in women.

3.5.3. Dowie and Willows (2016)

Dowie and Willows (2016) examine the extent to which South African unit trust investors are overconfident. Their data collection process involved sending out surveys to South African university staff who invest in unit trusts. The Wilcoxon sign test is used to analyse the responses from the surveys in order to test for overconfidence among investors' estimates of their fund returns. The results indicate that investors underestimate their fund returns instead of overestimating their fund returns. Therefore, Dowie and Willows (2016) conclude that South African unit trust investors are under-confident as opposed to being overconfident, and therefore, the pessimism bias is present. A possible reason for their findings, according to the authors, is because, at the time of the study, the 2008 subprime financial crisis was still fresh in the minds of investors (Dowie and Willows, 2016). Additionally, the study finds that younger investors are less able to estimate their fund returns in comparison to older investors.

3.5.4. Charteris and Musadziruma (2017)

Charteris and Musadziruma (2017) study the prevalence of feedback trading in the market of South African futures contracts by analysing the price series of the JSE Top 40 and JSE Top 40 mini futures contracts. The study uses a sample period of 2006 to 2016 for the JSE Top 40 futures and a sample period of 2008 to 2016 for the JSE Top 40 mini futures. Using the Sentana and Wadhvani (1992) model, Charteris and Musadziruma (2017) report that feedback trading is not present in the JSE Top 40 and JSE Top 40 mini futures contracts. These findings hold even after controlling for the 2008 financial crisis. Thus, this absence of positive feedback trading in the South African futures market suggests that overconfident trading may not be present in South Africa's futures market.

3.5.5. Charteris and Rupande (2017)

Charteris and Rupande (2017) investigate the presence of feedback trading on the JSE. The study observes all stocks trading on the JSE's main board for the period 2004 to 2013. Using the empirical model proposed by Sentana and Wadhvani (1992), Charteris and Rupande (2017) find weak evidence of the existence of feedback trading on the JSE. Specifically, the results reveal that only 23% of the observed stocks display feedback trading. Nevertheless, negative feedback trading is found to be more pronounced than positive feedback trading. This low level of positive feedback trading could imply that overconfident trading is less pronounced on the JSE since overconfident trading lead to positive feedback trading.

3.6. Summary

Chapter 3 reviewed existing empirical research surrounding the presence of investor overconfidence in different asset classes as well as the effect of overconfidence on volatility. This was followed by a discussion of empirical studies that analysed South African securities and could be used to make inferences on the presence of investor overconfidence. Table 3.1 (page 49) summarises the empirical studies discussed in Chapter 3.

It is evident from Table 3.1 that the empirical research surrounding the presence of investor overconfidence is inconsistent. Whilst Statman, *et al.* (2006) and Metwally and Darwish (2015) find a strong presence of overconfident trading in stock markets, Zaiane and Abaoub (2009) report that weak evidence of the existence of overconfidence in the Tunisian stock market. Bailey, *et al.* (2011) find that mutual fund investors display overconfident trading behaviour, however, Dowie and Willows (2016) report that South African unit trust investors are underconfident as opposed to being overconfident. Aharon and Qadan (2018) find the presence of overconfidence in commodity markets, whilst Yung and Liu (2009) report that the overconfidence bias is not present in the commodity futures market.

Table 3.1: Summary of Studies Reviewed in This Chapter

Finding	Author(s)
Strong evidence that investor overconfidence is present in the respective market.	Statman, <i>et al.</i> (2006); Griffin, <i>et al.</i> (2007); Lin, <i>et al.</i> (2010); Bailey, <i>et al.</i> (2011); Metwally and Darwish (2015); Aharon and Qadan (2018); Gupta, <i>et al.</i> (2018); Chen and Sabherwal (2019); Alsabban and Alarfaj (2020).
Weak evidence that investor overconfidence is present in the respective market.	Zaiane and Abaoub (2009); Da Dalt, <i>et al.</i> (2019).
Investor overconfidence is not present in the respective market.	Yung and Liu (2009); Dowie and Willows (2016).
Trading volume induced by overconfidence exhibits a significant positive effect on volatility.	Chuang and Lee (2006); Abbas (2013); Jlassi, <i>et al.</i> (2014).
Trading volume induced by overconfidence exhibits an insignificant positive effect on volatility.	Sheikh and Riaz (2012).
Investor overconfidence has no significant relationship with volatility during the 2008 global financial crisis.	Abbas (2013).
A significant positive effect of overconfident trading on volatility before, during, and after the 2008 global financial crisis.	Jlassi, <i>et al.</i> (2014).
Overconfident trading insignificantly impacts market volatility after the 2008 global financial crisis.	Mushinada and Veluri (2018)

Regarding the effect of overconfident trading, Chuang and Lee (2006), Abbas (2013) and Jlassi, *et al.* (2014) report that trading volume induced by investor overconfidence exhibits a significant, positive effect on volatility. On the contrary, Sheikh and Riaz (2012) find that trading related to overconfidence shows an insignificant, positive effect on volatility. In addition, Jlassi, *et al.* (2014) document the prevalence of the positive relationship between overconfidence and volatility before, during, and after the 2008 subprime financial crisis. However, Abbas (2013) reports that trading induced by overconfidence exhibits no significant relationship with volatility during the 2008 global financial crisis. Additionally, Mushinada and

Veluri (2018) document that investor overconfidence does not significantly contribute to market volatility after the 2008 global financial crisis.

The mixed evidence surrounding the presence and effect of investor overconfidence highlights the need for further analysis relating to the overconfidence bias. Furthermore, the lack of evidence from an ETFs market perspective supports the current study. Specifically, the study conducted by Da Dalt, *et al.* (2019) is the only study, known to the author of this thesis that covers overconfident trading by ETF investors. Additionally, the lack of research focusing on the overconfidence bias in the South African market provides a basis for further analysis. Overall, the uniqueness of the present study is achieved through its application of the Statman, *et al.* (2006) model to a South African financial market, specifically, the ETF market, which, to the knowledge of the author, has never been done before.

CHAPTER 4: DATA AND METHODOLOGY

4.1. Overview

The previous chapter reviewed existing empirical evidence surrounding the overconfidence bias and the empirical methodologies employed by each study were outlined in an attempt to develop appropriate methods to answer the research questions of this study. The current chapter begins by describing the data used in this study, while the latter part of this chapter explains the empirical methodologies used to achieve this study's objectives. Specifically, the presence of market-wide investor overconfidence is investigated by examining the relationship between the current turnover of the market and its market returns using Vector Autoregressive (VAR) models. Thereafter, panel VAR models are used to determine whether investor overconfidence influences the trading activities of individual ETFs. Finally, Exponential Generalised Autoregressive Conditional Heteroskedasticity (EGARCH) models are estimated to analyse the effect of overconfident trading on volatility.

4.2. Dataset

The following sections provide a discussion of the various aspects relating to the data that is used in this study. This discussion includes, amongst other aspects; the sample period, data frequency, sample of ETFs, computation of the main variables, and the preliminary data analysis.

4.2.1. Sample Period

Patterns in trading behaviour can be detected more efficiently through the analysis of more extended periods of time. Furthermore, longer periods of study provide results that are more valid and more accurate and avoids the possibility of problems caused by data mining (Chuang and Lee, 2006). Therefore, in an attempt to provide the most reliable results, the analysis in this study is undertaken from the date of each ETF's inception until August 30, 2019. The first South African ETF was the Satrix 40 ETF which tracks the JSE Top 40 (J200) index and which was launched on November 27, 2000. Thus, the sample period for the market of South African ETFs tracking domestic benchmarks⁸ begins November 27, 2000 and ends August 30, 2019. Regarding the market of South African ETFs tracking international benchmarks⁹, the sample period begins October 10, 2005 and ends August 30, 2019 since the first South African ETF tracking an

⁸ Market of South African ETFs tracking domestic benchmarks refers to the portfolio of South African ETFs tracking domestic benchmarks.

⁹ Market of South African ETFs tracking international benchmarks refers to the portfolio of South African ETFs tracking international benchmarks.

international benchmark was launched on the 10th of October, 2005¹⁰. It is important to note that, as a new ETF is listed or delisted in the South African ETF market, its respective portfolio is rebalanced accordingly.

Additionally, the 2008 global financial crisis is employed as a structural break in an attempt to analyze the effect of trading volume induced by overconfidence on market volatility before, during, and after the 2008 global financial crisis. Accordingly, the full sample period is divided into subsamples of pre-, during-, and post- 2008 global financial crisis. Phillips and Yu (2011) acknowledge that the crisis became apparent during the third week of July 2007, while Trichet (2010) notes that the gradual phasing out of the crisis began December 03, 2009. Thus, following Horta, Lagoa and Martins (2016), August 01, 2007 is used as the crisis starting date and December 07, 2009 is used as the crisis ending date to allow for a gap of at least four days relative to the starting and ending dates of Phillips and Yu (2011) and Trichet (2010), respectively. The during-crisis starting date is further supported by Gerlach (2011), who argues that the crisis started in August 2007 when there was a drastic tightening of interbank liquidity. On this background, the during-crisis period starts 01 August 2007 and ends 07 December 2009 for both markets. Accordingly, the pre-crisis period for the market of JSE-listed ETFs replicating local benchmarks begins 27 November 2000 and ends 31 July 2007, while the pre-crisis period for the market of JSE-listed ETFs replicating offshore benchmarks begins 10 October 2005 and ends 31 July 2007. The post-crisis period begins 08 December 2009 and ends 31 August 2019 for both markets. Table 4.1 provides a summary of these subsamples.

Table 4.1: Summary of Subsamples

	Pre-crisis	During-crisis	Post-crisis
Market of ETFs Tracking Domestic Benchmarks	27 November 2000 to 31 July 2007	01 August 2007 to 07 December 2009	08 December 2009 to 31 August 2019
Market of ETFs Tracking International Benchmarks	10 October 2005 to 31 July 2007	01 August 2007 to 07 December 2009	08 December 2009 to 31 August 2019

4.2.2. Data Frequency

The use of monthly observations to detect the presence of investor overconfidence is supported by Odean (1999), Gervais and Odean (2001), Statman, *et al.* (2006), and Zaiane and Abaoub (2009) who claim that

¹⁰ The reason for not combining ETFs tracking domestic and international benchmarks into one market portfolio is discussed in Section 4.2.4.

fluctuations in investor overconfidence tend to be more evident over monthly horizons. As a result, this study employs a data frequency that is of a monthly nature in order to determine whether investor overconfidence is present in the South African ETF market using VAR models¹¹. This results in 226 monthly observations for the full sample period of the market¹² of ETFs replicating local benchmarks and 167 monthly observations for the entire sample period of the market of ETFs replicating international benchmarks.

A daily data frequency is employed to study the effect of investor overconfidence on the volatility of the South African ETF market. Charteris (2013) discovers that the mispricing of South African ETFs persists in the short-run only, specifically, deviations do not persist for more than two trading days. Moreover, Darrat, Rahman and Zhong (2003) argue that the volume-volatility relation should be examined using high-frequency data since financial markets exhibit high speeds of adjustment. However, Gwilym and Sutcliffe (2012) note that the use of intraday data can be problematic due to non-synchronous prices, stale prices, missing values, inaccurate times, inaccurate volumes as well as market microstructure effects at the intraday level. On this basis, the effect of trading volume induced by investor overconfidence on the volatility of the returns of the South African ETF market is assessed using high-frequency data, that is, daily data¹³. Therefore, following Abbes (2013) and Jlassi, *et al.* (2014), daily data is used to estimate the EGARCH empirical model that examines that effect of overconfidence on volatility¹⁴. Ledoit and Wolf (2008) argue that if daily data is used, GARCH-type models are more efficient in predicting volatility. Overall, the full sample period for the market of JSE-listed ETFs replicating domestic benchmarks has 4894 daily observations, while the entire sample period of the market of JSE-listed ETFs replicating international benchmarks has 3624 daily observations.

4.2.3. Survivorship Bias

According to Gilbert and Strugnell (2010), financial research that ignores failed assets and only includes assets that had the capacity to survive leads to a phenomenon known as survivorship bias. Since surviving assets possess different characteristics to that of delisted assets, the results of such studies could be biased (Gilbert and Strugnell, 2010). Therefore, to mitigate the issue of the survivorship bias, research on financial markets should include delisted assets in its investigation. Moreover, Pawley (2006) analyse the South

¹¹ The VAR empirical model used in this study is discussed further in Section 4.3.

¹² Due to the differing inception dates of individual ETFs, the number of monthly observations of each individual ETF will differ from that of its respective market.

¹³ Daily observations are based on a 5 day week.

¹⁴ A detailed discussion of the EGARCH model ensues in Section 4.5.

African unit trust industry and report that samples containing a survivorship bias tend to overstate the average performances of the sample. Similarly, Gilbert and Strugnell (2010) examine the JSE and find that the inclusion of delisted assets has a significant effect on the results obtained. Regarding investor overconfidence, the delisting or failure of an ETF may cause investors to lose confidence in the South African ETF market. Conversely, the survival of ETFs could cause investors to gain confidence in the South African ETF market. Thus, the survival or failure of ETFs could influence investors' confidence in the South African ETF market. This study, therefore, includes delisted South African ETFs to eliminate the negative influence of the survivorship bias.

4.2.4. Sample of ETFs

The analysis of this study is conducted on the market of South African ETFs tracking domestic (that is, South African) benchmarks as well as the market of South African ETFs tracking international benchmarks. The trading activity of ETFs replicating domestic benchmarks and of ETFs tracking international benchmarks may differ for several reasons. According to Johnson (2009), ETFs tracking international benchmarks exhibit higher tracking errors due to mismatched market trading times and exchange rate volatility. Consistently, Steyn (2019) finds that JSE-listed ETFs tracking domestic benchmarks track their benchmarks more efficiently relative to JSE-listed ETFs tracking international benchmarks because the returns of ETFs that follow offshore indices are constrained by the treatment of dividends, missed dividend income due to withholding taxes, market trading times that do not overlap with the benchmark index, and exchange rate volatility. Given that that tracking ability of JSE-listed ETFs replicating local benchmarks and JSE-listed ETFs replicating offshore benchmarks differs significantly, the confidence levels of ETF investors and traders may differ when dealing with JSE-listed ETFs replicating local benchmarks and JSE-listed ETFs replicating international benchmarks. Specifically, investors may be more overconfident when investing in JSE-listed ETFs replicating local benchmarks since these ETFs have a higher tracking ability in comparison to JSE-listed ETFs tracking international benchmarks. As such, ETFs tracking domestic and international benchmarks could exhibit different trading patterns, and therefore, are not combined into one market portfolio because this could distort the results of this study. Hence, this study's analysis of the South African ETF market is categorized into the market of JSE-listed ETFs replicating local (domestic) benchmarks and the market of JSE-listed ETFs replicating offshore (international) benchmarks.

To avoid the small sample problem, and thus, to provide reliable results, each ETF included in the surveyed markets has to satisfy the minimum observations requirement of this study. Following Gupta, *et al.* (2018), the minimum requirement for each security (that is, ETF) is to have at least 30 monthly observations. This minimum observations requirement applies to both listed and delisted ETFs, such that, the surveyed ETFs

must have been registered on the JSE for at least 30 months. Subsequently, this data requirement results in a sample that includes 49 of the 78 listed ETFs and 6 of the 9 delisted ETFs as at the 30th of August 2019. Therefore, the sample comprises of a total of 55 South African ETFs. For ease of reference, Table 4.2 provides a list of the ETFs included in the sample of each market. Appendix 1 (page 144) provides a comprehensive summary of all South African ETFs included in the sample and those that are excluded from the sample.

Table 4.2: Sample of ETFs

Domestic Portfolio			
Satrix 40 ETF	NewFunds Shari'ah Top 40 ETF	NewFunds GOVI ETF	CoreShares Top 50 ETF
Satrix FINI ETF	Ashburton Inflation ETF	NewFunds ILBI ETF	Krugerrand Custodial Certificate
Satrix INDI ETF	NewFunds S&P GIVI SA Financial 15 ETF	NewFunds TRACI 3 Month ETF	AfricaRhodium ETF
NewGold ETF	NewFunds S&P GIVI SA Industrial 25 ETF	CoreShares PrefTrax ETF	Satrix ILBI ETF
Satrix SWIX Top 40 ETF	NewFunds S&P GIVI SA Resource 15 ETF	Ashburton Mid Cap ETF	Satrix Property ETF
Satrix RESI ETF	Stanlib Top 40 ETF	Stanlib SA Property ETF	NewFunds Rand*
Satrix DIVI ETF	Stanlib SWIX 40 ETF	NewPlat ETF	Zshares Govi*
CoreShares PropTrax SAPY ETF	NewFunds MAPPS Growth ETF	AfricaPalladium ETF	BettaBeta Equally Weighted TOP40*
NewFunds S&P GIVI SA Top 50 ETF	NewFunds MAPPS Protect ETF	NewPalladium ETF	BettaBeta CIS BGreen Portfolio*
Ashburton Top 40 ETF	CoreShares PropTrax Ten ETF	AfricaGold ETF	CoreShares Low Volatility ETF*
Satrix RAFI 40 ETF	NewFunds SWIX ETF	AfricaPlatinum ETF	Coreshares Equally Weighted Top 40 ETF*
NewFunds NewSA ETF	NewFunds Equity Momentum ETF	CoreShares DivTrax ETF	
International Portfolio			
Sygnia DJ EuroStoxx 50 ETF	Sygnia MSCI Japan ETF	Sygnia MSCI World ETF	Coreshares S&P Global Property ETF
Sygnia FTSE 100 ETF	Sygnia MSCI USA ETF	CoreShares S&P 500 ETF	Dollar Custodial Certificate - 10 Year

Notes:

1. Table compiled on the 30th of August 2019.
2. * denotes delisted ETFs.

4.2.5. Types of Data and Sources of Data

To meet the objectives of this study, the secondary data required are:

- Daily and monthly closing prices of all ETFs.
- Dividends paid for each ETF.
- Daily and monthly number of shares traded for each ETF.
- Daily and monthly total number of outstanding shares for each ETF.

Except for the dividend data, all secondary data used in this study is obtained from the IRESS database. Data for the dividends paid on each ETF is obtained from the Infront Analytics database.

4.2.6. Computation of Main Variables

To answer research question one, the presence of market-wide¹⁵ investor overconfidence is examined by employing basic VAR models to study the long-run association between the current trading activity of the market and its past market returns while controlling for market volatility and dispersion. This study uses turnover as a proxy for trading activity, and therefore, research question one requires the computation of monthly market turnover, market return, market volatility and market dispersion. For research question two, panel VAR models are employed to determine whether the overconfidence bias influences the trading activities of individual ETFs while controlling for individual ETF return volatility. As such, research question two requires the computation of the monthly turnover, return, and volatility of each individual ETF in addition to its respective monthly market return. Lastly, research questions three and four also require the computation of market return and market turnover, but on a daily frequency. The following subsections outline the computation of each variable used in this study.

4.2.6.1. Security and Market Turnover

Previous researchers have measured trading activity using trading volume (in shares) or the turnover ratio. However, according to Statman, *et al.* (2006), an examination of trading activity has to take into consideration that the number of outstanding shares of each ETF fluctuates over time. Hence, for research question one, two, three and four, this study employs the turnover ratio as a proxy of trading activity.

Following Statman, *et al.* (2006), market turnover is the average turnover of all ETFs in the market at time t and is aggregated based on the market-value weight of each ETF at time t . The ETFs included in the

¹⁵ Market relates to the respective ETF market, that is, either the market of JSE-listed ETFs replicating local benchmarks or the market of JSE-listed ETFs replicating offshore benchmarks.

value-weighted market portfolio are continuously rebalanced to account for any new ETFs listed or delisted in the respective South African ETF market. Therefore, the market turnover ($mturn_{m,t}$) of each portfolio is calculated in the following manner:

$$mturn_{m,t} = \sum_{i=1}^N w_{i,t} turn_{i,t} \quad (4.1)$$

In equation 4.1, $mturn_{m,t}$ denotes the market turnover of the respective market for month t . N denotes the total number of ETFs in the market at the end of month t and $turn_{i,t}$ signifies the turnover of the i th ETF during month t , which is calculated using equation 4.3. For month t , the weight, $w_{i,t}$, of i th ETF is computed as the market capitalisation of the i th ETF divided by the combined capitalisation of all ETFs in the respective market, such that:

$$w_{i,t} = \frac{P_{i,t} S_{i,t}}{\sum_{i=1}^N P_{i,t} S_{i,t}} \quad (4.2)$$

In equation 4.2, $S_{i,t}$ is the total number of outstanding shares for the i th ETF at the end of month t and $P_{i,t}$ is the ex-dividend price of the i th ETF at the end of month t . $P_{i,t} S_{i,t}$ represents the monthly market capitalisation of i th ETF. Additionally, in equation 4.2, N denotes the total number of ETFs in the respective market at the end of month t .

The turnover, $turn_{i,t}$, of each individual ETF is computed using the computation proposed by Lo and Wang (2000), such that:

$$turn_{i,t} = \frac{V_{i,t}}{S_{i,t}} \quad (4.3)$$

In equation 4.3 above, $turn_{i,t}$ represents the turnover of the i th ETF for month t . $V_{i,t}$ denotes the number of shares traded for the i th ETF during month t and $S_{i,t}$ is as defined in equation 4.2.

Overall, for research question one, the monthly market turnover is as defined in equation 4.1, and for research question two, the individual ETF turnover is as defined in equation 4.3. For research questions three and four, the daily market turnover of each market is computed in the same manner as in equation 4.1, however, daily observations are used and not monthly observations which are used in equation 4.1. This being so, the individual ETF turnover and the weight of each ETF is calculated using daily data observations. The daily market turnover series for research questions three and four are denoted as V_{τ} where V_{τ} represents the daily market turnover on day τ .

4.2.6.2. Security and Market Returns

The market VAR models estimated for research question one requires the computation of the monthly market return for each ETF market. Additionally, the panel VAR models employed for research question

two requires the calculation of the monthly return of each individual ETF in the respective market. The return from an ETF can be calculated using simple returns or continuous compounding returns. Compounded returns contain a non-linear transformation, and thus, the compounded returns of individual assets cannot be aggregated to compute portfolio returns (Brooks, 2008). Hence, Liu and Strong (2008) propose that the return from a portfolio is computed as the weighted average of the simple returns of each asset in the portfolio. Therefore, for research question one, simple returns are computed for each ETF since the returns are aggregated to calculate the return of the value-weighted market portfolio.

Following Statman, *et al.* (2006), Liu and Strong (2008), Meier (2018) and Alsabban and Alarfaj (2020), the ETF's daily closing prices and dividends are used to compute their simple dividend-adjusted returns in the following manner:

$$r_{i,\tau} = \frac{P_{i,\tau} - P_{i,\tau-1} + D_{i,\tau}}{P_{i,\tau-1}} \quad (4.4)$$

In equation 4.4 above, $r_{i,\tau}$ denotes the simple return of the i th ETF on day τ . $P_{i,\tau}$ and $P_{i,\tau-1}$ represent the current and previous ETF closing prices for day τ , while $D_{i,\tau}$ represents the dividend paid on day τ . If there are no dividends paid on day τ then $D_{i,\tau}$ is equal to zero. The simple daily returns, $r_{i,\tau}$, of each ETF are then used to compute the daily returns of the value-weighted portfolio.

The simple return of the value-weighted portfolio is computed using the following specification:

$$r_{m,\tau} = \sum_{i=1}^N w_{i,\tau} r_{i,\tau} \quad (4.5)$$

In equation 4.5 above, $r_{m,\tau}$ represents the simple return on the market portfolio for day τ and $r_{i,\tau}$ is defined as in equation 4.4. Additionally, N denotes the total number of ETFs in the respective market on day τ and $w_{i,\tau}$ represents the weight of each ETF which is calculated in a similar manner to equation 4.2 but using daily values.

The daily market return, $R_{m,\tau}$, series of each portfolio is then derived by transforming the simple return calculated in equation 4.5 into a log return, using the following computation:

$$R_{m,\tau} = \ln(r_{m,\tau} + 1) \quad (4.6)$$

In equation 4.6 above, \ln denotes the natural logarithm and $r_{m,\tau}$ is as defined in equation 4.5. For research question one, the daily market returns, $R_{m,\tau}$, are aggregated to calculate the value of the monthly market returns, $mret_t$. This process is followed for both the market of JSE-listed ETFs replicating local benchmarks as well as the market of JSE-listed ETFs replicating offshore benchmarks. Notably, the daily market return for research questions three and four follow the definition given in equation 4.6 above.

For the individual security panel VAR models employed to answer research question two, the return of individual ETFs is used as an endogenous variable. In this case, since the return is computed for a single asset (ETF) and not a portfolio, continuous compounding returns are calculated. The benefit of using compounded returns for individual assets that are not combined into portfolios is that the returns are more statistically tractable (Tsay, 2005). Accordingly, the daily returns for each ETF are calculated using the following equation:

$$etfret_{i,\tau} = \ln\left(\frac{P_{i,\tau} + D_{i,\tau}}{P_{i,\tau-1}}\right) \quad (4.7)$$

In equation 4.7, $etfret_{i,\tau}$ represents the return on the i th ETF for day τ . The natural logarithm is denoted by \ln while $P_{i,\tau}$, $P_{i,\tau-1}$ and $D_{i,\tau}$ follows the definition prescribed in equation 4.4. For research question two, each ETF's daily returns, $etfret_{i,\tau}$, are combined to calculate the value of its monthly return, $ret_{i,t}$, where $ret_{i,t}$ signifies the return on the i th ETF for month t .

4.2.6.3. Security and Market Volatility

For research question one, market volatility, $msig_t$, refers to the monthly realized volatility of the returns from the value-weighted ETF market portfolio. The volatility of the market returns are computed using the method proposed by French, Schwert and Stambaugh (1987) and is calculated using the month's daily returns, after correcting for realized autocorrelation:

$$msig_{m,t}^2 = \sum_{\tau=1}^Z R_{m,\tau}^2 + 2 \sum_{\tau=1}^Z R_{m,\tau} R_{m,\tau-1} \quad (4.8)$$

In equation 4.8, $R_{m,\tau}$ is day τ 's market return as defined in equation 4.6 and Z is the number of trading days in month t . Additionally, for research question two, individual ETF volatility, $sig_{i,t}$, relates to the monthly realized volatility of the returns of each individual ETF.

Using the method prescribed by French, *et al.* (1987), individual ETF volatility, $sig_{i,t}$, is computed using the following equation:

$$sig_{i,t}^2 = \sum_{\tau=1}^Z etfret_{i,\tau}^2 + 2 \sum_{\tau=1}^Z etfret_{i,\tau} etfret_{i,\tau-1} \quad (4.9)$$

In equation 4.9, $etfret_{i,\tau}$ refers to day τ 's return on the i th ETF as defined in equation 4.7 and Z is the number of trading days in month t .

4.2.6.4. Market Dispersion

Dispersion relates to the cross-sectional standard deviation of the market returns. Following Statman, *et al.* (2006), for the VAR models estimated to answer research question one, the monthly dispersion for each market is measured as in equation 4.10 (page 60).

$$disp_{m,t} = \sqrt{\frac{\sum_{\tau=1}^Z (R_{m,\tau} - \mu_t)^2}{Z-1}} \quad (4.10)$$

In equation 4.10, $disp_{m,t}$ denotes the cross-sectional standard deviation of the market returns for month t . $R_{m,\tau}$ is day τ 's market return as defined in equation 4.6, μ_t is the average daily return ($R_{m,\tau}$) during month t and Z is the number of trading days in month t .

4.2.7. Preliminary Data Analysis

4.2.7.1. Unit Root Tests

An important requirement of the empirical models used in this study is that all time series meet the condition of stationarity. The stationarity of market turnover, market return, market volatility and market dispersion (both daily and monthly) is examined using the Phillips-Perron (henceforth, PP) unit root test. However, the individual security version of the VAR empirical model (used to ascertain the overconfidence effect) is of a panel nature, and thus, a panel unit root test is employed to examine the stationarity of the panel variables (that is, security turnover, security return, and security volatility). The panel unit root test used in this study is the Choi panel unit root test. The unit root tests are discussed in the following subsections.

a. Phillips-Perron Unit Root Test

The stationarity of each time series relating to market variables is examined using the PP unit root test. A key advantage of the PP test, developed by Phillips and Perron (1988), is that the test is more robust than the augmented Dickey-Fuller test since it employs Newey-West (1987) heteroscedasticity- and autocorrelation-consistent standard errors (Escobari, Garcia and Mellado, 2017). The test is subject to the following null and alternative hypotheses:

$H_0 =$ *There exists a unit root in the univariate time series.*

$H_1 =$ *There exists stationarity or a deterministic trend in the univariate time series.*

Eviews presents the PP test statistic, Z_α , based on the equation given below (Eviews, 2019):

$$Z_\alpha = t_\alpha \left(\frac{\gamma_0}{f_0} \right)^{\frac{1}{2}} - \frac{T(f_0 - \gamma_0)(s\epsilon(\hat{\alpha}))}{2f_0^{\frac{1}{2}}s} \quad (4.11)$$

In equation 4.11, T denotes the sample size, $\hat{\alpha}$ represents the estimate, t_α is the t-ratio of α and $s\epsilon(\hat{\alpha})$ is the coefficient standard error. γ_0 is the estimate of error variance, f_0 represents an estimate of the residual spectrum computed at frequency zero, and s represents the standard error of the test regression. The PP unit root test follows the critical values provided by Fuller (1976). Additionally, Eviews reports probability values for the PP test statistic. The null hypothesis of a unit root in the univariate time series is rejected

when the associated p-value is less than the specified level of significance. If the null hypothesis is rejected, the univariate time series is considered stationary. Contrarily, when the null hypothesis is accepted, the series is deemed to be non-stationary. If a time-series is found to be non-stationary, the natural log transformation is used to eliminate the correlation amongst the trend and the volatility of the trend (Statman, *et al.*, 2006; Beard, Marsden, Brown, Tombor, Stapleton, Michie and West, 2019).

b. Choi Unit Root Test

The Choi panel unit root test developed by Choi (2001) combines the probability values from individual time series unit root tests. Notably, the main limitation of alternative panel root tests (such as the Levin, Lin and Chu (2002) and Im, Peseran and Shin (2003) panel unit root tests) is that the alternative hypothesis of these tests hypothesize that none of the individual series in the panel contains a unit root. In contrast, the main advantage of the Choi panel unit root test is that its alternative hypothesis hypothesizes that some series in the panel are stationary while other series are non-stationary (Baltagi, Bresson and Pirotte, 2007; Hurlin, 2010). Furthermore, according to Baltagi, *et al.* (2007), Choi's test for a panel unit root assumes that the individual time series spans are different for all series in the panel. Thus, the Choi panel unit root test is appropriate for the unbalanced panel data used in this study.

Choi's (2001) test is subject to the following null and alternative hypotheses (assuming a finite number of individual series):

$$H_0 = \text{All time series in the panel are unit root non – stationary.}$$

$$H_1 = \text{At least one time series in the panel is non – stationary whilst other are not.}$$

Choi's (2001) panel unit root test calculates the following test statistics:

$$P = -2 \sum_{i=1}^N \ln (p_i) \tag{4.12}$$

$$Z = \frac{1}{\sqrt{N}} \sum_{i=1}^N \Phi^{-1}(p_i) \tag{4.13}$$

In equations 4.12 and 4.13 above, P represents the inverse Chi-square test statistics and Z represents the inverse normality test statistic. Additionally, p_i denotes the p-value derived from any individual unit root test such as the Augmented Dickey-Fuller (ADF) or PP tests and Φ^{-1} denotes the inverse of the standard normal cumulative distribution function. In this study, p_i is obtained from the PP unit root test since the PP test is the individual unit root test employed in this study due to the advantages discussed in Section 4.2.7.1 (a).

According to Choi (2006), the null hypothesis of all series being unit root non-stationary is rejected against the alternative hypothesis at a level of significance equal to α when the following inequality holds:

$$P > c_{p\alpha} \quad (4.14)$$

$$Z > c_{z\alpha} \quad (4.15)$$

In equation 4.14 and 4.15, $c_{p\alpha}$ represents the critical value of the upper tail of the chi-square distribution and $c_{z\alpha}$ represents the critical value of the lower tail of the normal distribution.

4.2.7.2. Descriptive Statistics

The descriptive statistics that are analysed include the following:

a. Mean

The mean value is the most commonly used measure of central tendency as it represents the average value (Thompson, 2009). The mean value is calculated by dividing the sum of the value of all observations by the total number of observations.

b. Standard Deviation

Standard deviation measures the average distance from the mean that each observation lies (Deakin and Kildea, 1999). As such, standard deviation measures dispersion and, therefore, is regarded as a measure of the level of inherent risk. A small standard deviation suggests a lower level of inherent risk since the observations lie closer to the mean. Equation 4.16 outlines the computation of standard deviation:

$$\text{Standard Deviation } (\sigma) = \sqrt{\frac{\sum(X - \bar{X})^2}{N-1}} \quad (4.16)$$

In equation 4.16, X is a random observation in the series, \bar{X} is the mean value of the series and N is the total number of observations in the series.

c. Skewness

Skewness measures the distributional asymmetry of a series (Kim and White, 2004). Specifically, skewness provides an indication of the side of the distribution that has a longer tail. A series that is normally distributed has a skewness value of zero. The value of skewness is calculated as depicted in equation 4.17 below:

$$\text{Skewness } (S) = \frac{1}{N} \sum_{i=1}^N \left(\frac{X - \bar{X}}{\hat{\sigma}} \right)^3 \quad (4.17)$$

where X , \bar{X} and N follows the definition as in equation 4.16. $\hat{\sigma}$ is an estimator of the sample's standard deviation based on a biased approximation of the variance, calculated as in equation 4.18 (page 63).

$$\hat{\sigma} = \sigma \sqrt{\frac{(N-1)}{N}} \quad (4.18)$$

where σ is defined in equation 4.16.

d. Kurtosis

Kurtosis is a measure of how peaked or flat a series distribution is (Ruppert, 1987). The kurtosis value is computed as follows:

$$\text{Kurtosis } (K) = \frac{1}{N} \sum_{i=1}^N \left(\frac{X_i - \bar{X}}{\hat{\sigma}} \right)^4 \quad (4.19)$$

In equation 4.19, X , \bar{X} , N and $\hat{\sigma}$ is as previously defined. A normal distribution has a kurtosis value of 3, thus, a kurtosis value that is greater (less) than 3 indicates that the series distribution is peaked (flat).

e. Jarque-Bera

The Jarque-Bera (hereafter, JB) test is a goodness-of-fit test that examines a series departure from normality, based on its skewness and kurtosis (Jarque and Bera, 1980). The JB test statistic is computed using the following equation:

$$\text{JB test statistic} = \frac{\eta}{6} \left(S^2 + \frac{(K-3)^2}{4} \right) \quad (4.20)$$

In equation 4.20, η denotes the number of coefficients used to create the series, while S and K represent the series skewness and kurtosis values, respectively. Under the null hypothesis of a normally distributed series, the JB test follows a chi-square distribution with 2 degrees of freedom. Additionally, the probability value (p-value) associated with the JB test statistic can be used to determine whether the null hypothesis is rejected. Accordingly, the null hypothesis of a normal distribution is rejected when the specified level of significance is greater than the associated probability value.

4.2.7.3. Correlation Analysis

According to Rodgers and Nicewander (1988), the Pearson correlation coefficient measures the linear association between variables. Therefore, to examine the correlation between different variables, this study computes the Pearson correlation coefficient, ρ , as follows:

$$\rho = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\left[\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2 \right]^{1/2}} \quad (4.21)$$

In equation 4.21, n is the total number of observations, X_i is the i th observation of the X variable and Y_i is the i th observation of the Y variable. Additionally, \bar{X} is the mean of the X observations and \bar{Y} is the average of the Y observations. The correlation coefficient ranges from -1 to +1. A positive correlation between two variables implies that the variables move in the same direction, that is, as one variable increases (decreases),

the other variable also increases (decreases). Conversely, a negative correlation between two random variables suggests that the variables move in opposite directions, such that, as one variable increases (decreases), the other variable decreases (increases). Conversely, if the movements in the variables have no association, the correlation coefficient is equivalent to zero.

A high correlation between the explanatory variables could lead to a phenomenon referred to as multicollinearity. Multicollinearity occurs when there is a high correlation between a regression's independent variables and could lead to coefficient estimates with signs or magnitudes that are implausible (Farrar and Glauber, 1967; Shieh, 2010; Thompson, Kim, Aloe and Becker, 2017). However, the presence of near multicollinearity amongst independent variables does not violate the Best Linear Unbiased Estimator (BLUE) properties of OLS estimations, and therefore, estimators will still be consistent, unbiased, and efficient (Shieh, 2010).

4.2.7.4. Model Diagnostics Tests

Diagnostics tests provide evidence on the appropriateness of the assumptions made in fitting a model (Weisberg, 1983; Atkinson, 1986). Thus, to ensure that the estimated models are efficient, unbiased, and reliable, diagnostics tests must be conducted. Model diagnostics tests can include informal procedures (such as inspecting a plot of the residuals) or more formal tests of the underlying assumptions, particularly, residual autocorrelation, heteroskedasticity, and non-normality. However, in this analysis, the model diagnostics tests for the VAR model, include; the Breusch–Godfrey Lagrange Multiplier test (henceforth, LM test) and the White heteroskedasticity test. Tests for non-normality are not conducted since normality is not necessary for the asymptotic validity of VAR-related statistical procedures (Kilian and Demiroglu, 2000). Given that research questions one and two are linked, diagnostics tests are conducted only on the estimated market VAR models (discussed under Section 4.3.1) since the respective panel VAR models (discussed under Section 4.4.2) are estimated using individual ETFs that are included in the computation of the market variables.

The reliability of the EGARCH models estimated for research questions three and four are assessed by running diagnostic tests for ARCH effects since the conditional variance equations are correctly specified when there are no ARCH effects present in the residuals (Serletis, 2007: 236). Notably, research questions three and four are linked, and therefore, the diagnostics test for ARCH effects are only conducted on the EGARCH models estimated for the full sample because the EGARCH models estimated for the sub-period analysis are estimated using subsamples of the full sample period. The following sections outline the various diagnostics tests used in this study.

a. Test for Serial Correlation

According to Mantalos and Shukur (2005), one of the most important tests for serial correlation is the Breusch-Godfrey test put forward by Breusch (1978) and Godfrey (1978). The Breusch-Godfrey test statistic for serial correlation at lag order h is computed by running an auxiliary regression of the residuals ε_t on the original right-hand regressors and the lagged residuals ε_{t-1} where the missing first h values of ε_{t-h} are equal to zero (Charemza and De Adman, 1997). The R^2 from the auxiliary regression is then used to compute the LM test statistic, such that:

$$\text{LM test statistic} = (T - h)R^2 \tag{4.22}$$

In equation 4.22, T represents the number of observations in the basic series, h denotes the number of lags of the error term and R^2 denotes the coefficient of determination from the auxiliary regression. The LM test statistic¹⁶ follows the chi-square distribution, and thus, the null hypothesis of no serial correlation at lag h is rejected when the LM test statistic is greater than its critical chi-square value (Brooks, 2002: 166).

b. Test for Heteroskedasticity

The White (1980) test is employed to test for heteroskedasticity. The regression used in the test regresses each cross product of the residuals on the cross products of the regressors (Simo-Kengne, Balcilar, Gupta, Reid and Aye, 2013). Thereafter, the joint significance of the regression is tested. By employing the no cross-terms option, the regressors for the test's regressions include the levels and squares of the original explanatory variables as well as a constant term. Eviews presents the LM chi-square statistic for the joint significance of all the regressors in the system of the test regression. Under the null hypothesis of no heteroskedasticity, the LM test statistic follows a chi-square distribution (Simo-Kengne, *et al.*, 2013).

c. Test for ARCH effects

According to Engle (1982), OLS estimators are not consistent when ARCH effects are present in the residuals. Thus, when the variance equation is correctly specified, there should be no ARCH effects in the residuals (Lundbergh and Teräsvirta, 2002). Engle (1982) proposed a LM test to identify the presence of ARCH effects in the residuals of the estimated model. The test statistic is given by (Demos and Sentana, 1998):

$$\xi = TR^2 \tag{4.23}$$

where T is the number of observations and R^2 is the coefficient of determination from an auxiliary regression of the squared residuals on lagged squared residuals up to order q and a constant. Under the null

¹⁶ It is important to note that the Eviews statistical software package computes the likelihood ratio (LR) version of the LM test to account for Edgeworth expansion correction. This modified test statistic is denoted as LRE on Eviews.

hypothesis of no ARCH effects, the LM test statistic, ξ , follows a chi-square distribution with q degrees of freedom (Demos and Sentana, 1998). The null hypothesis is accepted when the test statistic is less than the critical chi-square value. Following Bollerslev (1986), in this study, ARCH effects up to lag order 8 is tested, thus, q equals 8.

4.3. Method Used to Detect Investor Overconfidence

In an attempt to answer the first research question of this study, the association between current trading activity and lagged market returns is inspected using VAR models. Section 4.3.1 outlines the different components of the VAR models used in this analysis.

4.3.1. Market VAR Models

There are two markets in this study's analysis, specifically, the market of JSE-listed ETFs replicating local benchmarks as well as the market of JSE-listed ETFs replicating offshore benchmarks. Thus, the VAR model discussed in this section is estimated individually for both markets in an attempt to answer research question one. VAR models, first proposed by Sims (1980), are used to estimate multiple equations simultaneously, unlike univariate time series models. The primary advantage of VAR models is that they do not require specifications of which variables are endogenous or exogenous; however, theoretical models may impose restrictions on variables in order to separate the influence of exogenous shocks on the VAR system (Abrigo and Love, 2016). It is important to note that the values of exogenous variables are determined outside the VAR system. Another key advantage of VAR models is that the autoregressive and lag terms in the system may help to uncover more features of the data, which makes this modelling technique essential in understanding the interactions between the variables.

The overconfidence hypothesis proposed by Gervais and Odean (2001) posits that overconfident investors associate positive market returns with their personal ability to process information and identify undervalued stocks, and thus, trade more frequently in periods after market gains. In other words, an increase in market gains result in an increase in investors' confidence, and consequently, traders trade more excessively after periods of positive market returns. Therefore, in the presence of investor overconfidence, historical market returns influence current trading activity positively. As such, the model proposed by Statman, *et al.* (2006) investigates the presence of investor overconfidence by examining how historical market returns relates to current trading activity using VAR models and the impulse response functions associated with each VAR model. The Statman, *et al.* (2006) overconfidence model has proven to be a seminal work, with Statman's (2006) method being employed by several subsequent studies on the topic (including, Griffin, *et al.*, 2007;

Zaiane and Abaoub, 2009; Lin, *et al.*, 2010; Metwally and Darwish, 2015; Zia, *et al.*, 2017; Gupta, *et al.*, 2018; Alsabban and Alarfaj, 2020).

Following Statman, *et al.* (2006), this study detects the presence of investor overconfidence by estimating VAR models which follow equation 4.24 below¹⁷:

$$\begin{bmatrix} mturn_t \\ mret_t \end{bmatrix} = \begin{bmatrix} \alpha_{mturn} \\ \alpha_{mret} \end{bmatrix} + \sum_{p=1}^P A_p \begin{bmatrix} mturn_{t-p} \\ mret_{t-p} \end{bmatrix} + \sum_{s=0}^S B_s \begin{bmatrix} msig_{t-s} \\ disp_{t-s} \end{bmatrix} + \begin{bmatrix} \varepsilon_{mturn,t} \\ \varepsilon_{mret,t} \end{bmatrix} \quad (4.24)$$

where $mturn_t$ represents the market turnover for month t calculated using equation 4.1 and $mret_t$ is the market return for month t which is computed by aggregating the daily returns calculated using equation 4.6. Additionally, in equation 4.24, $msig_t$ represents the market return volatility for month t which is calculated using equation 4.8 and $disp_t$ represents month t 's market return dispersion as defined in equation 4.10. In equation 4.24, α is a vector of intercepts, ε is a vector of residuals, and A_p and B_s represent regression coefficients. The endogenous variables are $mturn$ and $mret$ while $msig$ and $disp$ are the exogenous variables. P and S denote the optimal lag lengths of the endogenous and exogenous variables, respectively. Since formal theories relating to overconfidence do not prescribe a time period for the lead-lag association between current turnover and past returns, optimal lag lengths are selected using the information criteria and likelihood ratio, which is discussed in section 4.3.1.1.

The VAR model presented in equation 4.24 is developed on the basis that investor overconfidence can be detected by examining the interaction between current trading activity and lagged market returns. However, Statman, *et al.*, (2006) argues that any analysis of trading activity must take into consideration that the number of outstanding shares changes over time. Therefore, consistent with Statman, *et al.* (2006), this study employs turnover as a proxy of trading activity. Subsequently, the market VAR model in equation 4.24 employs market turnover and market returns as endogenous variables. The finding of a positive and significant association between the current turnover of the market and past market returns provides evidence of the presence of overconfident trading by investors in the respective market (Odean, 1998; Gervais and Odean, 2001; Statman, *et al.*, 2006). Notably, the association between lagged market returns and current market turnover is positive and statistically significant when the $mturn$ equation in the estimated VAR model exhibits positive and statistically significant lagged $mret$ coefficients. The significance of this relationship is determined using the t-test, specifically, by examining the test statistics and probability values obtained for each coefficient. However, to ensure that the model is reliable, the model needs to consider alternative factors that influence trading activity.

¹⁷ Equation 4.24 is the same equation 3.3 but has been repeated for ease of reference.

Consistent with the Statman, *et al.* (2006) model, the market VAR model used in this study includes market volatility and dispersion as control variables in an attempt to account for alternative explanations of trading activity. Market volatility, *msig*, is the first control variable and is included to account for the relationship between trading volume and volatility. This volume-volatility relationship has several explanations, including the Mixture of Distribution hypothesis and the Sequential Arrival of Information hypothesis. The Mixture of Distribution hypothesis introduced by Clark (1973) posits that price changes and changes in trading volume are induced by the same underlying information arrival process. Consequently, the volatility of prices and trading volume have a contemporaneous relationship. On the contrary, the Sequential Arrival of Information hypothesis proposed by Copeland (1976) argues that information disseminates sequentially. Thus, variability in returns have the ability to predict trading volume, and vice versa. Subsequently, the Sequential Arrival of Information hypothesis claims that there is a lead-lag interrelation between the volatility of returns and trading activity. Dispersion (*disp*) is employed as the second control variable in the market VAR model to account for potential activities in trading that may be due to investors rebalancing their portfolios. For example, according to Sheikh and Riaz (2012), as the spread between individual security returns widens, investors may engage in trading activities as they seek to maintain fixed portfolio weights.

Prior to the estimation of the market VAR models, the preliminary data analysis involves checking the stationarity of each series by using PP unit root test. To ensure that the estimated market VAR models are reliable and unbiased, diagnostics tests are performed on the estimated models as discussed under Section 4.2.7.4. Finally, research question one is answered by examining the results of the estimated VAR models, Granger causality tests, and their associated impulse response functions.

4.3.1.1. Optimal Lag Order Selection

The optimal lag order of the endogenous variables is selected by inspecting the information criteria and likelihood ratio (LR henceforth). The information criteria employed in this study, include; the Akaike's (1973) Information Criterion (AIC), Schwarz (1978) Information Criteria (SC), and Hannan and Quinn (1979) criterion (HQ). The computation of the information criteria is defined below:

$$AIC = \log|\hat{\Sigma}| + 2\hat{k}/T \quad (4.25)$$

$$SC = \log|\hat{\Sigma}| + \frac{\hat{k}}{T} \log(T) \quad (4.26)$$

$$HQ = \log|\hat{\Sigma}| + \frac{2\hat{k}}{T} \log(\log(T)) \quad (4.27)$$

In equations 4.25, 4.26, and 4.27 above, $\hat{\Sigma}$ is the covariance matrix of the residuals, \hat{k} denotes the total number of parameters in the VAR system and T represents the total number of observations. The

information criteria are computed for lags 0 to \bar{k} (in this study, $\bar{k} = 8$)¹⁸ and the optimal lag length is the number of lags that minimizes the value of the information criterion. Following Zia, *et al.* (2017), when the information criteria are not minimized at the same lag length, the optimal lag length is selected based on the AIC. This is because the AIC tends to suggest more lags. Therefore, the higher number of lags provide a better insight into the VAR structure and helps to uncover more features of the data (Lin, *et al.*, 2010). On the contrary, the SC imposes a larger penalty for additional lag coefficients, and therefore, tends to suggest fewer lags.

In addition to the information criteria, the LR test is used to select the optimal lag lengths, and is calculated as follows (Hatemi and Hacker, 2009):

$$LR = T|\log|\widehat{\Sigma}_r| - \log|\widehat{\Sigma}_u|| \quad (4.28)$$

In equation 4.28, T denotes the sample size, $|\widehat{\Sigma}_r|$ denotes the covariance matrix of the residuals of the restricted VAR, and $|\widehat{\Sigma}_u|$ denotes the covariance matrix of the residuals of the unrestricted model. The LR test statistic follows the chi-square critical values with degrees of freedom equal to the total number of restrictions (Hatemi and Hacker, 2009). LR test statistics are computed for each lag, and the optimal lag length is the last lag that rejects the null hypothesis which hypothesises that all elements in the coefficient matrix are insignificant (Hatemi and Hacker, 2009). Specifically, the null hypothesis that all elements in the coefficient matrix are insignificant is rejected when the LR test statistic is greater than its critical value. In this study, the information criteria and likelihood ratio test statistics for the market VAR models will be computed using the Eviews statistic package.

Following Alsabban and Alarfaj (2020), the optimal lag length of the endogenous variables is selected using the information criteria and likelihood ratio of the VAR model which imposes no lag on the exogenous variables. Thereafter, the optimal lag length of the exogenous variables is selected by constructing VAR models which include the chosen optimal lag length of the endogenous variables and an addition of lagged exogenous variables (Alsabban and Alarfaj, 2020). Specifically, the first model begins with 1 lag of each exogenous variable and each subsequent model includes an additional lag of the exogenous variables. This process is repeated until the model has 8 lags of the exogenous variables. The AIC and SC of these models are then compared to select the optimal lag length of the exogenous variables that minimises the information criteria. In choosing the optimal lag length for the exogenous variables, only AIC and SC is used because Eviews does not provide HQ and LR statistics for the estimated VAR regressions.

¹⁸ According to Narayan, Narayan and Prasad (2008), using a maximum of 8 lags for the unit root tests is standard practice.

4.3.1.2. Granger Causality Tests

The overconfidence hypothesis suggests that the overconfidence effect is present in a market when lagged market returns can explain the current market turnover. Granger causality tests are performed to ascertain the causal relationship between market turnover and market returns. Granger (1969) defines causality as a condition in which the historical values of one variable can be used to predict another variable. Following Granger (1969), the Granger causality test used in this analysis assumes a simple causal model:

$$mturn_t = \sum_{j=1}^m \alpha_j mturn_{t-j} + \sum_{j=1}^n \beta_j mret_{t-j} + \varepsilon_t \quad (4.29)$$

$$mret_t = \sum_{j=1}^m c_j mturn_{t-j} + \sum_{j=1}^n d_j mret_{t-j} + \epsilon_t \quad (4.30)$$

In equation 4.29 and 4.30, the error terms, ε_t and ϵ_t , represent two uncorrelated white-noise series. $mturn_t$ represents the market turnover at period t while $mret_t$ represents the market return at period t .

Based on the definition of causality, $mret$ is Granger causing $mturn$ if some β_j are not zero. Likewise, $mturn$ is causing $mret$ if some c_j are not zero. If both these events occur, there exists a feedback causality between $mret$ and $mturn$. Conversely, if there is no causality in any direction, $mret$ and $mturn$ exhibit independence. In this study, Granger causality tests are used to confirm whether there is a significant causal relationship from market return to market turnover. This is because Granger causality tests demonstrate the influence of the lagged terms of one variable on the current values of another variable; however, Granger causality tests do not reveal the duration and sign of the causality (Liddle and Lung, 2013). Therefore, to account for the limitations of Granger causality tests, this study employs impulse response analysis.

4.3.1.3. Impulse Response Analysis

An important component of this analysis is investigating how one variable responds to an impulse in another variable (Rahman and Shahbaz, 2013). More specifically, the emphasis of this study is on how current market turnover responds to shocks in market returns. However, individual VAR coefficients do not fully capture the effect of other endogenous and exogenous variables in the system (Swanson and Granger, 1997). By contrast, impulse response functions illustrate the impact of a residual shock by using all the coefficient estimates in the VAR system, thereby, fully capturing the effect of other variables in the system (Swanson and Granger, 1997). As such, impulse response functions provide an indication of how a shock to one of the variables impacts the values of the endogenous variables (both current and future values). Thus, whilst the parameters in the VAR model are quite dense, the impulse response functions provide a more precise illustration of how endogenous variables ($mret$ and $mturn$) respond to each other. Notably, the overconfidence hypothesis is confirmed when $mturn$ exhibits significant positive responses to shocks in

mret. The impulse responses are presented within a given confidence interval in order to analyse the statistical significance of the estimated responses (Hatemi, 2014).

Overall, the VAR models, Granger causality tests and impulse response functions are employed to study the presence of market-wide investor overconfidence in an attempt to answer research question one. The next section outlines the empirical methodology employed to answer research question two which examines the presence of investor overconfidence at the individual security level.

4.4. Method Used to Determine Whether Investor Overconfidence Influences the Trading Activities of Individual ETFs

Panel VAR models are employed to investigate how the turnover of individual ETFs is related to the return of the market in order to determine whether the overconfidence bias influences the trading activities of individual ETFs, and thus, to answer research question two. As such, the ensuing discussion describes the method used to ensure that the market-wide overconfidence effect (tested using the methods discussed in Section 4.3.1) is not a summation of the disposition effect.

4.4.1. Individual Security VAR Model

A positive relationship between trading activity and historical returns is consistent with the overconfidence effect but also with the disposition effect (Shefrin and Statman, 1985). On the one hand, high market returns may induce increased trading because of an increase in investors' confidence about their security-picking skills or because investors enjoy realising paper gains on individual ETFs. On the other hand, negative market returns may reduce trading activities because of a decrease in investors' confidence about their security-picking skills or because investors want to hold on to individual ETFs and not realise the loss. A complete disentanglement of the overconfidence effect from the disposition effect is difficult because one could argue that when investors sell other securities for the disposition effect, they raise cash that can then be used to purchase the security in question, thus increasing its volume (Statman, *et al.*, 2006). Nevertheless, in spite of this drawback in interpretation, an assessment of the trading activity of individual ETFs will at least confirm whether or not the observed overconfidence effect is a sum of the disposition effect (Statman, *et al.*, 2006; Metwally and Darwish, 2015; Gupta. *et al.*, 2018).

To ensure that the patterns observed in the market-wide turnover are not a direct summation of the disposition effect, Statman, *et al.* (2006) propose an individual security VAR model which examines patterns in individual security turnover. Therefore, this study employs the individual security VAR model

proposed by Statman, *et al.* (2006) to ascertain the validity of the market-wide results. Shefrin and Statman (1985) argue that the disposition effect motivates the examination of the impact of individual security return on individual security turnover. Hence, individual security return is included as an endogenous variable in the individual security VAR model to control for the disposition effect. Additionally, market return is included as an endogenous variable in the VAR system to observe the response of individual security turnover to past market returns. Thus, the VAR model for individual securities employs individual security turnover ($turn$), individual security return (ret), and market return ($mret$) as endogenous variables. Regarding the control variables, individual security volatility (sig) is included as an exogenous variable to account for the volume-volatility relationship previously discussed under Section 4.3.1. However, the dispersion variable, which is used in the market VAR model to account for portfolio rebalancing activities, is dropped as a control variable because the individual security VAR model is applied to each ETF individually.

The individual security VAR model estimated for each security in the respective market follows the following specification¹⁹:

$$\begin{bmatrix} turn_t \\ ret_t \\ mret_t \end{bmatrix} = \begin{bmatrix} \alpha_{turn} \\ \alpha_{ret} \\ \alpha_{mret} \end{bmatrix} + \sum_{p=1}^P A_p \begin{bmatrix} turn_{t-p} \\ ret_{t-p} \\ mret_{t-p} \end{bmatrix} + \sum_{s=0}^S B_s sig_{t-s} + \begin{bmatrix} e_{turn,t} \\ e_{ret,t} \\ e_{mret,t} \end{bmatrix} \quad (4.31)$$

In equation 4.31, $turn_t$ represents the individual security (that is, ETF) turnover for month t , ret_t is the individual security (ETF) return for month t , $mret_t$ is the market return for month t , and sig_t represents the individual security (ETF) volatility for month t . A_p and B_s are regression coefficients while P and S denote the optimal lag lengths of the endogenous and exogenous variables, respectively.

Statman, *et al.* (2006) estimate 1878 individual security VAR models whereby each VAR model has a complete set of estimated coefficients with accompanying standard errors. As such, the output produced is extensive and difficult to report on an individual security basis. For brevity, Statman, *et al.* (2006) report the cross-sectional mean coefficient for all endogenous variables. However, using the cross-sectional standard deviation to compute the standard error of the mean coefficient underestimates the standard error, and thus, the statistical significance of each average coefficient is difficult to determine because the standard error is biased. To overcome this drawback, Statman, *et al.* (2006) use a bootstrap procedure to estimate the standard errors of the average coefficient. Given that VAR models have a time-series nature and the database comprises of a large number of securities, this bootstrap procedure is econometrically complicated and computationally intensive. In this study, there are 47 ETFs replicating local benchmarks as well as 8

¹⁹ Equation 4.31 is the same as equation 3.4 but has been repeated for ease of reference.

ETFs replicating international benchmarks, making it difficult to report the results of a total of 55 individual security VAR models. Whilst reporting the cross-sectional average coefficient is a solution to this problem, the standard errors of the mean coefficients are biased and introducing a bootstrap procedure would be complex. Therefore, to address this problem, this study estimates the individual security VAR models using a panel data approach, that is, panel VAR models. Further benefits of the panel data approach are discussed in the next section which outlines the panel VAR model used in this study.

4.4.2. Individual Security VAR Model with Panel Data Approach

Panel data refers to data containing elements of both time series and cross-sectional data (Chipunza and McCullough, 2018). This being so, panel data studies exhibit a greater capacity for modelling the complexity of human behaviour in comparison to a single time-series or cross-sectional analysis (Hsiao, 2005). Hsiao (2007) argues that the analysis of panel data leads to more accurate inferences of the model estimates because panel data contains more sample variability. Moreover, by combining the intra-individual dynamics and inter-individual differences, panel data approaches allow the researcher to control for the effects of omitted variables (Hsiao, 2007). Hence, a key advantage of panel data methodologies is that they are able to handle heterogeneity better than cross-sectional data methods (Baltagi and Pesaran, 2007).

The challenge for panel data techniques is controlling for the effects of the unobserved heterogeneity. Specifically, the unobserved heterogeneity effects can either be assumed as fixed parameters, random variables, or a combination of the two (Hsiao, 2007). Accordingly, Lin, Law, Ho and Sambasivan (2019) note that there are two classes of panel methodologies, namely; the fixed effects model or the random effects model. Fixed effects models assume that the error term has a non-stochastic variation across each entity and over time (Lin, *et al.*, 2019). In contrast, random effects models assume that the error term varies stochastically. However, the individual security-level panel VAR model estimated by Zia, *et al.* (2017) does not control for unobserved heterogeneity. Moreover, Canova and Ciccarelli (2013) note that panel VAR models used for microeconomic studies typically disregard cross-sectional interdependencies and assume cross-sectional homogeneity. Thus, following Zia, *et al.* (2017), the security-level panel VAR models estimated in this study disregards cross-sectional heterogeneity, and therefore, does not control for fixed or random heterogeneity effects.

Panel VAR models can be estimated using either a balanced panel or an unbalanced panel. A balanced panel contains an equal number of time series observations for each cross-sectional element whilst an unbalanced panel, as used in this study, consists of cross-sectional elements with an unequal number of time series observations (Wooldridge, 2019). Given that panel data approaches have a higher capacity for

capturing the complexity of human behaviour and result in more accurate inferences, equation 4.31 is adapted to accommodate a panel VAR model. Relative to the bootstrapping method employed by Statman, *et al.* (2006), the panel VAR methodology used in this study simplifies the model estimation process and improves the accuracy of the statistical inferencing. The basic form of the panel VAR estimated by Zia, *et al.* (2017) takes the following specification:

$$Y_t = \alpha + \sum_{k=1}^K A_k Y_{t-k} + \sum_{l=1}^L B_l X_{t-1} + e \quad (4.32)$$

In equation 4.32, Y_t represents a vector of endogenous variables and X_t represents a vector of the exogenous variable.

Rewriting equation 4.31 using the panel data approach of Zia, *et al.* (2017), the individual security panel VAR model estimated in this study is as follows:

$$\begin{bmatrix} turn_{i,t} \\ ret_{i,t} \\ mret_{i,t} \end{bmatrix} = \begin{bmatrix} \alpha_{turn} \\ \alpha_{ret} \\ \alpha_{mret} \end{bmatrix} + \sum_{p=1}^P A_p \begin{bmatrix} turn_{i,t-p} \\ ret_{i,t-p} \\ mret_{i,t-p} \end{bmatrix} + \sum_{s=0}^S B_s sig_{t-s} + \begin{bmatrix} e_{turn,it} \\ e_{ret,it} \\ e_{mret,it} \end{bmatrix} \quad (4.33)$$

In equation 4.33, there are three endogenous variables: $turn_{i,t}$ (the individual security turnover for the i th ETF for month t calculated using equation 4.3), ret_t (the individual security return for the i th ETF for month t computed by aggregating its daily returns calculated using equation 4.7), and $mret_t$ (ETF i 's respective market return for month t calculated by aggregating the daily returns computed using equation 4.6). The lagged return volatility of the i th ETF for month t (calculated using equation 4.9) is denoted by $sig_{i,t-s}$. A_p and B_s are regression coefficients while P and S indicate the optimal lag lengths of the endogenous and exogenous variables, respectively.

To determine whether the market-wide overconfidence effect is a simple aggregation of the disposition effect, it is crucial to examine how individual ETF turnover responds to individual ETF return and its respective market return. Since the disposition effect relates to investors' attitudes towards individual securities in their portfolios, a significant positive relationship between past individual ETF return and current individual ETF turnover is indicative of the presence of the disposition effect (Shefrin and Statman, 1985; Odean, 1998; Statman, *et al.*, 2006). On the contrary, investor overconfidence relates to investors' attitudes about the aggregate ETF market rather than an individual ETF, and therefore, a significant positive effect of historical market returns on current individual ETF turnover is indicative of the presence of the overconfidence effect at individual security level (Odean, 1998; Statman, *et al.*, 2006; Metwally and Darwish, 2015). Therefore, if investor overconfidence influences the trading activities of individual ETFs, lagged $mret_t$ coefficients in the $turn_{i,t}$ equation will be positive and statistically significant. The statistical significance of this relationship is analysed using the t-test. It is important to note that the disposition effect

and the overconfidence bias may exist concurrently in which case the overconfidence effect accounts for the trading activity that is not captured by the disposition effect. When the overconfidence bias induces trading activities in addition to the disposition bias, the regression with individual security turnover as the dependent variable exhibits significant and positive lagged market return coefficients, even after lagged individual ETF returns are included in the regression (Statman, *et al.*, 2006).

The individual security models are estimated for both the market of JSE-listed ETFs replicating local benchmarks as well as the market of JSE-listed ETFs replicating offshore benchmarks²⁰. Consistent with Statman, *et al.* (2006) and Gupta, *et al.* (2018), the optimal lag lengths of the endogenous and exogenous variables in the individual security VAR models follow the optimal lag lengths selected for the respective market VAR model. Statman, *et al.* (2006) argue that the uniformity in the lag length will ensure consistency in the comparison of individual ETFs since the optimal lag lengths may vary substantially across individual ETFs. In addition to the estimated individual security panel VAR models, Granger causality tests (discussed in Section 4.3.1.2)²¹ and impulse response functions (discussed in Section 4.3.1.3) are examined to trace how individual security turnover responds to shocks in market return over time, thus, providing evidence of the presence of investor overconfidence at individual security level. Overall, the panel VAR models, Granger causality tests and their associated impulse response functions are employed to provide further evidence of the prevalence of overconfident trading by investors in the respective ETF markets in South Africa. The next section outlines the empirical method used to investigate how overconfident trading effects the volatility of the respective market.

4.5. Method Used to Examine the Effect of Investor Overconfidence on Market Volatility

Numerous studies (Odean, 1998; Benos, 1998; Chuang and Lee, 2006; Abbes, 2013; Jlassi, *et al.*, 2014) document that the volatility of the returns from assets increases as the level of investor overconfidence increases. It is important to note that, whilst these studies argue that excessive trading by overconfident investors contributes to excess market volatility, overconfident trading is not the only source of excess market volatility. To answer research question three, the effect of investor overconfidence on the volatility of the returns from the respective South African ETF market is examined using the Chuang and Lee (2006) model. This model is also employed by Abbes (2013) and Jlassi, *et al.* (2014) in their attempts to investigate

²⁰ Prior to the estimation of the individual security models, the stationarity of each panel variable is confirmed using the Choi unit root test discussed under Section 4.2.7.1.

²¹ The Granger causality test estimated for the individual security model differs from equation 4.29 and 4.30. Specifically, $mturn_t$ is replaced with $turn_t$ and there are three explanatory variables, namely; $turn_{t-j}$, $mret_{t-j}$ and ret_{t-j} .

the influence of overconfident behaviour on market volatility. In this study, the analysis of the effect of investor overconfidence on the market return volatility is conducted for the market of JSE-listed ETFs replicating local benchmarks as well as the market of JSE-listed ETFs replicating international benchmarks.

The first step in identifying whether excess market volatility results from excessive trading caused by overconfident investors is to distinguish between trading volume that is related to investors' overconfidence levels due to historical market returns and trading volume that is not associated with investors' overconfidence. Following Chuang and Lee (2006), trading volume, V_τ , is decomposed into the two components using the following computation²²:

$$V_\tau = \alpha + \sum_{j=1}^J \beta_j R_{\tau-j} + \varepsilon_\tau = \left[\sum_{j=1}^J \beta_j R_{\tau-j} \right] + [\alpha + \varepsilon_\tau] = OVER_\tau + NONOVER_\tau \quad (4.34)$$

In equation 4.34 above, the trading volume, V_τ , is defined as the market turnover on day τ and is calculated using the computation discussed under Section 4.2.6.1. R_τ represents the return on the market for day τ and is computed using equation 4.6. $R_{\tau-j}$ is the return on the market on day $\tau - j$ where J is the optimal number of lags, β_j is the coefficient that captures the relationship between past market returns and current trading volume, α is the constant term, and ε_τ denotes the error term. Chuang and Lee (2006) define the component of trading volume that is not related to investors' overconfidence ($NONOVER_\tau$) as the sum of the constant and error terms. As such, the component of trading volume that is related to investors' overconfidence ($OVER_\tau$) is computed as the trading volume minus the sum of the constant and error terms, that is, the difference between V_τ and $NONOVER_\tau$. Since formal theories relating to investor overconfidence do not stipulate a time frame for the relation between trading volume and returns, the optimal lag length, J , is selected based on the AIC, SC, and HQ information criteria which was discussed under Section 4.3.1.1. Specifically, this study first estimates equation 4.34 with $j = 1$, and for each subsequent model, j increases by 1 up to a maximum of $j = 8$. The optimal lag length is then selected based on the model that minimizes the AIC, SC, and HQ criteria.

Following Chuang and Lee (2006), Abbes (2013), and Jlassi, *et al.* (2014), the two trading volume components are then merged into the conditional variance equation of an EGARCH (1,1) model. The primary advantage of the EGARCH model, which was introduced by Nelson (1991), is that the model captures asymmetric effects in which negative return shocks have a greater impact on volatility in comparison to positive return shocks of the same magnitude (Corbet and Gurdgiev, 2019). Another advantage of the EGARCH model is that since the conditional variance equation is in log-linear form, the model allows coefficients to be negative without the conditional variance becoming negative (Halkos and

²² Equation 4.34 is the same as equation 3.13 but has been repeated for ease of reference.

Tsirivis, 2019). The use of the EGARCH model is further supported by Labuschagne, Oberholzer and Venter (2017a) who find that the EGARCH (1,1) model is the best fitting model for modelling the volatility of the returns from CARBS (Canada, Australia, Russia, Brazil, and South Africa) indices. Similarly, Labuschagne, Oberholzer and Venter (2017b) report that the univariate EGARCH model is the best fitting model for equity indices in Indonesia, Turkey and South Africa. As such, the EGARCH (1,1) model is estimated to determine how overconfident trading effects the volatility of the returns from the South African ETF market. The EGARCH (1,1) model requires the estimation of a mean equation and a conditional variance equation, both of which are discussed in the ensuing sections.

Consistent with Chuang and Lee (2006), Abbes (2013), and Jlassi, *et al.* (2014), the mean equation of the EGARCH (1,1) model estimated to examine the effect of overconfident trading on market volatility is as follows:

$$R_{\tau} = \mu_{\tau} + \eta_{\tau} \quad (4.35)$$

where R_{τ} in equation 4.35 represents day τ 's return on the market (calculated using equation 4.6) and μ_{τ} is the average daily return on the market. η_{τ} denotes day τ 's error generated from the mean equation and follows a Generalised Error Distribution (GED) in order to account for the non-normality in the distribution of the returns.

A key advantage of GARCH-type models is their ability to capture serial dependence by allowing the conditional variance to depend on past residuals and historical conditional variance (Kroner and Lastrapes, 1993). Given this, the general form of the conditional variance equation for the EGARCH (1,1) model is as follows (Samouilhan, 2006):

$$\ln(\sigma^2) = \omega + f_1 \frac{|\eta_{\tau-1}|}{\sqrt{\sigma_{\tau-1}^2}} + f_2 \frac{\eta_{\tau-1}}{\sqrt{\sigma_{\tau-1}^2}} + f_3 \ln(\sigma_{\tau-1}^2) \quad (4.36)$$

In equation 4.36, the natural logarithm of the conditional variance (σ^2) is a function of the constant term (ω), the natural logarithm of past conditional variance ($\sigma_{\tau-1}^2$) and both the absolute and level values of the standardised residuals, that is, $\frac{|\eta_{\tau-1}|}{\sqrt{\sigma_{\tau-1}^2}}$ and $\frac{\eta_{\tau-1}}{\sqrt{\sigma_{\tau-1}^2}}$, respectively. η_{τ} in equation 4.36 represents the error term of the mean equation on day τ as defined in equation 4.35. f_1 , f_2 , and f_3 are parameters representing the Autoregressive Conditional Heteroskedasticity (ARCH) effect, the asymmetry effect, and the Generalised Autoregressive Conditional Heteroskedasticity (GARCH) effect, respectively.

The EGARCH (1,1) model achieves covariance stationarity when the sum of f_1 and f_3 is less than one (Atoi, 2014). However, the higher the sum of f_1 and f_3 , the higher the degree of persistence in the

conditional volatility because significant f_1 and f_3 coefficients indicate that current conditional volatility is influenced by lagged residuals and historical conditional volatility, respectively. Moreover, asymmetric volatility effects exist when $f_2 \neq 0$. Particularly, when $f_2 < 0$, leverage effects are present and negative return shocks have a greater impact on volatility in comparison to positive return shocks of the same magnitude (Samouilhan, 2006). Contrarily, when $f_2 > 0$, positive return shocks have a greater impact on volatility relative to negative return shocks of the same magnitude. If $f_2 = 0$, asymmetric volatility effects do not exist (Samouilhan, 2006).

Following Chuang and Lee (2006), Abbes (2013), and Jlassi, *et al.* (2014), the effect of trading volume induced by investor overconfidence on the volatility of the market is examined by incorporating the two trading volume components into the conditional variance equation of the EGARCH (1,1) model as defined below:

$$\ln(\sigma^2) = \omega + f_1 \frac{|\eta_{\tau-1}|}{\sqrt{\sigma_{\tau-1}^2}} + f_2 \frac{\eta_{\tau-1}}{\sqrt{\sigma_{\tau-1}^2}} + f_3 \ln(\sigma_{\tau-1}^2) + f_4 OVER_{\tau} + f_5 NONOVER_{\tau} \quad (4.37)$$

In equation 4.37, ω , f_1 , f_2 , and f_3 is as defined in equation 4.36 and $OVER_{\tau}$ and $NONOVER_{\tau}$ is as defined in equation 4.34. In equation 4.37, f_4 captures the effect of investor overconfidence on the conditional market volatility whilst f_5 reflects the impact of other possible explanations (except investor overconfidence) on the conditional market volatility. According to Chuang and Lee (2006), trading volume induced by investor overconfidence exhibits a positive impact on market volatility when f_4 is positive and statistically significant. In other words, if f_4 is positive and significant, an increase in investor overconfidence (and its subsequent trading volume) leads to a rise in market volatility. Moreover, if f_4 is significantly positive and greater than f_5 then trading volume caused by overconfidence contributes to excess market volatility. The statistical significance of these relationships are examined using the z-test, specifically, the z-statistics and their associated probability values.

Overall, the impact of investor overconfidence on the volatility of the returns from the respective South African ETF market is examined by estimating an EGARCH (1,1) model which follows the mean equation defined in equation 4.35 and the conditional variance equation defined in equation 4.37. The reliability of the estimated model is assessed by running diagnostic tests for ARCH effects as discussed in Section 4.2.7.4. The next section discusses the method employed in the sub-period analysis of the effect of investor overconfidence on market volatility.

4.6. Sub-Period Analysis of the Effect of Investor Overconfidence on the Volatility of the South African ETF Market

Overconfidence theories maintain that investors who are overconfident about their price trend predictions overestimate the precision of their information and underestimate the likelihood of extreme events. Accordingly, overconfident investors trade more aggressively on winning stocks during market uptrends, causing asset prices to increase above their intrinsic values, subsequently, leading to a persistent shift in their fundamental values (Gervais and Odean, 2001). Consequently, there is an abnormal increase in market volatility. Jlassi, *et al.* (2014) argue that the 2007 housing bubble was fuelled by overconfident trading because overestimated asset prices increased market volatility asymmetrically. Abbes (2013) proposes that behavioural finance is crucial to understanding the sources of increased volatility experienced during the 2008 global financial crisis as well as to identify potential policy responses that could be used to mitigate the behaviours that lead to financial market instability.

Given the discussion above, research question four attempts to examine the effect of trading volume induced by overconfidence on ETF market volatility before, during, and after the 2008 global financial crisis. As mentioned in Chapter 1, several market experts have predicted a passive investment bubble which could subsequently lead to an ETF market crash, thus, the objective of the sub-period analysis is to determine whether trading induced by the overconfidence bias in the market of JSE-listed ETFs could increase the volatility experienced during an ETF market crash. To achieve this objective, a sub-period analysis of the effect of investor overconfidence on market volatility is conducted by estimating the EGARCH (1,1) model outlined in Section 4.5 using the during-crisis data samples. Additionally, the effect of overconfidence on market volatility is examined before and after the financial crisis by estimating the EGARCH (1,1) model using the pre-crisis and post-crisis data samples, respectively. The sub-period analysis is performed for both the market of JSE-listed ETFs replicating South African benchmarks and the market of JSE-listed ETFs replicating non-South African benchmarks, and the pre-, during-, and post-crisis sample periods of each market are tabulated in Table 4.1 (Section 4.2.1). It is important to note that, to ensure consistency in each day's overconfident trading value ($OVER_{\tau}$) and non-overconfident trading value ($NONOVER_{\tau}$), the respective values computed for the full sample (using equation 4.34) is also used in the sub-period analysis, although, the series is segregated into subsamples.

4.7. Validity and Reliability

Validity refers to the extent to which the results reflect how things are out in the real world (Sale, Lohfeld and Brazil, 2002; Thanasegaran, 2009). In this study, the prevalence of overconfident trading by investors in South Africa's ETF market is tested by examining the association between the current turnover of the

market and past market return. The level of overconfidence is linked to previous market returns since overconfidence levels fluctuate based on investors' historical performances (Gervais and Odean, 2001). The level of overconfidence is also linked to market turnover because higher overconfidence levels cause investors to trade more aggressively (Odean, 1998). Therefore, investors trade more frequently when historical market returns are favourable because of increased confidence levels since these investors attribute market gains to their personal ability to pick securities.

Moreover, to ensure that the overconfidence effect is not an aggregation of the disposition effect, an individual security model is employed. The overconfidence effect present at the market level is verified when individual security turnover responds positively to historical market returns, that is, when an increase in historical market returns increases the turnover of individual ETFs. Additionally, the trading volume is decomposed into a component relating to investors' overconfidence and a component unrelated to investors' overconfidence to examine the effect of these two components on the conditional volatility of market returns -using the EGARCH model. EGARCH models account for heteroskedasticity in the error term as well as asymmetric volatility effects, thus, providing a more real-world representation of the patterns of financial market returns and volatility. Overall, the variables and methodologies employed in this study ensure that the results obtained reflect the reality of the South African ETF market.

Reliability refers to the degree of consistency in measuring a variable (Thanasegaran, 2009; Ihantola and Kihn, 2011). To confirm the reliability of the results, the data and methods employed need to be examined. With regards to the data collection process, data is only collected for the 55 South African ETFs that meet the minimum data requirements. As at the 30th of August 2020, the 55 ETFs included in the survivorship bias-free sample account for 63% of the presently listed ETFs and 67% of the currently delisted ETFs, and thus, is considered a reasonable proxy of the South African ETF market. Moreover, the data is obtained from the IRESS and Infront Analytics databases, and the models are estimated using the Eviews statistical software package, all of which are widely used in academic research. Regarding the methods employed, VAR models are analysed in order to examine the response of market turnover to past market returns. To make the model more reliable, the market-wide VAR models control for alternative explanations of trading activity, such as market volatility and dispersion. Furthermore, to ensure that the overconfidence effect is not a direct aggregation of the disposition effect, the relationship between individual security turnover and past market returns is examined using panel VAR models. In addition, popular techniques used to model volatility, specifically, the EGARCH model, is used to detect the effect of overconfidence on the volatility of the market returns. During the model estimations process, diagnostics tests are carefully performed to

further assure the reliability of the results obtained. Therefore, the data and methods employed in this study support the notion of reliable results.

4.8. Summary

The current chapter provided a detailed discussion of the empirical methodologies used to study the presence and effect of overconfident trading by investors in South Africa's ETF market. The analysis of overconfidence is conducted on the market of JSE-listed ETFs replicating domestic benchmarks as well as the market of JSE-listed ETFs replicating international benchmarks. Since the first South African ETF tracking a local benchmark was introduced in November 2000, the analysis of ETFs tracking domestic benchmarks begins November 2000, whereas the analysis of ETFs tracking international benchmarks begins October 2005 because the first South African ETF tracking an offshore benchmark was launched in October 2005. Notably, the sample periods end on the 30th of August 2019 for both markets.

Regarding the methodologies employed, market VAR models are estimated for both markets to investigate the lead-lag relationship between current market turnover and lagged market returns. The market VAR models estimated account for other possible explanations of market turnover, viz. market volatility and dispersion. To increase the reliability of the market-wide results, individual security-level VAR models are estimated with a panel approach in order to determine whether overconfidence is also present at the individual security-level. Additionally, EGARCH (1,1) models are employed to explore the effect of trading volume induced investor overconfidence on the volatility of the returns from the respective ETF market. The EGARCH models are estimated for both markets over the full sample periods and for all sub-periods to examine the overconfidence effect before, during and after the 2008 global financial crisis. Table 4.3 (page 82) provides a summary of the empirical methods employed to answer the research questions of this study.

Prior to the estimation of the empirical regressions, preliminary data analysis (that is, analysis of unit root tests, descriptive statistics, and correlations) is conducted to ensure that the time series is appropriate for use in the estimations. Moreover, after each model is estimated, diagnostics tests (for serial correlation and heteroskedasticity) are performed to confirm the reliability of the results obtained. The next chapter provides a discussion of the results obtained.

Table 4.3: Summary of Empirical Methods

Research Question	Method Employed	Regression
Are investors overconfident when trading in the South African ETF market?	VAR model	$\begin{bmatrix} mturn_t \\ mret_t \end{bmatrix} = \begin{bmatrix} \alpha_{mturn} \\ \alpha_{mret} \end{bmatrix} + \sum_{p=1}^P A_p \begin{bmatrix} mturn_{t-p} \\ mret_{t-p} \end{bmatrix} + \sum_{s=0}^S B_s \begin{bmatrix} msig_{t-s} \\ disp_{t-s} \end{bmatrix} + \begin{bmatrix} \varepsilon_{mturn,t} \\ \varepsilon_{mret,t} \end{bmatrix}$
Does investor overconfidence influence the trading activity of individual ETFs?	Panel VAR model	$\begin{bmatrix} turn_{i,t} \\ ret_{i,t} \\ mret_{i,t} \end{bmatrix} = \begin{bmatrix} \alpha_{turn} \\ \alpha_{ret} \\ \alpha_{mret} \end{bmatrix} + \sum_{p=1}^P A_p \begin{bmatrix} turn_{i,t-p} \\ ret_{i,t-p} \\ mret_{i,t-p} \end{bmatrix} + \sum_{s=0}^S B_s sig_{t-s} + \begin{bmatrix} e_{turn,it} \\ e_{ret,it} \\ e_{mret,it} \end{bmatrix}$
Does investor overconfidence exhibit a positive or negative effect on the volatility of the returns from the South African ETF market?	EGARCH (1,1) model	<p>Mean equation: $R_\tau = \mu_\tau + \eta_\tau$ Conditional variance equation:</p> $\ln(\sigma^2) = \omega + f_1 \frac{ \eta_{\tau-1} }{\sqrt{\sigma_{\tau-1}^2}} + f_2 \frac{\eta_{\tau-1}}{\sqrt{\sigma_{\tau-1}^2}} + f_3 \ln(\sigma_{\tau-1}^2) + f_4 OVER_\tau + f_5 NONOVER_\tau$
Does the volatility of the South African ETF market respond positively or negatively to trading volume induced by investor overconfidence before, during, and after the 2008 global financial crisis?	EGARCH (1,1) model	<p>Mean equation: $R_\tau = \mu_\tau + \eta_\tau$ Conditional variance equation:</p> $\ln(\sigma^2) = \omega + f_1 \frac{ \eta_{\tau-1} }{\sqrt{\sigma_{\tau-1}^2}} + f_2 \frac{\eta_{\tau-1}}{\sqrt{\sigma_{\tau-1}^2}} + f_3 \ln(\sigma_{\tau-1}^2) + f_4 OVER_\tau + f_5 NONOVER_\tau$

CHAPTER 5: DATA ANALYSIS AND RESULTS

5.1. Overview

The current chapter attempts to answer the research questions of this study by analyzing the results obtained from the estimations of the empirical methodologies which were outlined in Chapter 4. Prior to the analysis of the results obtained from the estimated models, a preliminary data analysis is reported to ensure that the data employed is suitable for this study. This is followed by a discussion of the empirical results. Specifically, the results from the market VAR models are discussed in order to determine whether investor overconfidence is present in the South African ETF market (research question one). Thereafter, the results of the individual security panel VAR models are provided to examine whether the overconfidence bias influences the trading activities of individual ETFs (research question two). The latter part of Chapter 5 outlines the results obtained from the estimated EGARCH models in an attempt to investigate the effect of overconfidence on volatility (research question three) as well as a sub-period analysis of this effect (research question four).

5.2. Preliminary Data Analysis

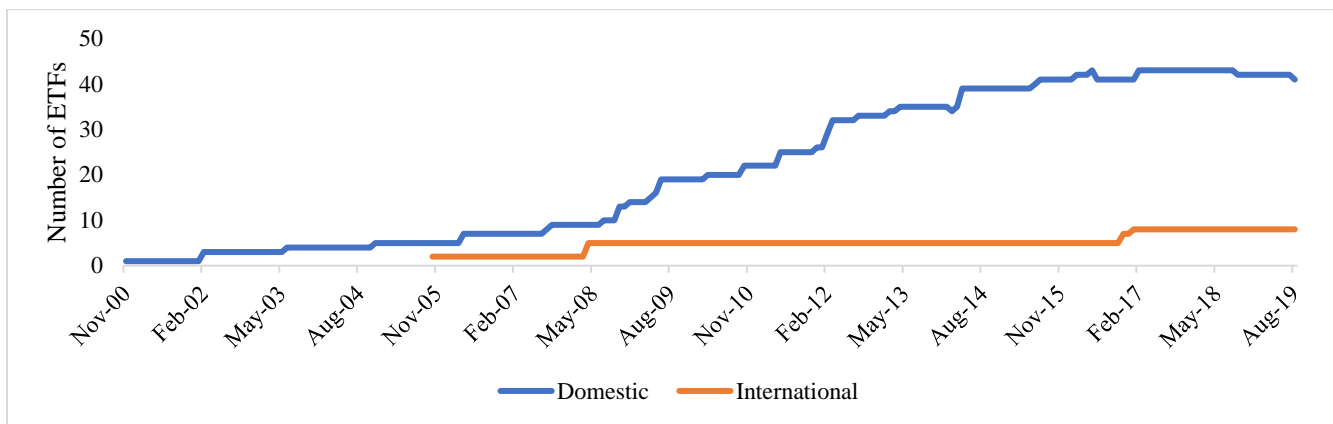
5.2.1. Number of ETFs Included in This Study

In order to provide empirical results that represent the entire South African ETF market, the initial starting point for the data analysis is to include all ETFs that traded on the JSE, including ETFs that were delisted. However, because of this study's minimum data requirements, ETFs with less than 30 months of data were excluded from the sample. To account for the listing or delisting of a new ETF, the respective market portfolio is reconstituted each time a new ETF is listed or delisted on a market-value month-by-month basis where monthly observations are used and a day-by-day basis where daily observations are used. For the market-value month-by-month reconstitutions, the resulting number of ETFs included in the market portfolio of JSE-listed ETFs replicating South African benchmarks and the market of JSE-listed ETFs replicating non-South African benchmarks are illustrated in Figure 5.1 (page 84).

Figure 5.1 shows that the number of ETFs which adhered to the minimum data requirements, and thus, is included in the respective market portfolio has increased over the sample period. This increase in the number of portfolio constituents is a result of the increase in the number of ETFs trading on the JSE. Initially, in November 2000, the market of JSE-listed ETFs replicating South African benchmarks comprised of 1 ETF and, by the end of August 2019, the respective market constituted of 41 listed ETFs. However, over the 226 month sample period, the market of JSE-listed ETFs replicating South African benchmarks comprised of a total sample of 47 ETFs, including the 41 ETFs listed at the end of August

2019 and 6 ETFs delisted during the sample period. The market portfolio representing South African ETFs tracking international benchmarks initially comprised of 2 ETFs in October 2005 and, by the end of August 2019, the respective market constituted of 8 ETFs. No ETFs in the market of JSE-listed ETFs replicating international benchmarks were delisted during the sample period.

Figure 5.1: Number of ETFs Included in the Market-Value Weighted Portfolios



5.2.2. Stationarity of Variables

Secondary data described in Section 4.2.5 are used to compute the monthly market turnover (*mturn*), market return (*mret*), market volatility (*msig*), and market dispersion (*disp*) for the market portfolio²³ of each ETF category. Figure 5.2 (page 85) provides a graphical representation of the endogenous variables (that is, monthly market turnover and market return) for the market of ETFs with local benchmarks and Figure 5.3 (page 85) illustrates these monthly market variables for the market of ETFs with offshore benchmarks.

Figures 5.2 and 5.3 illustrate each market portfolio’s monthly market turnover and return series that is used as endogenous variables in their respective market VAR models estimated to answer research question one. However, to avoid biased or spurious regressions, each series used in the market VAR model (described in Section 4.3.1) has to be stationary. From a visual inspection of Figure 5.2 and Figure 5.3, it is evident that, although the market return series exhibit periods of volatility clustering, the series frequently cross their mean values. Volatility clustering is commonly observed in financial asset returns and refers to the tendency of large fluctuations in returns to be followed by large fluctuations in returns and small fluctuations in returns to be followed by small fluctuations in returns (Ning, Xu and Wirjanto, 2015). Notably, the low market returns of the market of ETFs with domestic benchmarks after 2010 is attributed to the thin trading

²³In this study, the abbreviation ‘dom’ relates to the market of ETFs with local benchmarks and ‘int’ relates to the market of ETFs with international benchmarks.

bias because there exist several days in which ETFs in the respective market portfolio were not traded. As a result, the price of these ETFs did not change. Subsequently, on days when ETFs were not traded, the daily returns of these ETFs equalled zero. Thus, the return of the market-value weighted portfolio of ETFs tracking domestic benchmarks is relatively low after 2010. Alternatively, these low returns may be because the returns on ETFs with larger market capitalizations are offsetting the returns of ETFs with smaller market capitalizations. According to Harvey (1995), the influence of infrequent trading can be mitigated by making use of monthly data frequencies, thus, further supporting the use of monthly data for the VAR models estimated in this study.

Figure 5.2: Graphical Representation of the Monthly Market Variables for the Market of ETFs With Domestic Benchmarks

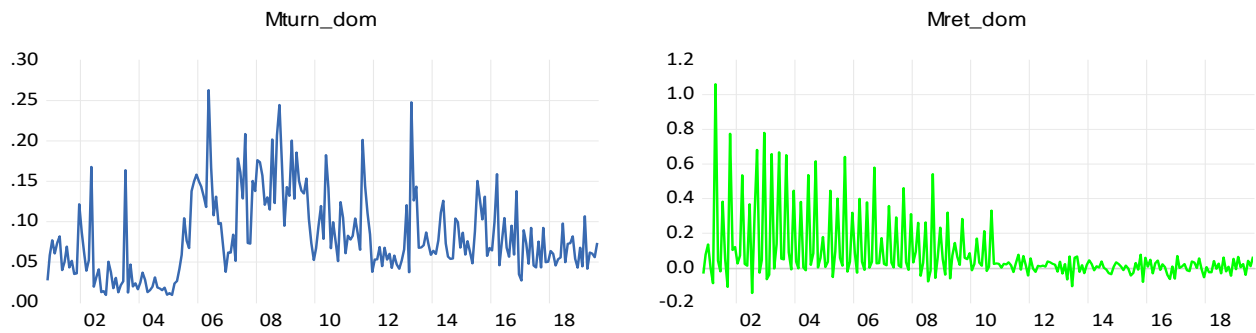
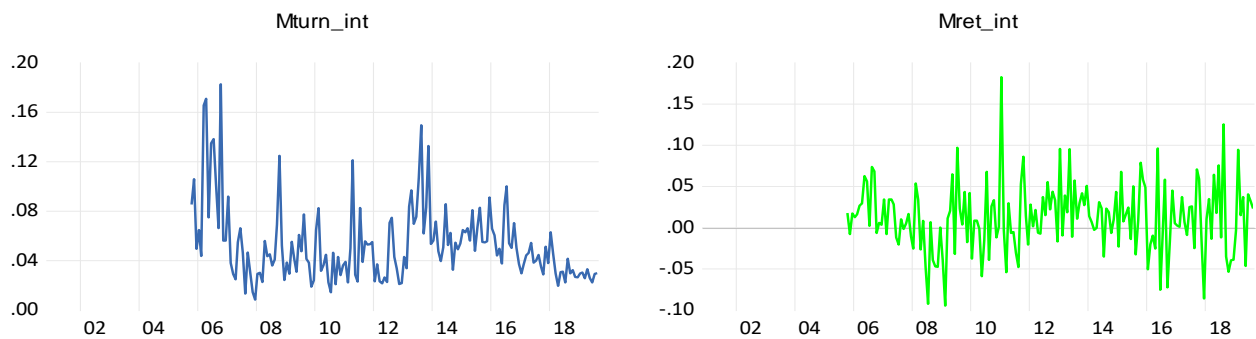


Figure 5.3: Graphical Representation of the Monthly Market Variables for the Market of ETFs With International Benchmarks.



In Figure 5.2 and Figure 5.3, there are no visible trends in the monthly market turnover and market return series of both market portfolios. Therefore, for both markets, the market turnover and market return series may be stationary. However, to ascertain the stationarity of each series, the PP unit root test is conducted on each series. The PP unit root test is used to confirm the stationarity of each series by testing the null

hypothesis that the respective series contains a unit root, and thus, is non-stationary. This being so, a series that is stationary rejects the null hypothesis of non-stationarity and, instead, accepts the alternative hypothesis of stationarity in the respective series. The PP unit root test is conducted at levels for each monthly series, and the results are presented in Table 5.1.

Table 5.1: Results from the PP Unit Root Test

Variable	PP test stat.	Prob.		Variable	PP test stat.	Prob.
Mturn_dom	-8.5714	0.0000		Mturn_int	-8.5664	0.0000
Mret_dom	-18.2524	0.0000		Mret_int	-13.1548	0.0000
Msig_dom	-13.0114	0.0000		Msig_int	-9.0379	0.0000
Disp_dom	-17.8079	0.0000		Disp_int	-9.3947	0.0000

Notes:

1. PP test stat. and *prob.* refers to the PP unit root test statistic and probability value, respectively.

Table 5.1 presents the results from the PP unit root test conducted on each series at their levels. In Table 5.1, the probability value associated with each PP test statistic is less than 0.01, and therefore, the null hypothesis of a unit root in the series is rejected at a 1% level of significance. This being so, the alternative hypothesis which hypothesizes the presence of stationarity in the time series is accepted for all series. The significant PP test statistics for all series indicate that the variables are stationary at levels. Thus, the stationary time series can be used to estimate the VAR models since they will not lead to biased or spurious regressions. Moreover, since all the series are stationary at levels, cointegration between market turnover and market return is not a problem, thus, VAR models can be used to appropriately estimate the long-run relationship between the variables and, therefore, the use of VECM is not required.

It is also important to ensure that the individual security variables are stationary in order to avoid biased or spurious regressions in the panel VAR models (described in Section 4.4.2). However, given the panel nature of the individual security variables, the Choi panel unit root test is employed to assess the stationarity of monthly security turnover, security return, market return and security volatility. Table 5.2 (page 87) presents the results of the Choi panel unit root test conducted at levels for each panel variable.

Table 5.2 contains the results from the Choi panel unit root test conducted on each panel variable at their levels. It is evident from Table 5.2 above that all the chi-square statistics and the Choi z-statistics are significant since their associated probability values are less than 0.01. This being so, for all panel variables, the null hypothesis that all series in the panel are non-stationary is rejected at a 1% level of significance,

suggesting that at least one series in the panel is non-stationary while other series are stationary. Overall, for both market portfolios, the panel monthly security turnover, security return, market return, and security volatility variables are stationary at a 1% level of significance, and thus, can be used in the estimation of the panel VAR models.

Table 5.2: Results from the Choi Panel Unit Root Test

Variable	PP-Fisher Chi-square	Prob.	PP-Choi Z-stat	Prob.
Turn_dom	1950.06	0.0000	-39.5476	0.0000
Ret_dom	2662.26	0.0000	-46.6325	0.0000
Mret_dom	2708.70	0.0000	-46.7140	0.0000
Sig_dom	2155.18	0.0000	-41.3632	0.0000
Turn_int	376.207	0.0000	-17.5100	0.0000
Ret_int	519.504	0.0000	-20.7070	0.0000
Mret_int	526.123	0.0000	-21.2667	0.0000
Sig_int	350.850	0.0000	-16.8752	0.0000

Notes:

1. PP-Fisher chi-square, PP-Choi z-stat. and *prob.* refers to the inverse chi-square test statistic, inverse normality test statistic, and probability value, respectively.

Additionally, the stationarity of the daily market return and market turnover series used in the EGARCH model (described in Section 4.5) is assessed using the PP unit root test, and the results are provided in Table 5.3. Table 5.3 indicates that all the PP test statistics are significant at a 1% level of significance since their associated p-values are less than 0.01. The significant PP test statistics for all series reject the null hypothesis of a unit root in the respective series, and therefore, indicates that the series are stationary at levels, and thus, can be used in the estimation of the EGARCH (1,1) models.

Table 5.3: Results from the PP Unit Root Test

Variable	PP test stat.	Prob.		Variable	PP test stat.	Prob.
V_dom	-84.0084	0.0000		V_int	-64.5823	0.0001
R_dom	-70.5471	0.0001		R_int	-63.9839	0.0001

Notes:

1. PP test stat. and *prob.* refers to the PP unit root test statistic and probability value, respectively.

5.2.3. Descriptive Statistics

Table 5.4 provides the descriptive statistics for each market portfolio's monthly market variables while Table 5.5 contains the descriptive statistics for each portfolio's monthly individual security variables which are of a panel nature. Additionally, Table 5.6 (page 89) presents the descriptive statistics of the daily market turnover and daily market return for both market portfolios.

Table 5.4: Descriptive Statistics of Each Monthly Market Variable

Variable	Mean	Standard Deviation	Skewness	Kurtosis	Jarque-Bera Statistic	Prob. (Jarque-Bera)
Mturn_dom	0.0819	0.0502	0.9890	3.7739	42.4802	0.0000
Mret_dom	0.0835	0.1834	2.4586	9.2793	598.9822	0.0000
Msig_dom	0.1952	0.1907	2.8035	13.174	1270.787	0.0000
Disp_dom	0.0224	0.0353	2.7865	10.5019	822.4091	0.0000
<hr/>						
Mturn_int	0.0528	0.0307	1.7620	6.7896	186.3416	0.0000
Mret_int	0.0121	0.0415	0.3601	4.3333	15.9788	0.0003
Msig_int	0.0452	0.0230	2.6901	18.2019	1809.473	0.0000
Disp_int	0.0104	0.0049	2.8658	15.1577	1257.095	0.0000

Notes:

1. *Prob.* refers to the probability value.

Table 5.5: Descriptive Statistics of Each Monthly Security Variable

Variable	Mean	Standard Deviation	Skewness	Kurtosis	Jarque-Bera Statistic	Prob. (Jarque-Bera)
Turn_dom	0.0887	0.6361	61.6560	4088.597	3.36e+09	0.0000
Ret_dom	0.0275	0.4230	0.1636	768.2180	1.18e+08	0.0000
Sig_dom	0.0880	0.5606	22.0619	524.8923	55183727.00	0.0000
<hr/>						
Turn_int	0.0574	0.0568	3.7758	28.2209	24403.71	0.0000
Ret_int	0.0082	0.0486	-0.0844	3.4308	7.5370	0.0231
Sig_int	0.0512	0.0278	3.3897	24.8222	18384.63	0.0000

Notes:

1. *Prob.* refers to the probability value.

Table 5.4 provides the descriptive statistics of each variable that is used in the respective market VAR models. On average, the monthly market turnover ratio for the market of JSE-listed ETFs replicating local benchmarks is 0.0819, whilst its average monthly market return is 0.0835%. With respect to the market of

ETFs tracking international benchmarks, the average monthly market turnover ratio is 0.0528 and its average monthly market return is 0.0121%. Interestingly, Table 5.5 shows that, in both markets, the average monthly return of individual ETFs is less than its respective average monthly market return. Specifically, the average individual security return is 0.0275% and 0.0082% for the market of ETFs with South African and non-South African benchmarks, respectively. Moreover, in Table 5.4 and Table 5.5, it is evident that variables with high mean values tend to exhibit high standard deviations while variables with low mean values tend to exhibit low standard deviations. Overall, for both the market and individual security variables, the market of JSE-listed ETFs with South African benchmarks exhibits the highest average monthly turnover, return, volatility, and dispersion. Similarly, as can be seen in Table 5.6, the average daily turnover and return of the market of ETFs with local benchmarks are higher than the market of ETFs with international benchmarks.

Table 5.6: Descriptive Statistics of Each Daily Market Variable

Variable	Mean	Standard Deviation	Skewness	Kurtosis	Jarque-Bera Statistic	Prob. (Jarque-Bera)
V_dom	0.0038	0.0054	5.2214	49.4398	462015.70	0.0000
R_dom	0.0039	0.0417	13.4311	214.3192	9253207.00	0.0000
V_int	0.0026	0.0044	7.300030	81.1549	913697.30	0.0000
R_int	0.0006	0.0113	0.371022	18.8377	36335.56	0.0000

Notes:

1. *Prob.* refers to the probability value.

As noted in Section 4.2.7.2, a normally distributed series contains a skewness value of 0 and a kurtosis value of 3. With the exception of the *ret_int* series in Table 5.5, the remaining series in Table 5.4, 5.5, and 5.6 exhibit skewness values that are positive, thus, indicating that the series are rightly skewed and positively distributed. On the contrary, the *ret_int* series displays negative skewness suggesting that the series is skewed to the left and negatively distributed. Regarding the kurtosis values, all the series contained in Tables 5.4, 5.5, and 5.6 exhibit kurtosis values that are greater than 3, and therefore, the series are leptokurtic and peaked. The final value that is important when determining whether a series is normally distributed is the JB test statistic. For all the series in Tables 5.4, 5.5 and 5.6, the JB test statistics are significant at a 5% level of significance since the p-values are less than 0.05. The significant JB statistics reject the null hypothesis of a normally distributed series, therefore, suggesting that the series are non-normally distributed. Overall, the skewness values, kurtosis values, and JB test statistics indicate that all the series are not normally distributed. However, according to Adu, Alagidede, and Karimu (2015), the

assumption of normality in asset returns is highly unrealistic, particularly for BRICS countries since these countries exhibit large growth swings and vulnerability to political and regulatory changes which may lead to asset returns that significantly deviate from the normality assumption. As a result, this financial time series analysis will proceed using the aforementioned non-normally distributed variables.

5.2.4. Correlation Analysis

The correlation coefficients between different variables are computed to provide an indication of the strength of the linear relationship between the variables used in each regression. For both markets, Table 5.7 displays the correlation coefficients computed between the market variables used in the market VAR model. Additionally, Table 5.8 presents the correlation coefficients calculated for the variables used in the panel VAR models and Table 5.9 (page 91) shows the correlation coefficients calculated between the variables employed in the estimation of the EGARCH models.

Table 5.7: Correlation Coefficients Between Each Portfolio's Market Variables

	Mturn_dom	Mret_dom	Msig_dom	Disp_dom
Mturn_dom	1.0000			
Mret_dom	-0.0785	1.0000		
Msig_dom	-0.0482	0.7315	1.0000	
Disp_dom	-0.0681	0.9623	0.7829	1.0000
	Mturn_int	Mret_int	Msig_int	Disp_int
Mturn_int	1.0000			
Mret_int	0.0362	1.0000		
Msig_int	0.0861	-0.0801	1.0000	
Disp_int	0.0789	-0.0122	0.7366	1.0000

Table 5.8: Correlation Coefficients Between Variables in Panel VAR Model

	Turn_dom	Ret_dom	Mret_dom	Sig_dom
Turn_dom	1.0000			
Ret_dom	0.0014	1.0000		
Mret_dom	-0.0099	0.2036	1.0000	
Sig_dom	-0.0039	-0.0146	0.1071	1.0000
	Turn_int	Ret_int	Mret_int	Sig_int
Turn_int	1.0000			
Ret_int	-0.0271	1.0000		
Mret_int	-0.0273	0.8070	1.0000	
Sig_int	0.1902	-0.1135	-0.1302	1.0000

Table 5.9: Correlation Coefficients Between Variables in the EGARCH Model

	V_dom	R_dom			V_int	R_int
V_dom	1.0000			V_int	1.0000	
R_dom	-0.0325	1.0000		R_int	0.0206	1.0000

Table 5.7 shows that contemporaneous market turnover and market returns exhibit a low, negative correlation for the market of ETFs with local benchmarks. This relation holds for both monthly and daily observations as can be seen in Table 5.7 and 5.9, respectively. This negative correlation suggests that, for the market of JSE-listed ETFs with local benchmarks, as market turnover increases (decreases), its market return decreases (increases), and vice versa. Additionally, the market return for the market of JSE-listed ETFs with local benchmarks also exhibits a low, negative correlation with individual security turnover since the correlation coefficient is -0.0099 in Table 5.8. On the contrary, there is a positive correlation between current market turnover and market returns for the market of ETFs with international benchmarks, however, this correlation is still low. Table 5.7 and 5.9 shows that this low, positive correlation between market turnover and market returns holds for both monthly and daily observations, respectively. Notably, Table 5.8 shows that the individual security turnover for the market of JSE-listed ETFs with offshore benchmarks is negatively correlated (-0.0273) with its market returns. The correlation statistics in Table 5.8 also suggests that individual security return is positively correlated (0.0014) with individual security turnover for the market of JSE-listed ETFs with local benchmarks, although, for the market of JSE-listed ETFs with international benchmarks, these variables are negatively correlated (-0.0271).

The positive correlation between the turnover and returns of the market of JSE-listed ETFs replicating non-South African benchmarks suggests that, as the turnover of the respective market portfolio increases (decreases), its market returns increases (decreases), and vice versa. However, this positive correlation between market turnover and market returns cannot be interpreted as the presence of investor overconfidence for two main reasons. Firstly, the correlation coefficient is computed using current market turnover and current market returns whereas the overconfidence hypothesis postulates that, in the presence of investor overconfidence, current market turnover exhibits a positive relationship with lagged market returns, and thus, contemporaneous market returns are not considered in the overconfidence hypothesis. Secondly, correlation does not imply causality, and therefore, whilst market turnover and market returns may be correlated, the market turnover and market return may not cause each other (Wolde-Rufael, 2007).

As can be seen in Table 5.7, the correlation between market turnover and the exogenous variables (market volatility and market dispersion) is negative for the market of JSE-listed ETFs with South African

benchmarks but positive for the market of JSE-listed ETFs with non-South African benchmarks. Regarding individual security turnover, Table 5.8 indicates that the individual security turnover and volatility is negatively correlated for the market of ETFs with local benchmarks but positively correlated for the market of ETFs with offshore benchmarks. Additionally, for both market portfolios, Table 5.7 suggests that market volatility and market dispersion are positively correlated. However, the correlation coefficient between market volatility and market dispersion is 0.7829 for the market of JSE-listed ETFs replicating domestic benchmarks and 0.7366 for the market of JSE-listed ETFs replicating international benchmarks, thus, indicating that these variables are highly correlated. Similarly, in Table 5.8, there is a high, positive correlation between the individual security return of the ETFs tracking offshore benchmarks and its market return. This high correlation between the explanatory variables could lead to a phenomenon referred to as multicollinearity.

In this study, the high correlation observed is due to market volatility and market dispersion being a product of market returns. Therefore, multicollinearity amongst the variables is not a cause for concern. This is because, multicollinearity can safely be ignored when one or more of the variables are computed using another variable in the regression (O'Brien, 2007; Allison, 2012; and Johnston, Jones, and Manley, 2018). O'Brien (2007) argues that the primary purpose of multiple regressions analysis is to control for the effects of relevant independent variables, and thus, multicollinearity is an insufficient reason to drop a highly correlated control variable. Moreover, Allison (2012) contends that, when variables are computed using another variable in the regression, multicollinearity has no adverse consequences since the p-values of the variables will not be affected. Furthermore, Statman, *et al.* (2006) maintains that the inclusion of the highly correlated exogenous variables (market volatility and market dispersion) to account for alternative explanations of trading activity does not materially change the results of their study on investor overconfidence. Overall, according to Juselius (2015), the VAR model is a solution to the multicollinearity between exogenous variables problem.

5.2.5. Model Diagnostics Tests

Diagnostics tests are performed to ensure that the estimated models which are analysed in this study are unbiased and reliable. For research questions one and two, diagnostics tests for serial correlation and heteroskedasticity are performed only on the estimated market VAR models since the respective panel VAR models are estimated using individual ETFs that form part of the computation of the market variables. Additionally, for research questions three and four, diagnostics tests for ARCH effects are performed only on the EGARCH models estimated for the full sample because the EGARCH models analysed in the sub-

period analysis are estimated using subsamples of the full sample period. The results from the tests for serial correlation, heteroskedasticity and ARCH effects are discussed below.

5.2.5.1. Test for Serial Correlation

The LM test for serial correlation is conducted with 10 lags for the market VAR models. The results for the market VAR model estimated for the market of ETFs with domestic benchmarks are presented in Table 5.10 while Table 5.11 (page 94) contains the results of the market VAR model estimated for the market of ETFs with international benchmarks.

The results in Table 5.10 and Table 5.11 report that the probability values associated with each LRE statistic are greater than 0.01 for all lags in both markets. Given that the level of significance (that is, 1%) is less than the associated probability values, the null hypothesis of no serial correlation cannot be rejected at a 1% level of significance for all lags in both markets. Therefore, the VAR models estimated for both markets do not display serial correlation at a 1% level of significance. As such, the residuals of the VAR models estimated for research questions one and two do not exhibit serial correlation.

Table 5.10: LM Test Results for the Market VAR Model Estimated for the Market of ETFs With Domestic Benchmarks

Lag (<i>h</i>)	LRE stat	Df	<i>Prob.</i>
1	4.5930	4	<i>0.3317</i>
2	1.4699	4	<i>0.8320</i>
3	8.8724	4	<i>0.0644</i>
4	5.8292	4	<i>0.2123</i>
5	5.0846	4	<i>0.2787</i>
6	4.8900	4	<i>0.2988</i>
7	1.8966	4	<i>0.7548</i>
8	1.2539	4	<i>0.8691</i>
9	4.8087	4	<i>0.3075</i>
10	1.2043	4	<i>0.8774</i>

Notes:

1. Null hypothesis: No serial correlation at lag *h*.
2. Eviews reports the LRE statistic which is a modified version of the LM statistic and is denoted by the LRE stat in the Table. *Prob.* represents the probability value associated with each LRE statistic.

Table 5.11: LM Test Results for the Market VAR Model Estimated for the Market of ETFs With International Benchmarks

Lag (<i>h</i>)	LR stat	Df	<i>Prob.</i>
1	5.0847	4	<i>0.2787</i>
2	1.0519	4	<i>0.9018</i>
3	3.8630	4	<i>0.4249</i>
4	2.3967	4	<i>0.6632</i>
5	1.0111	4	<i>0.9081</i>
6	2.5192	4	<i>0.6412</i>
7	6.0321	4	<i>0.1968</i>
8	2.3444	4	<i>0.6727</i>
9	1.2903	4	<i>0.8630</i>
10	2.1605	4	<i>0.7063</i>

Notes:

1. Null hypothesis: No serial correlation at lag *h*.
2. Eviews reports the LRE statistic which is a modified version of the LM statistic and is denoted by the LRE stat in the Table. *Prob.* represents the probability value associated with each LRE statistic.

5.2.5.2. Test for Heteroskedasticity

The White test for heteroskedasticity is conducted on both market VAR models, and the results for the joint test are presented in the ensuing tables. Table 5.12 contains the joint test results for the market VAR model estimated for the market of ETFs with local benchmarks whilst Table 5.13 (page 95) presents the results for the market VAR model estimated for the market of ETFs with international benchmarks.

Table 5.12: Heteroskedasticity Test Results for the Market VAR Model Estimated for the Market of ETFs With Domestic Benchmarks

Chi-sq	df	<i>Prob.</i>
161.9138	132	<i>0.0394</i>

Notes:

1. The results presented are for the joint heteroskedasticity test.
2. Null hypothesis: No heteroskedasticity.
3. Chi-sq denotes the chi-squared value and *prob.* represents its associated probability value.

In Tables 5.12 and 5.13, the probability values associated with the chi-square statistics are greater than 0.01. Thus, at a 1% level of significance, the null hypothesis of a constant variance in the errors of the respective market VAR model cannot be rejected. This being so, the VAR models estimated for both markets do not exhibit heteroskedasticity since their errors are homoskedastic at a 1% level of significance. Given this, the

residuals of the VAR models estimated for research questions one and two do not exhibit heteroskedasticity. Overall, the results of the diagnostics tests conducted on the market VAR models suggest that the estimated VAR models for research questions one and two are unbiased and reliable since the errors of these models do not display serial correlation or heteroskedastic behaviour.

Table 5.13: Heteroskedasticity Test Results for the Market VAR Model Estimated for the Market of ETFs With International Benchmarks

Chi-sq	df	Prob.
155.3076	132	0.0810

Notes:

1. The results presented are for the joint heteroskedasticity test.
2. Null hypothesis: No heteroskedasticity.
3. Chi-sq denotes the chi-squared value and *prob.* represents its associated probability value.

5.2.5.3. Test for ARCH Effects

To ensure that the estimated EGARCH (1,1) models are reliable, the ARCH-LM diagnostics test is conducted on the full sample EGARCH (1,1) models, and the results are provided in Table 5.14 and Table 5.15.

Table 5.14: ARCH-LM Diagnostics Test Results for the EGARCH (1,1) Model Estimated for the Market of ETFs With Domestic Benchmarks

F-stat.	Prob.	Obs*R-squared	Prob.
0.0535	0.9999	0.4287	0.9999

Notes:

1. Null hypothesis: No ARCH effects present in the residuals.
2. F-stat, *prob* and obs*R-squared refers to the F-statistic, probability value, and the number of observations multiplied by the R-squared, respectively.

Table 5.15: ARCH-LM Diagnostics Test Results for the EGARCH (1,1) Model Estimated for the Market of ETFs With International Benchmarks

F-stat.	Prob.	Obs*R-squared	Prob.
0.8342	0.5723	6.6782	0.5717

Notes:

1. Null hypothesis: No ARCH effects present in the residuals.
2. F-stat, *prob* and obs*R-squared refers to the F-statistic, probability value, and the number of observations multiplied by the R-squared, respectively.

In Tables 5.14 and 5.15, Obs*R-squared represents the LM-test statistic as specified in equation 4.23. The probability values associated with the F-statistics and the LM-statistics are greater than 0.01 in both tables. Since the probability values are greater than the 1% level of significance, the null hypothesis of no ARCH effects in the model's residuals up to lag order 8 cannot be rejected at a 1% level of significance. This being so, the null hypothesis of no ARCH effects in the residuals is accepted for both models, thereby, suggesting that the EGARCH (1,1) models estimated for both markets do not exhibit ARCH effects in the residuals. Therefore, the conditional variance equations are correctly specified and the EGARCH (1,1) models estimated for research questions three and four are reliable. Overall, the models estimated in this study are reliable, and thus, can be analysed. Hence, the ensuing sections provide an analysis of the results obtained from the estimated models.

5.3. Empirical Results of Market-Wide Investor Overconfidence

For research question one, the presence of market-wide investor overconfidence is examined for the market of South African ETFs tracking domestic benchmarks as well as the market of South African ETFs tracking international benchmarks. Market-wide investor overconfidence is detected by examining the long-run relationship between current market turnover and lagged market returns using VAR models, Granger causality tests and impulse response functions. However, prior to the estimation of the VAR models, the optimal lag lengths of the endogenous and exogenous variables are selected based on the information criteria and the likelihood ratio (LR) test. The next section outlines the selection of the optimal lag lengths for the VAR models.

5.3.1. Optimal Lag Length Selection

For each VAR model, the optimal lag lengths of the endogenous and exogenous variables are selected based on the information criteria and the likelihood ratio (LR) test. The information criteria and LR statistics used to select the optimal lag lengths of the endogenous variables are based on the unrestricted VAR which imposes no lag on the exogenous variables. Table 5.16 (page 97) outlines the information criteria and LR test statistics used to select the optimal lag length of the endogenous variables in the VAR model estimated for the market of ETFs tracking domestic benchmarks while Table 5.17 (page 97) presents these statistics for the VAR model estimated for the market of ETFs tracking international benchmarks.

For the unrestricted VAR model estimated for the market of ETFs with domestic benchmarks, Table 5.16 shows that the AIC is minimised at 6 lags whilst the SC and HQ are minimised at 2 lags. In this case, the information criteria present conflicting results, and thus, the optimal lag length is selected based on the AIC

Table 5.16: Lag Order Selection Criteria for the Endogenous Variables in the VAR Model Estimated for the Market of ETFs Tracking Domestic Benchmarks

Lag	LR	AIC	SC	HQ
0	NA	-6.3128	-6.2196	-6.2752
1	86.7387	-6.6833	-6.5281	-6.6206
2	23.0281	-6.7558	-6.5384*	-6.6680*
3	8.1257	-6.7579	-6.4785	-6.6451
4	3.3212	-6.7373	-6.3957	-6.5993
5	15.4328	-6.7759	-6.3722	-6.6128
6	15.8763*	-6.8174*	-6.3516	-6.6293
7	5.1376	-6.8062	-6.2784	-6.5930
8	4.3739	-6.7915	-6.2016	-6.5532

Notes:

1. * indicates lag order selected by the criterion
2. LR refers to the LR test statistic at 5% level, AIC refers to the Akaike information criterion, SC refers to the Schwarz information criterion and HQ refers to the Hannan-Quinn information criterion.

Table 5.17: Lag Order Selection Criteria for the Endogenous Variables in the VAR Model Estimated for the Market of ETFs Tracking International Benchmarks

Lag	LR	AIC	SC	HQ
0	NA	-7.7614	-7.6456	-7.7144
1	37.1495	-7.9523	-7.7593	-7.8740
2	12.4069	-7.9837	-7.7134	-7.8739
3	26.4088*	-8.1094*	-7.7620*	-7.9683*
4	1.5646	-8.0697	-7.6450	-7.8972
5	1.5586	-8.0300	-7.5282	-7.8262
6	2.5068	-7.9971	-7.4181	-7.7620
7	4.6289	-7.9794	-7.3232	-7.7129
8	3.3472	-7.9530	-7.2195	-7.6551

Notes:

1. * indicates lag order selected by the criterion
2. LR refers to the LR test statistic at 5% level, AIC refers to the Akaike information criterion, SC refers to the Schwarz information criterion and HQ refers to the Hannan-Quinn information criterion.

since AIC reveals more features of the data (Lin, *et al.*, 2010). Based on the AIC, the optimal lag length of the endogenous variables is 6 lags. The LR test further supports 6 lags as being optimal for the endogenous variables in the VAR model estimated for the market of ETFs replicating local benchmarks. Table 5.17 shows that the information criteria and the LR test provide consistent results with regards to the lag order

selection of the endogenous variables in the VAR model estimated for the market of ETFs replicating offshore benchmarks. Regarding the information criteria, the AIC, SC, and HQ are minimised at 3 lags. Moreover, the LR test selects 3 lags as being optimal for the endogenous variables. Thus, 3 lags are chosen as the optimal lag length of the endogenous variables in the VAR model employed for the market of ETFs tracking international benchmarks.

Following Alsabban and Alarfaj (2020), the optimal lag length of the exogenous variables is selected by constructing VAR models that include the chosen optimal lag length of the endogenous variables and an addition of lags of the exogenous variables -as discussed in Section 4.3.1.1. Table 5.18 provides the AIC and SC values for each additional lag of the exogenous variables in the VAR model estimated for the market of ETFs with domestic benchmarks and Table 5.19 (page 99) provides these values for the VAR model estimated for the market of ETFs with international benchmarks.

Table 5.18: Lag Order Selection Criteria for the Exogenous Variables in the VAR Model Estimated for the Market of ETFs Tracking Domestic Benchmarks

Lag	AIC	SC
1	-6.7636	-6.2391*
2	-6.7564	-6.1703
3	-6.7940	-6.1461
4	-6.8047*	-6.0952
5	-6.7916	-6.0203
6	-6.7683	-5.9353
7	-6.7610	-5.8635
8	-6.7663	-5.8037

Notes:

1. * indicates lag order selected by the criterion
2. AIC refers to the Akaike information criterion and SC refers to the Schwarz information criterion.

The information criteria values contained in Table 5.18 and Table 5.19 indicates that the AIC and SC provide inconsistent results regarding the optimal number of lags for the exogenous variables. Hence, the optimal lag length is chosen based on the AIC which offers better insight into the data. The AIC selects 4 and 7 lags of the exogenous variables as being optimal for the VAR models estimated for the market of ETFs tracking domestic and international benchmarks, respectively. Overall, the lag order selection criteria select 6 lags of the endogenous variables and 4 lags of the exogenous variables for the VAR model estimated for the market of ETFs tracking domestic benchmarks. For the VAR model estimated for the market of ETFs tracking international benchmarks, the lag order selection criteria choose 3 lags of the endogenous

variables and 7 lags of the exogenous variables as being optimal. The lag lengths selected as optimal are used to estimate the VAR models which are discussed in the forthcoming section.

Table 5.19: Lag Order Selection Criteria for the Exogenous Variables in the VAR Model Estimated for the Market of ETFs Tracking International Benchmarks

Lag	AIC	SC
1	-7.9554	-7.5395*
2	-7.9113	-7.4199
3	-7.9832	-7.4162
4	-7.9360	-7.2907
5	-7.8796	-7.1553
6	-7.9603	-7.1564
7	-8.0571*	-7.1730
8	-8.0340	-7.0689

Notes:

1. * indicates lag order selected by the criterion
2. AIC refers to the Akaike information criterion and SC refers to the Schwarz information criterion.

5.3.2. Market Vector Autoregression (VAR) Models

The overconfidence hypothesis proposed by Gervais and Odean (2001) posits that overconfident investors attribute market gains to their own security picking abilities, and, as a result, they trade more aggressively after periods of positive market gains. Thus, the overconfidence hypothesis predicts that, in the presence of overconfident trading, historical market returns exhibit a positive influence on current trading activity. To assess the lead-lag relationship between trading activity and returns, this study's primary tool of analysis is VAR models which contain market turnover, *mturn*, and market returns, *mret*, as endogenous variables. In addition, the VAR models control for market volatility, *msig*, and cross-sectional return dispersion, *disp*.

The VAR model specified in equation 4.24 is estimated for the market of JSE-listed ETFs tracking South African benchmarks and the results for the regression with market turnover as the dependent variables are presented in Table 5.20 (page 100) while Appendix 2 (page 147) provides the results of the regression with market return as the dependent variable. In addition, Table 5.21 (page 101) provides the results of the regression with market turnover as the dependent variable in the VAR model estimated for the market of JSE-listed ETFs tracking non-South African benchmarks while Appendix 3 (page 148) provides the results of the regressions with market return as the dependent variable. In Table 5.20, Table 5.21, Appendix 2 and Appendix 3, the dependent variables (*mturn* and *mret*) are organised in rows while the lagged dependent

variable and exogenous variable coefficients appear in columns. For each coefficient, the t-test statistic and probability value are reported.

Table 5.20: VAR Model Estimates for the Market of ETFs Tracking Domestic Benchmarks

	Mturn_dom (-1)	Mturn_dom (-2)	Mturn_dom (-3)	Mturn_dom (-4)	Mturn_dom (-5)	Mturn_dom (-6)
Mturn_dom	0.3245*** (4.5732)	0.2523*** (3.4389)	0.1232 (1.6453)	-0.0588 (-0.7736)	0.0627 (0.8476)	0.1042 (1.4747)
	Mret_dom (-1)	Mret_dom (-2)	Mret_dom (-3)	Mret_dom (-4)	Mret_dom (-5)	Mret_dom (-6)
Mturn_dom	0.1143* (1.9303)	0.1500** (2.5113)	-0.0238 (-0.3833)	-0.0879 (-1.4281)	-0.0187 (-0.8043)	0.0128 (0.5079)
	C	Msig_dom	Msig_dom (-1)	Msig_dom (-2)	Msig_dom (-3)	Msig_dom (-4)
Mturn_dom	0.0200** (2.5358)	-0.0191 (-0.7153)	-0.0275 (-0.9515)	0.0168 (0.5885)	0.0621** (2.1797)	-0.0117 (-0.4440)
	Disp_dom	Disp_dom (-1)	Disp_dom (-2)	Disp_dom (-3)	Disp_dom (-4)	
Mturn_dom	0.2783 (1.3801)	-0.4583 (-1.2724)	-0.6918** (-1.9961)	-0.4453 (-1.2033)	0.4126 (1.1515)	

Notes:

1. Values in brackets '()' represents t-statistics.
2. ***, **, * indicate statistical significance at 1%, 5%, and 10% levels, respectively.

Table 5.20 shows that, for the market of ETFs tracking domestic benchmarks, market turnover is autocorrelated since current market turnover can be explained by lagged market turnover. With the exception of the fourth lag ($mturn_{t-4}$), there is a positive relationship between current market turnover and lagged market turnover, although only the first and second $mturn$ lag coefficients are statistically significant at a 1% level of significance. Moreover, from the third and higher $mturn$ lags, the coefficients are statistically insignificant, indicating that market turnover is only influenced by its own behaviour in the past two months. Thus, autocorrelation in the trading activity of the market of ETFs tracking domestic benchmarks persists only for a short period of time. Notably, the statistically significant lagged $mturn$ coefficients are generally decreasing in magnitude for the market of ETFs with domestic benchmarks. In contrast, for the market of ETFs tracking international benchmarks, the significant lagged $mturn$ coefficients are generally increasing in magnitude. Like the market of ETFs with domestic benchmarks, for

the market of ETFs tracking international benchmarks, there is a positive relationship between current market turnover and lagged market turnover, however, only the first and third *mturn* lag coefficients are significant at a 99% confidence interval whilst the second *mturn* lag coefficient is statistically insignificant. Therefore, the market turnover of both markets is autocorrelated.

Table 5.21: VAR Model Estimates for the Market of ETFs Tracking International Benchmarks

	Mturn_int (-1)	Mturn_int (-2)	Mturn_int (-3)	Mret_int (-1)	Mret_int (-2)	Mret_int (-3)
Mturn_int	0.2771*** (3.6968)	0.0278 (0.3654)	0.3341*** (4.7361)	0.0535 (1.2253)	0.1037** (2.3271)	0.0590 (1.2841)
	C	Msig_int	Msig_int (-1)	Msig_int (-2)	Msig_int (-3)	Msig_int (-4)
Mturn_int	0.0087 (1.0510)	0.2408** (1.9847)	0.1965 (1.5992)	0.0537 (0.4337)	-0.4650*** (-3.7590)	0.0528 (0.4144)
	Msig_int (-5)	Msig_int (-6)	Msig_int (-7)	Disp_int	Disp_int (-1)	Disp_int (-2)
Mturn_int	0.0353 (0.2825)	0.1830 (1.4964)	-0.0017 (-0.0141)	-0.0299 (-0.0542)	-0.9216* (-1.6612)	-0.5369 (-0.9660)
	Disp_int (-3)	Disp_int (-4)	Disp_int (-5)	Disp_int (-6)	Disp_int (-7)	
Mturn_int	2.0763*** (3.7451)	-0.4384 (-0.7650)	-0.0257 (-0.0452)	-0.2367 (-0.4205)	-0.5914 (-1.0947)	

Notes:

1. Values in brackets ‘()’ represents t-statistics.
2. ***, **, * indicate statistical significance at 1%, 5%, and 10% levels, respectively.

The finding of autocorrelated market trading volume is consistent with Covrig and Ng (2004) who document that 95% of NYSE and AMEX stocks exhibit statistically significant trading volume autocorrelation. Similarly, Blau and Smith (2014) report that the daily short-sale volume of common shares trading on the NYSE is autocorrelated. Autocorrelation in ETF market trading volume can be explained by the flow of new information to ETF markets. Specifically, He and Wang (1995) argue that trading in competitive markets take place when traders receive new information, either private or public. On the one hand, trading tends to be correlated across time when the flow of new public information to the market is serially correlated (He and Wang, 1995). On the other hand, when information is private, clustering in

trading may be due to the arrival of independent information (He and Wang, 1995). Autocorrelation in trading volume may also be generated by liquidity constraints (Blau and Smith, 2014). Blau and Smith (2014) mention that an investor who wants to trade a large number of shares may have to spread the large trade across time as liquidity become available.

As can be seen from the results contained in Table 5.20 and Table 5.21, market turnover is also dependent on lagged market returns, even after controlling for lagged market turnover, market volatility, and dispersion. Specifically, for the market of ETFs tracking domestic benchmarks, when *mturn* is the dependent variable, the first *mret* lag coefficient is equal to 0.1143 and statistically significant at a 10% level of significance while the second lagged *mret* coefficient is equal to 0.1500 and significant at a 5% level of significance. Notably, from the third and higher lags of *mret*, the coefficients are statistically insignificant. For the market of ETFs tracking international benchmarks, the coefficient of the second lagged *mret* is equal to 0.1037 and significant at a 5% level of significance, although, the first and third lags of *mret* are statistically insignificant. For both markets, the positive and significant relationship between current market turnover and lagged market returns supports the overconfidence hypothesis, and therefore, represents the first key empirical finding of this study. This positive relationship between past market returns and current trading activity suggests that an increase in past market returns increases current trading because investors' confidence in their trading skills increases after market gains. Similarly, a decrease in lagged market returns reduces current trading because investors' confidence in their trading skills decreases after periods of negative market returns. A more detailed discussion of this key finding ensues in the context of impulse response functions (refer to Section 5.3.4).

Consistent with Karpoff (1987), the results indicate that market volatility is a significant explanatory variable when market turnover is the dependent variable. For the market of ETFs with international benchmarks, current market volatility has a significant, positive effect on market turnover since the coefficient of contemporaneous *msig* is 0.2408 and significant at a 5% level of significance. This finding suggests that, for the market of JSE-listed ETFs replicating international benchmarks, an increase (decrease) in market volatility leads to an increase (decrease) in the contemporaneous market turnover. Additionally, the third lag of *msig* is highly significant, however, the coefficient is negative. This implies that there is an inverse relationship between current market turnover and the market volatility of three months prior, such that, an increase (decrease) in the market volatility of three months prior will reduce (increase) current market turnover. On the contrary, for the market of ETFs with domestic benchmarks, there exists a significant, positive relationship between current market turnover and the market volatility of three months

prior since the third lag of *msig* is equal to 0.0621 and significant at a 5% level of significance. Notably, the remaining *msig* variables are statistically insignificant.

This relationship between trading activity and market volatility is also documented in overconfidence studies by Statman, *et al.* (2006) and Gupta, *et al.* (2018), however, these studies examine stock-level data. Interestingly, in line with Statman, *et al.* (2006), the market volatility coefficients obtained for the market of ETFs replicating offshore benchmarks indicate that the volume-volatility relationship changes signs in subsequent months. The volume-volatility relationship -either positive or negative- can be attributed to the degree to which ETF prices convey information, the manner in which information about ETFs are disseminated, the rate at which ETF information flows to the market as well as the existence of any short-sale restrictions on ETFs (Karpoff, 1987). Particularly, ETF market volatility can be interpreted as the market's assessment of information about ETFs while the corresponding trading volume is indicative of the degree to which investors differ in the way they interpret the information, and subsequently, form ETF investment decisions (Beaver, 1968). The relationship between contemporaneous volatility and turnover for the market of ETFs tracking international benchmarks is consistent with the Mixture of Distribution hypothesis while the lead-lag relationship between current market turnover and lagged market volatility for both markets is consistent with the Sequential Arrival of Information hypothesis. For a detailed discussion of the Mixture of Distribution hypothesis and the Sequential Arrival of Information hypothesis refer to Section 4.3.1.

Regarding the second control variable (market dispersion), for both markets, there is no significant relationship between current market turnover and contemporaneous market dispersion since contemporaneous *disp* is an insignificant explanatory variable in the regression with *mturn* as the dependent variable. However, there is a significant relationship between current market turnover and lagged market return dispersion. Particularly, for the market of ETFs with domestic benchmarks, there exists a significant, negative relationship between current turnover and only the second lagged *disp* variable whilst the remaining lags of *disp* are not statistically significant. This implies that, for the respective market, an increase (decrease) in the dispersion of ETF market returns two months prior results in a decrease (increase) in current market turnover. Similarly, there is a significant lead-lag relationship between current turnover and lagged dispersion for the market of ETFs with international benchmarks. However, only the first and third lagged *disp* coefficients are significant at a 10% and 1% level, respectively. Notably, the coefficient of the first lag of *disp* is negative whilst the third *disp* lag coefficient is positive and more than twice the magnitude of the first lagged *disp* coefficient.

Consistent with Statman, *et al.* (2006), this study reports a significant sign-changing relationship between current market turnover and the control variable, market dispersion, for both markets. The significant, positive lead-lag relationship between current trading activity and lagged dispersion is also found by Chordia, Huh, and Subrahmanyam (2006), however, their study is conducted at a company-stock level. This positive association between current trading activity and lagged dispersion can be attributed to the portfolio rebalancing needs of investors. Specifically, an increase in the dispersion of returns induces trading activity because of investors' portfolio rebalancing needs which are triggered by the historical price performance of ETFs. On the contrary, this study also documents a negative relationship between trading activity and lagged return dispersion. According to Domowitz, Glen, and Madhavan (2001), risk-averse investors may reduce their trading activity after the dispersion of returns increases, resulting in a negative relationship between market turnover and lagged market dispersion. Overall, the control variables, market volatility and dispersion, have a significant impact on the market turnover of both markets.

Interestingly, the results presented in Appendix 2 (page 147) and Appendix 3 (page 148) suggest that the surveyed ETFs are not even in a state of weak-form market efficiency since there exist statistically significant explanatory variables when market return is the dependent variable in the regression. Particularly, for the market of ETFs tracking domestic benchmarks, market return exhibits significant autocorrelation since the third, fifth, and sixth lagged *mret* coefficients are statistically significant when market return is the dependent variable. However, the market return prediction anomaly exhibits a significant, positive autocorrelation at the third and sixth lagged *mret* whilst there is a significant, negative autocorrelation at the fifth *mret* lag. Notably, the magnitude of the significant, positive autocorrelation coefficients is more than twice the magnitude of the significant, negative autocorrelation coefficient. Contrarily, the market return of the market of ETFs with international benchmarks exhibits no significant autocorrelation.

The presence of significant autocorrelation in the market returns of the market of ETFs replicating domestic benchmarks may be caused by non-synchronous trading, time-varying expected returns, or conditional risk premia as well as feedback trading strategies. According to Lo and MacKinlay (1990), autocorrelation in the returns of a market may be caused by non-synchronous trading since some securities in the market do not trade frequently. Generally, the returns of ETFs that are frequently traded respond to new news immediately whilst the returns of ETFs that are infrequently traded incorporate the respective news with a lag, thus, inducing autocorrelation in the returns of the market of ETFs tracking domestic benchmarks. In addition, changes in the expected returns or risk premia of ETFs may induce autocorrelation in the returns of the respective market. This is because, variables used to forecast expected returns may be characterized

by autoregressive processes, and as a result, the predetermined variables associated with realised ETF returns exhibit high autocorrelation (Conrad and Kaul, 1988). Moreover, Sentana and Wadhvani (1992) posit that returns will be autocorrelated if investors follow feedback strategies (that is, they react to price changes). Positive feedback-traders induce negative autocorrelation of ETF returns but negative feedback-traders induce positive autocorrelation of ETF returns (Sentana and Wadhvani, 1992; McKenzie and Faff, 2003). As reviewed in Section 3.5.1, Charteris, *et al.*, (2014) examine emerging market ETFs and document that South African ETF investors display feedback trading strategies which may account for the ETF market return autocorrelation observed in this study.

The results obtained also indicate that the control variables, market volatility and dispersion, have an impact on market return. With regards to market volatility, there exists a significant, inverse relationship between the current market returns and its market volatility of four months prior for the market of ETFs tracking domestic benchmarks. Interestingly, for the market of ETFs tracking international benchmarks, there is a significant relationship between current market return and its market volatility of seven months prior, however, this relationship is positive. Whilst Nelson (1991) argues that the relationship between market return and market volatility can be either positive or negative, the positive relationship between the returns of the market of ETFs tracking international benchmarks and its volatility is consistent with many traditional asset pricing models. These asset pricing models assert that, for any asset, an increase in risk (measured by volatility) is accompanied by an increase in returns. On the contrary, the negative relationship between the returns of the market of ETFs tracking domestic benchmarks and its lagged volatility can be attributed to the volatility feedback effect put forward by French, *et al.* (1987). The volatility feedback effect suggests that, if volatility is priced, increases in anticipated market volatility causes the required rate of return on ETFs to increase. As a result, ETF prices decline, leading to negative ETF market returns. This finding suggests that the time-varying risk premium theory of traditional asset pricing models cannot be used to describe the behaviour of ETFs tracking domestic benchmarks.

Regarding the cross-sectional market return dispersion, the return of the market of ETFs tracking domestic benchmarks displays a significant relationship with both contemporaneous and the fifth lagged *disp* variables. However, the relationship between market returns and contemporaneous dispersion is positive whilst there is a negative relationship between market returns and the fifth *disp* lag coefficient. Similarly, for the market of ETFs tracking international benchmarks, current market returns exhibit a significant, negative relationship with the seventh lagged *disp* coefficient. According to Jiang (2010) and Chen, Demirer and Jategaonkar (2015), the cross-sectional dispersion of returns captures a dimension of uncertainty that is related to fundamental economic transitions and economic restructuring that cannot be

accounted for by market risk factors. Hence, the return dispersion carries a significantly positive price of risk, which accounts for the positive relationship between the market return dispersion and the market return of the market of ETFs tracking domestic benchmarks. However, for both markets, the inverse relationship between lagged market return dispersion and current market returns is inconsistent with the theory that ETF investors demand compensation for bearing risk associated with imperfect diversification (Sehgal and Garg, 2016). Nevertheless, the negative relationship between lagged market return dispersion and current market returns is consistent with the findings of Verousis and Voukelatos (2018) who examine stock-level data and report that dispersion is associated with a negative risk premium because there exists a negative relationship between the cross-sectional dispersion of returns and the expected returns.

Interestingly, the results presented in Appendix 2 and Appendix 3 suggest that, for both markets, historical market turnover cannot be used to predict market return since the lagged *mturn* coefficients are not statistically significant when *mret* is the dependent variable. Similarly, Tapa and Hussin (2016) report a statistically insignificant relationship between current return and both current and lagged trading volume. However, this finding contradicts Attari, Rafiq and Awan (2012) and Darwish (2012) who find that lagged trading volume can be used to predict market returns. To further investigate the causality between the market turnover and the market returns of each market, Granger causality tests are conducted, and the results are discussed in the next section.

5.3.3. Granger Causality Tests

The Granger causality test proposed by Granger (1969) is employed to examine the causal relationship between market turnover and market returns. The null hypotheses and the possible outcomes of this test were presented under Section 4.3.1.2. Results for the market of ETFs tracking domestic benchmarks are provided in Table 5.22 (page 107) while Table 5.23 (page 107) contains the results for the market of ETFs tracking international benchmarks.

The results presented in Table 5.22 and Table 5.23 show that, when market turnover is the dependent variable, the chi-square statistics associated with market return have p-values that are less than 0.10. Thus, for both markets, the null hypothesis that market return does not cause market turnover is rejected at a 10% level of significance. This implies that, for both the market of JSE-listed ETFs tracking South African benchmarks as well as the market of JSE-listed ETFs tracking non-South African benchmarks, lags of market return does Granger cause the market turnover. On the contrary, when market return is the dependent variable, the chi-square statistics associated with market turnover have probability values that are greater than 10%. Thus, the null hypothesis that lags of market turnover does not cause market return cannot be

rejected for both markets. Overall, for both markets, there is a unidirectional Granger causality from market return to market turnover. This finding is consistent with the results obtained from the VAR models in which the lags of market turnover are statistically insignificant when market return is the dependent variable.

Table 5.22: Granger Causality Test Results for the Market of ETFs Tracking Domestic Benchmarks

Dependent variable: Mturn_dom			
Excluded	Chi-sq	df	Prob.
Mret_dom	10.72660	6	0.0972*
All	10.72660	6	0.0972*
Dependent variable: Mret_dom			
Excluded	Chi-sq	df	Prob.
Mturn_dom	3.062349	6	0.8010
All	3.062349	6	0.8010

Notes:

1. Chi-sq represents the chi-square value and *prob.* denotes its associated probability value.
2. * indicates significance at a 10% level of significance.

Table 5.23: Granger Causality Test Results for the Market of ETFs Tracking International Benchmarks

Dependent variable: Mturn_int			
Excluded	Chi-sq	df	Prob.
Mret_int	7.459611	3	0.0586*
All	7.459611	3	0.0586*
Dependent variable: Mret_int			
Excluded	Chi-sq	df	Prob.
Mturn_int	3.880846	3	0.2746
All	3.880846	3	0.2746

Notes:

1. Chi-sq represents the chi-square value and *prob.* denotes its associated probability value.
2. * indicates significance at a 10% level of significance.

Despite the finding that, for both markets, lags of market return Granger causes market turnover, further analysis is needed to confirm the presence of investor overconfidence. This is because the overconfidence hypothesis posits that lagged market return has a positive impact on current market turnover, however, the Granger causality test does not provide any information regarding the sign of the causality (Liddle and Lung, 2013). To overcome this limitation of the Granger causality test, impulse response functions are inspected in the next section.

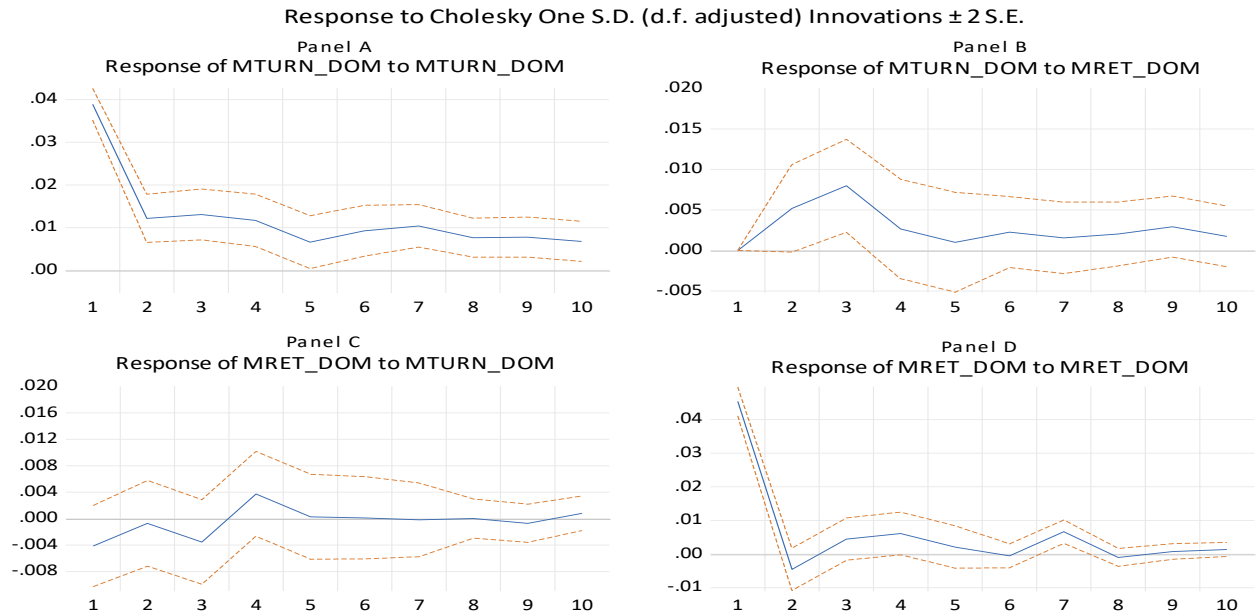
5.3.4. Market Impulse Response Functions

An important component of this analysis is to investigate how market turnover responds to shocks in market returns. However, the individual VAR coefficients do not fully capture the effect of other endogenous and exogenous variables (Swanson and Granger, 1997; Dewachter, Iania, Lyrio and de Sola, 2015). On the contrary, impulse response functions trace the impact of a residual shock by using all the VAR coefficient estimates, thereby, fully capturing the effect of other variables in the system (Swanson and Granger, 1997; Dewachter, *et al.*, 2015). Specifically, an impulse response function traces the impact of a shock to one of the variables on the current and future values of the endogenous variables. Based on the market-wide VAR models in Table 5.20 and Table 5.21, impulse response functions are estimated for the respective models²⁴. Figure 5.4 (page 109) illustrates the impulse response functions associated with the VAR model estimated for the market of ETFs tracking domestic benchmarks and the impulse response functions associated with the VAR model estimated for the market of ETFs tracking international benchmarks are illustrated in Figure 5.5 (page 109). Notably, the impulse response is statistically significant when both standard error bands are below or above zero on the y-axis.

In each figure, panel A shows the response of *mturn* to a one standard deviation shock in *mturn* and panel B plots the response of *mturn* to a one standard deviation shock in *mret*. It is important to note that, because the *mturn* variable did not undergo a log transformation, the vertical axis in both panels A and B measures the unit change in *mturn* (that is, the market turnover ratio) as opposed to a percentage measurement. Additionally, panels C and D illustrate the response of *mret* to a one standard deviation shock in *mturn* and *mret*, respectively. The values assigned to each impulse response in Figure 5.4 is provided in the impulse response table in Appendix 4 (page 148) and the values associated with each impulse response in Figure 5.5 is provided in the impulse response table in Appendix 5 (page 150).

²⁴ The exogenous variables are included in the estimation of the impulse response functions, however, are not reported because they do not aid in answering the research questions of this study.

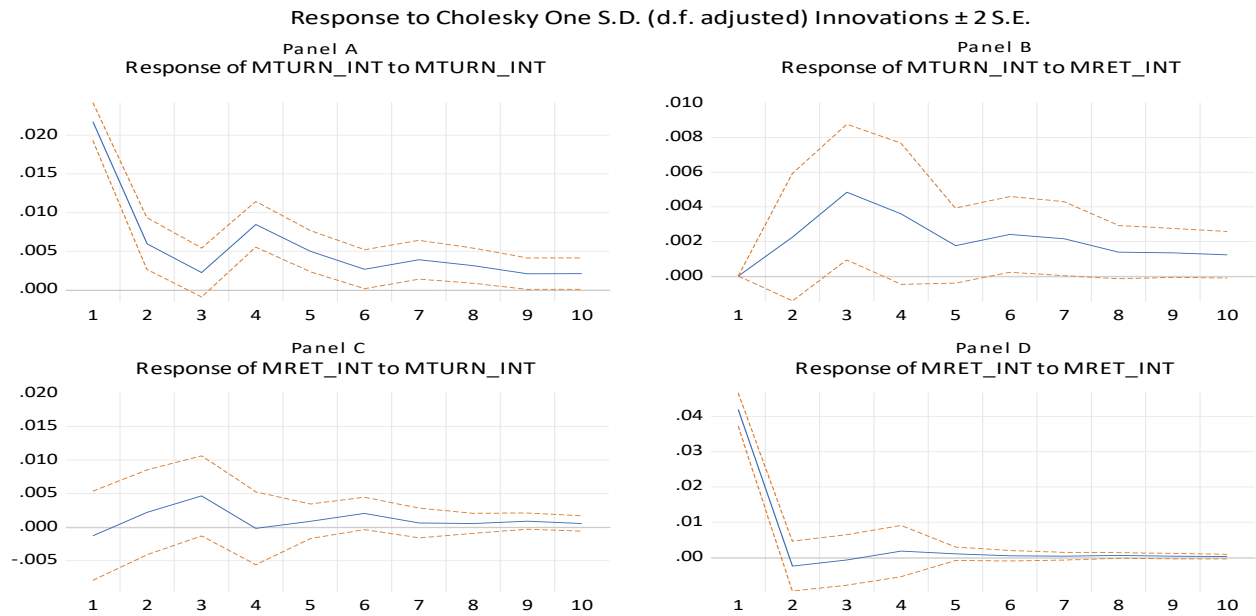
Figure 5.4: Impulse Response Functions for the VAR Model Estimated for the Market of ETFs with Domestic Benchmarks



Notes:

1. The blue line represents the impulse response and the red lines represent two standard error bands.

Figure 5.5: Impulse Response Functions for the VAR Model Estimated for the Market of ETFs with International Benchmarks



Notes:

1. The blue line represents the impulse response and the red lines represent two standard error bands.

The general downtrend observed in panel A of both figures suggests that the response of *mturn* to its own shocks decreases over time. Particularly, for the market of ETFs tracking domestic benchmarks, panel A in Figure 5.4 shows that a one standard deviation shock to *mturn* generates a 0.039 increase in the next month's market turnover ratio and an increase of approximately 0.012 in the following month's market turnover ratio. Similarly, panel A in Figure 5.5 shows that, for the market of ETFs with international benchmarks, a one standard deviation *mturn* shock results in an increase of 0.022 in the next month's market turnover ratio and a 0.006 increase in the following month's market turnover ratio. Except for the fourth period impulse response in the market of ETFs with offshore benchmarks, this downward trend generally persists for each additional period in both markets. Moreover, except for the fifth period response in the market of ETFs with domestic benchmarks and the third and ninth period responses in the market of ETFs with offshore benchmarks, the response of market turnover to its own shocks remain statistically significant and positive throughout the ten month period.

Overall, in both Figure 5.4 and 5.5, panel A confirms the serial dependence of market turnover since the positive impact of *mturn* shocks to *mturn* is significant for several months during the ten month period. This finding is consistent with the VAR results which document that both markets are subject to a significant positive relationship between current market turnover and lagged market turnover. Notably, this autocorrelation in trading volume may be generated by the manner in which new information flows to ETF markets or by liquidity constraints, and could indicate the presence of positive feedback trading, as discussed in Section 5.3.2.

Regarding the response of market turnover to shocks in market return, panel B in both Figures 5.4 and 5.5 show that, for both markets, a one standard deviation *mret* shock does not generate any impulse response in the market turnover of the next month. However, from the second period onwards, *mturn* responds positively to a one standard deviation shock in *mret*, and this positive response persists for at least ten months in both markets. Specifically, a one standard deviation shock in market return at the second lag results in an increase of 0.005 and 0.002 in the market turnover ratio for the market of ETFs tracking domestic and international benchmarks, respectively. Interestingly, for both markets, the highest response of market turnover to a one standard deviation *mret* shock is evident in period 3, and thereafter, the response of *mturn* starts decreasing but remains positive for all periods. Notably, for the market of ETFs tracking domestic benchmarks, only the third period response of *mturn* to *mret* is statistically significant whilst the third and sixth period responses of *mturn* to *mret* are statistically significant for the market of ETFs tracking international benchmarks.

For both markets, the statistically significant positive response of market turnover to shocks in market return provides evidence that improved market performance increases investors' confidence about their trading skills and security valuation abilities, causing them to trade more frequently. This finding corresponds with the results obtained from the VAR estimations and the Granger causality tests conducted. Therefore, for both markets, the results obtained from the VAR models, Granger causality tests and impulse response functions accept the overconfidence hypothesis which states that high market returns results in overconfident investors who trade more aggressively in subsequent periods. More specifically, this study provides evidence that investors trading in the market of South African ETFs that track local benchmarks suffer from the overconfidence bias. Additionally, the results obtained in this study suggest that investor overconfidence is also pronounced in the market of South African ETFs that track international benchmarks. The presence of investor overconfidence in both ETF markets can be attributed to positive market gains which cause investors to overestimate their own abilities or exaggerate the level of accuracy of their knowledge and information, and, as a result, they trade more aggressively in subsequent periods, generating an increase in market turnover.

Steyn (2019) finds that South African ETFs tracking international benchmarks track their benchmarks less efficiently because the returns of ETFs tracking offshore benchmarks are constrained by the treatment of dividends and the exchange rate volatility -as discussed in Section 4.2.4. Remarkably, this study finds that the overconfidence bias also influences the decisions made by traders in the market of JSE-listed ETFs tracking offshore benchmarks. This finding is consistent with Kyrychenko and Shum (2009) who report that overconfident investors are more likely to hold financial assets with foreign exposures. This is because most investors are interested in diversifying their portfolios instead of timing the market (Kyrychenko and Shum, 2009).

The first key empirical finding of this current study is that investor overconfidence is present in the market of South African ETFs with domestic benchmarks and in the market of South African ETFs with international benchmarks, which, when combined, makes up the total South African ETF market. This finding implies that the investment decisions of South African ETF investors are influenced by behavioural biases, suggesting that these investors are irrational. Moreover, the influence of behavioural biases, such as the observed overconfidence bias, on ETF investment choices contradicts the notion of market efficiency, thereby, suggesting that the South African ETF market is not efficient. This is further supported by the finding that returns from the South African ETF can be predicted -as discussed in Section 5.3.2. Overall, these results suggest that the South African ETF market is inefficient, thus, implying that smart money traders cannot arbitrage away the impact of overconfident traders. Therefore, overconfident traders play a

significant role in the South African ETF market, and their actions can move the prices of ETFs in either direction.

Furthermore, the overconfident trading present in the market of South African ETFs tracking domestic and international benchmarks may generate excess market volatility by distorting ETF prices further deteriorating the efficiency of the South African ETF market. More particularly, investor overconfidence causes investors to overestimate ETF prices, which could result in the formation of an ETF price bubble that could fuel the ETF market crash predicted by Michael Burry and several other prominent industry experts. Hence, the current study also examines the effect of investor overconfidence on the volatility of the South African ETF market. However, prior to the investigation of the effect of overconfidence on market volatility, it is important to ensure that the market-wide overconfidence observed is not a direct summation of the disposition effect. This is expressed as research question two, and the results of the analysis are presented in Section 5.4.

It is important to note that, this presence of investor overconfidence in the South African ETF market is consistent with findings from alternative asset class, such as²⁵; stocks (Statman, *et al.*, 2006), REITs (Lin, *et al.*, 2010), and mutual funds (Bailey, *et al.*, 2011). Moreover, whilst Dowie and Willows (2016) report that South African unit trust investors are underconfident, this study provides evidence the South African ETF investors are overconfident. This inconsistency may be because unit trusts have high management costs and trade only at their closing NAV, and therefore, are generally more costly to purchase in comparison to ETFs (Gastineau, 2001; Steyn, 2019). This cost advantage of ETFs may be realised in their higher market returns, thus, causing investors to be more overconfident when trading in ETFs. To the knowledge of the author of this study, there are no existing studies that specifically investigate investor overconfidence in ETF markets. Therefore, the implications of this key empirical finding are discussed in Chapter 6 under Section 6.3.

For completeness, in Figure 5.4 and Figure 5.5, panel C plots the response of market return to a one standard deviation shock in market turnover. For both markets, the response of *mret* to shocks in *mturn* is statistically insignificant for all periods. This finding conforms to the results of the estimated VAR models in which lagged *mturn* coefficients were statistically insignificant explanatory variables in the regressions with *mret* as the dependent variable. Panel D in Figures 5.4 and 5.5 plots the response of market return to its own shocks. The impulse response functions support the notion that the surveyed ETFs are not even

²⁵ These studies were discussed in Chapter 3.

weak-form efficient since market returns are autocorrelated. Specifically, the first period response of *mret* to a one standard deviation *mret* shock generates an increase of 0.045% and 0.042% in the next month's return of the market of JSE-listed ETFs with local and international benchmarks, respectively. Notably, for the market of ETFs tracking international benchmarks, only the first period response of *mret* to its own shocks is statistically significant whilst, for the market of ETFs tracking domestic benchmarks, the first and seventh period responses of *mret* to its own shocks are statistically significant. This statistically significant autocorrelation displayed by the market returns of both markets is consistent with the results reported for the VAR models in which lagged *mret* coefficients were significant explanatory variables in the regressions with *mret* as the dependent variable. This autocorrelation displayed by the market return of both markets may be generated through non-synchronous trading, time-varying expected returns or conditional risk premia as well as feedback trading strategies as discussed in Section 5.3.2.

In summary, for research question one, this study provides robust results in support of the overconfidence hypothesis. Specifically, for the market of South African ETFs tracking domestic benchmarks and for the market of South African ETFs tracking international benchmarks, the results obtained from the estimated VAR models indicate that, when market turnover is the dependent variable, lagged market return coefficients are statistically significant and positive. This finding suggests that there is a statistically significant and positive relationship between current market turnover and lagged market returns. Moreover, the results of the Granger causality tests indicate that there is a unidirectional Granger causality from market return to market turnover for both ETF markets. Additionally, the impulse response functions associated with the respective VAR models illustrate that the response of market turnover to shocks in market returns is statistically significant and positive in both markets. Therefore, the results from the estimated VAR models, Granger causality tests, and impulse response functions provide evidence in support of the overconfidence hypothesis and suggests that, in the South African ETF market, increased market gains increase investors' confidence in their trading skills, subsequently, leading to more frequent trades by investors. The next section provides evidence of the influence of the overconfidence bias on individual ETFs.

5.4 Empirical Results for the Influence of Investor Overconfidence on Individual ETFs

For research question two, the long-run relationship between the trading activity (turnover) of individual ETFs and market return is examined using panel VAR models, Granger causality tests and impulse response functions, and the results are discussed in the ensuing sections.

5.4.1. Individual Security Panel Vector Autoregression (VAR) Models

Section 5.3 documents the presence of investor overconfidence in the market of South African ETFs with domestic benchmarks and in the market of South African ETFs with international benchmarks because historical market returns positively impact current market trading activity in both markets. However, this finding may also be attributed to the manifestation of the disposition effect. As mentioned in Section 2.3.4, a key difference between the two behavioural biases is that, whilst the overconfidence bias relates to investors' attitudes about the general market, the disposition bias relates to investors' attitudes about individual securities in their portfolios (Statman, *et al.*, 2006). Therefore, to disentangle the two biases, the trading activity of individual ETFs is inspected. The lead-lag relationship between the current trading activity (turnover) of individual ETFs and lagged market returns is examined by estimating an individual security VAR model with a panel approach for each ETF market. The panel VAR models are estimated with the optimal lag lengths selected for the respective market VAR models in order to ensure consistency in the comparison of individual ETFs.

The panel VAR model specified in equation 4.33 is estimated for the market of JSE-listed ETFs tracking with local benchmarks, and the results for the regression with security turnover as the dependent variable are presented in Table 5.24 (page 115) while Appendix 6 (page 151) provides the results of the regressions with security return and market return as the dependent variables. Additionally, Table 5.25 (page 116) presents the results for the regression with security turnover as the dependent variable in the panel VAR model estimated for the market of JSE-listed ETFs with international benchmarks while Appendix 7 (page 152) provides the results of the regressions with security return and market return as the dependent variables. In Table 5.24, Table 5.25, Appendix 6 and Appendix 7, the dependent variables (*turn*, *ret* and *mret*) are organised in rows and the lagged dependent variable and exogenous variable coefficients appear in columns. For each coefficient, the t-statistic and probability value are reported.

The results presented in Tables 5.24 and 5.25 suggest that, for both markets, the turnover of individual ETFs is autocorrelated because lagged *turn* coefficients are significant explanatory variables when current *turn* is the dependent variable. For the market of ETFs tracking domestic benchmarks, lagged *turn* coefficients are generally declining in magnitude from the second and higher lags. Whilst all six *turn* lags have positive coefficients, only lags one to five are statistically significant at a 1% level of significance. Similarly, for the market of ETFs tracking international benchmarks, all the lagged *turn* coefficients have positive signs and are statistically significant. Similar findings are reported by Zia, *et al.* (2017) and Gupta, *et al.* (2018) who report that the turnover of individual securities is positively autocorrelated. This

significant, positive relationship between the turnover of individual ETFs and its historical turnover can be attributed to the manner in which information about individual ETFs is disseminated (He and Wang, 1995; Chung, Li and McInish, 2005) and could indicate the presence of positive feedback trading (Covrig and Ng, 2004).

Table 5.24: Panel VAR Model Estimates for ETFs Tracking Domestic Benchmarks

	Turn_dom (-1)	Turn_dom (-2)	Turn_dom (-3)	Turn_dom (-4)	Turn_dom (-5)	Turn_dom (-6)
Turn_dom	0.1044*** (7.0554)	0.0858*** (5.7894)	0.0853*** (5.7583)	0.0570*** (3.8133)	0.0541*** (3.7641)	0.0036 (0.9548)
	Ret_dom (-1)	Ret_dom (-2)	Ret_dom (-3)	Ret_dom (-4)	Ret_dom (-5)	Ret_dom (-6)
Turn_dom	0.0064 (0.5298)	0.0098 (0.7431)	-0.0025 (-0.2044)	-0.0089 (-0.9526)	-0.0076 (-0.9339)	-0.0041 (-0.5677)
	Mret_dom (-1)	Mret_dom (-2)	Mret_dom (-3)	Mret_dom (-4)	Mret_dom (-5)	Mret_dom (-6)
Turn_dom	0.0090 (0.2299)	0.0533 (1.3558)	0.0522 (1.3521)	0.0533 (1.4233)	0.0244 (0.6547)	-0.0051 (-0.1385)
	C	Sig_dom	Sig_dom (-1)	Sig_dom (-2)	Sig_dom (-3)	Sig_dom (-4)
Turn_dom	0.0425*** (12.4271)	0.0009 (0.1223)	0.0046 (0.5666)	0.0003 (0.0383)	-0.0098 (-1.3036)	-0.0005 (-0.0622)

Notes:

1. Values in brackets ‘()’ represents t-statistics.
2. ***, **, * indicate statistical significance at 1%, 5%, and 10% levels, respectively.

Table 5.24 indicates that, for the market of ETFs tracking domestic benchmarks, the turnover (*turn*) of individual ETFs do not exhibit statistically significant relationships with the historical returns of individual ETFs (*ret*), lagged market returns (*mret*) and contemporaneous and historical security volatility (*sig*). Interestingly, the first and second lagged *ret* coefficients are positive while the remaining *ret* lagged coefficients are negative. Overall, the results in table 5.24 suggest that the disposition effect does not significantly influence the trading activities of individual ETFs in the market of South African ETFs with domestic benchmarks since there is no significant relationship between current *turn* and lagged *ret*. Nevertheless, this finding is consistent with Statman, *et al.* (2006) who report statistically insignificant lagged *ret* coefficients in the regression with *turn* as the dependent variable.

Similarly, Table 5.24 suggests that the turnover of individual ETFs do not exhibit significant relationships with the lagged market return because none of the *mret* lagged coefficients are statistically significant. With the exception of the sixth *mret* lag coefficient, the remaining *mret* lags exhibit positive coefficients. Since the *mret* coefficients are statistically insignificant, the results obtained from estimated panel VAR models indicate that, after controlling for the disposition effect and the volume-volatility relationship, the overconfidence bias does not significantly influence the trading activities of individual ETFs in the market of JSE-listed ETFs with domestic benchmarks. These findings are consistent with Gupta, et al. (2018) who find that the overconfidence of Indian investors increases the turnover of the market but not individual securities. However, given the limitations of VAR models, the influence of *ret* and *mret* on *turn* are further examined and discussed under the analysis of the impulse response functions in Section 5.4.3.

Table 5.25: Panel VAR Model Estimates for ETFs Tracking International Benchmarks

	Turn_int (-1)	Turn_int (-2)	Turn_int (-3)	Ret_int (-1)	Ret_int (-2)	Ret_int (-3)
Turn_int	0.2132*** (6.3811)	0.0624* (1.8207)	0.2219*** (6.6371)	0.0652 (1.1800)	0.0804 (1.4127)	-0.1444*** (-2.6172)
	Mret_int (-1)	Mret_int (-2)	Mret_int (-3)	C	Sig_int (-1)	Sig_int (-1)
Turn_int	0.0066 (0.1073)	-0.0114 (-0.1799)	0.2891*** (4.6488)	0.0096 (1.6056)	0.3389*** (5.1893)	-0.0420 (-0.6745)
	Sig_int (-2)	Sig_int (-3)	Sig_int (-4)	Sig_int (-5)	Sig_int (-6)	Sig_int (-7)
Turn_int	-0.0574 (-0.9052)	-0.1797*** (-2.8253)	0.04804 (0.7766)	0.0957 (1.5721)	0.0296 (0.4920)	0.0065 (0.1102)

Notes:

1. Values in brackets ‘()’ represents t-statistics.
2. ***, **, * indicate statistical significance at 1%, 5%, and 10% levels, respectively.

On the contrary, the results of the estimated panel VAR model in Table 5.25 suggest that, after controlling for the disposition effect and the volume-volatility relationship, the overconfidence bias significantly influences the trading activities of individual ETFs in the market of JSE-listed ETFs with international benchmarks. This is because, in the market of ETFs tracking international benchmarks, the current turnover of individual ETFs exhibits a significant positive relationship with lagged market returns. Specifically, the third lagged *mret* coefficient is 0.2891 and significant at a 1% level of significance when *turn* is the dependent variable. Notably, there is also a significant relationship between current *turn* and the third

lagged *ret* coefficient. Particularly, the third lagged *ret* coefficient is -0.1444 and statistically significant at a 1% level of significance. However, the significant negative relationship between current *turn* and lagged *ret* is not consistent with the disposition effect since current *turn* and lagged *ret* exhibits a significant positive relationship in the presence of the disposition effect. This influence of *ret* and *mret* on *turn* is further discussed under the analysis of the impulse response functions in Section 5.4.3. Consistent with Statman, *et al.* (2006), the results of the estimated panel VAR model in Table 5.25 suggest that, not only is the influence of historical market returns on the trading activity of individual ETFs present after including security returns in the same regression, but the impact of lagged market returns is greater than that of lagged security returns (compare 0.2891 to -0.1444).

The turnover of individual ETFs in the market of ETFs tracking international benchmarks is also significantly influenced by the volatility of individual ETFs. Specifically, when *turn* is the dependent variable, current *sig* is 0.3389 and significant at a 1% level of significance. Additionally, the third lagged *sig* coefficient is -0.1797 and significant at a 1% level of significance while the remaining *sig* lags are statistically insignificant. Consistent with Statman *et al.* (2006), this study reports a sign-changing relationship between the turnover of individual ETFs and contemporaneous and historical security volatility. As discussed in Section 5.3.2, the volume-volatility relationship can either be positive or negative depending on how information about individual ETFs are disseminated as well as the extent to which ETF prices convey information (Karpoff, 1987).

Noticeably, Appendix 6 and Appendix 7 shows that the returns of individual ETFs (*ret*) in both markets as well as the returns of both markets (*mret*) are inefficient since these returns can be significantly explained by either lagged security turnover, lagged security return, lagged market return, or lagged security volatility. The critical results in Appendix 6 reveal that the returns of individual ETFs in the market of ETFs tracking domestic benchmarks exhibit significant autocorrelation since all the lagged security return coefficients (except the third lagged *ret* coefficient) are statistically significant. Specifically, the relationship between *ret* and the first, second, fourth, and fifth lagged *ret* coefficient is negative and statistically significant at a 1% level of significance. Likewise, for the market of ETFs tracking international benchmarks, Appendix 7 shows that, when *ret* is the dependent variable, the first lagged *ret* coefficient is -0.2951 and statistically significant at a 1% level of significance. This negative autocorrelation in the returns of individual ETFs in both markets could be attributed to the presence of positive feedback trading (Sentana and Wadhvani, 1992). Positive feedback trading by South African ETF investors is also documented by Charteris, *et al.* (2014). Overall, this positive feedback trading by ETF investors could induce overconfident trading because investors may overestimate their ability to predict price trends if this trajectory continues.

5.4.2. Granger Causality Tests

The causal relationship between the turnover of individual ETFs and market returns is inspected using the Granger (1969) causality test discussed under Section 4.3.1.2. Results for individual ETFs in the market of ETFs tracking domestic benchmarks are provided in Table 5.26 while Table 5.27 (page 119) contains the results for individual ETFs in the market of ETFs tracking international benchmarks.

Table 5.26: Granger Causality Test Results for Individual ETFs in the Market of ETFs Tracking Domestic Benchmarks

Dependent variable: Turn_dom			
Excluded	Chi-sq	df	Prob.
Ret_dom	2.6673	6	0.8493
Mret_dom	10.1179	6	0.1198
All	17.6534	12	0.1266
Dependent variable: Ret_dom			
Excluded	Chi-sq	df	Prob.
Turn_dom	3.1173	6	0.7940
Mret_dom	532.5223	6	0.0000***
All	543.1496	12	0.0000***
Dependent variable: Mret_dom			
Excluded	Chi-sq	df	Prob.
Turn_dom	9.2400	6	0.1605
Ret_dom	13.9678	6	0.0300**
All	23.2003	12	0.0261**

Notes:

1. Chi-sq represents the chi-square value, whilst *prob.* represents its associated probability value.
2. ***, **, * indicate statistical significance at 1%, 5%, and 10% levels, respectively.

The results contained in Table 5.26 show that, for the market of ETFs tracking domestic benchmarks, when security turnover is the dependent variable, the chi-square statistic associated with market return has a probability value that is greater than 0.10. Thus, the null hypothesis that market return does not cause security turnover cannot be rejected at a 10% level of significance, therefore, suggesting that lags of market return do not Granger-cause the turnover of individual ETFs in the market of ETFs tracking domestic benchmarks. On the contrary, Table 5.27 illustrates that, for the market of ETFs tracking international benchmarks, the probability value associated with the chi-square statistic of market return is less than 0.01 when security turnover is the dependent variable. Therefore, the null hypothesis that lags of market return

does not cause security turnover is rejected at a 1% level of significance. This implies that, with a 99% confidence interval, lags of market return does Granger cause the turnover of individual ETFs in the market of ETFs tracking international benchmarks. Noticeably, for the market of ETFs tracking international benchmarks, lags individual security return (*ret*) also Granger causes the turnover of individual ETFs in the market since its chi-square statistic of 11.8848 is statistically significant at a 1% level of significance. However, this causal relationship does hold in the market of ETFs tracking domestic benchmarks since the chi-square statistic of 2.6673 is statistically insignificant.

Table 5.27: Granger Causality Test Results for Individual ETFs in the Market of ETFs Tracking International Benchmarks

Dependent variable: Turn_int			
Excluded	Chi-sq	df	Prob.
Ret_int	11.8848	3	0.0078***
Mret_int	23.0445	3	0.0000***
All	36.8933	6	0.0000***
Dependent variable: Ret_int			
Excluded	Chi-sq	df	Prob.
Turn_int	7.1801	3	0.0664*
Mret_int	11.5227	3	0.0092***
All	18.7538	6	0.0046***
Dependent variable: Mret_int			
Excluded	Chi-sq	df	Prob.
Turn_int	3.6889	3	0.2971
Ret_int	3.1823	3	0.3644
All	6.4190	6	0.3779

Notes:

1. Chi-sq represents the chi-square value, whilst *prob.* represents its associated probability value.
2. ***, **, * indicate statistical significance at 1%, 5%, and 10% levels, respectively.

Noticeably, Table 5.26 indicates that there is also a significant bidirectional Granger causality between market return and security return implying that the return of individual ETFs in the market of ETFs tracking domestic benchmarks Granger causes its respective market return and market return Granger causes the returns of individual ETFs in the market. However, for the market of ETFs tracking international benchmarks, market return Granger causes the returns of individual ETFs in the market, but the return of individual ETFs does not significantly Granger cause the return of the market. Interestingly, the turnover of individual ETFs Granger causes the returns of individual ETFs since the chi-square statistic of 7.1801 is

significant at a 10% level of significance. Overall, the Granger causality tests are used to ascertain the existence of causal relationships between the endogenous variables and do not provide evidence regarding the overconfidence hypothesis and disposition effect since Granger causality tests do not provide information about the sign of the causality (Liddle and Lung, 2013). Hence, impulse response functions are analysed in the next section to examine the sign of the causal relationship.

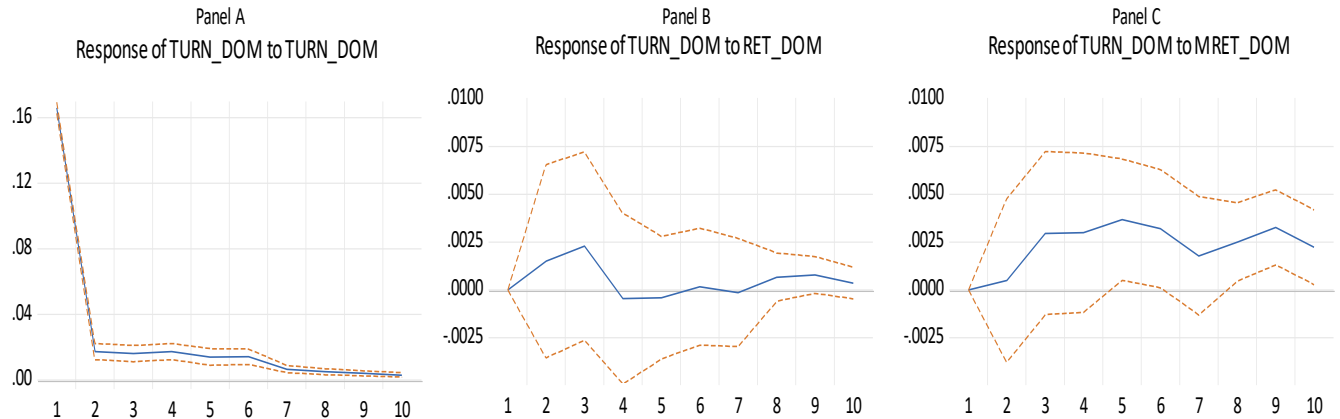
5.4.3. Impulse Response Functions

An important component of this analysis is to examine the sign and significance of the causal relationship between the turnover of individual ETFs and its security and market return. As such, impulse response functions are employed to trace how security turnover responds to shocks in security and market return. The response of individual security turnover, *turn*, to shocks in each of the endogenous variables associated with the panel VAR models estimated for the market of ETFs tracking domestic and international benchmarks are illustrated in Figure 5.6 (page 121) and Figure 5.7 (page 121), respectively. Additionally, the responses of individual security return (*ret*) and market return (*mret*) to shocks in each of the endogenous variables are illustrated for the market of ETFs tracking domestic and international benchmarks in Appendix 8 (page 153) and 9 (page 153), respectively. In Figure 5.6 and 5.7, panel A shows the response of *turn* to a one standard deviation shock in *turn*, panel B plots the response of *turn* to a one standard deviation shock in *ret* and panel C illustrates the response of *turn* to a one standard deviation shock in *mret*. It is important to note that, because the *turn* variable did not undergo a log transformation, the vertical axis in panels A, B and C measures the unit change in *turn* (that is, the security turnover ratio) as opposed to a percentage measurement. The values associated with each impulse response in Figure 5.6 are provided in the impulse response table in Appendix 10 (page 154) and the values associated with each impulse response in Figure 5.7 are provided in the impulse response table in Appendix 11 (page 154). Notably, the impulse response is statistically significant when both standard error bands are below or above zero on the y-axis.

Panel A in Figures 5.6 and 5.7 confirm the serial dependence of individual security turnover since the *turn* in both markets exhibit significant, positive responses to its own shocks. For both markets, the highest response of *turn* to a one standard deviation shock in *turn* is evident in the first period. As with market turnover, the response of individual security turnover to its own shocks follows a decreasing trend, although, the impact of these shocks remains positive throughout the ten month period. Moreover, like market turnover, autocorrelation in individual security turnover could indicate the presence of positive feedback trading in individual ETFs -as discussed in Section 5.4.1.

Figure 5.6: Impulse Response Functions for ETFs Tracking Domestic Benchmarks

Response to Cholesky One S.D. (d.f. adjusted) Innovations ± 2 S.E.

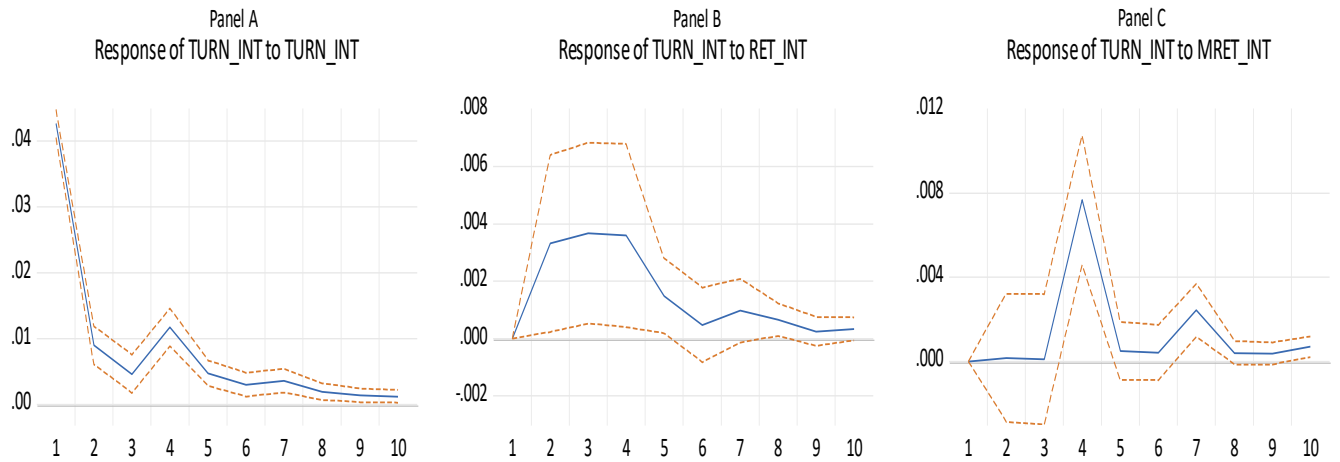


Notes:

1. The blue line represents the impulse response and the red lines represent two standard error bands.

Figure 5.7: Impulse Response Functions for ETFs Tracking International Benchmarks

Response to Cholesky One S.D. (d.f. adjusted) Innovations ± 2 S.E.



Notes:

1. The blue line represents the impulse response and the red lines represent two standard error bands.

The response of individual security turnover to a one standard deviation shock in individual security return is illustrated in panel B of both figures. For both markets, a one standard deviation *ret* shock does not generate any significant impulse response in the *turn* of the next month. For the second period, *turn* responds positively to a one standard deviation shock in *ret*, however, this response is only statistically significant for individual ETFs in the market of ETFs tracking international benchmarks. Noticeably, throughout the ten month period, the response of individual security turnover to individual security return

is statistically insignificant for individual ETFs in the market of ETFs tracking domestic benchmarks. As such, there is no significant evidence of the disposition effect in the market of South African ETFs tracking domestic benchmarks since the response of individual security turnover to individual security return is statistically insignificant.

On the contrary, the response of individual security turnover to individual security return is statistically significant and positive for individual ETFs in the market of ETFs tracking international benchmarks during periods 2, 3, 4, 5 and 8. This significant, positive response of individual security turnover to shocks in individual security return is consistent with the predictions of the disposition effect. In other words, the significant, positive dependence of individual security turnover on individual security return suggests that investment decisions made by investors trading in the market of JSE-listed ETFs with international benchmarks are influenced by the disposition effect. This finding indicates that positive returns increase security trading volume because investors enjoy realising paper gains on individual ETFs. Alternatively, negative returns may reduce trading activity because these investors hold on to individual ETFs since they do not want to realise losses. This finding is consistent with Choe and Eom (2009) who report strong evidence of the disposition effect in the Korean stock index futures market and Prosad, *et al.* (2017) who document evidence of both the overconfidence and disposition effect in the Indian equity market.

Consistent with the results of market turnover (discussed in Section 5.3.4), panel C in Figures 5.6 and 5.7 illustrate that, for both markets, a one standard deviation shock in market return does not generate any impulse response in the next month's security turnover. Noticeably, from the second period onwards, the response of *turn* to *mret* is positive in both markets. However, in the market of ETFs tracking domestic benchmarks, this positive response of *turn* to shocks in *mret* is only statistically significant during periods 5, 6, 8, 9 and 10. Thus, whilst the results of the panel VAR models (presented in Section 5.4.1) report no statistically significant relationship between current security turnover and lagged market returns, the impulse response function in panel C of Figure 5.6 illustrates that the response of individual security turnover to shocks in market returns is positive and statistically significant during periods 5, 6, 8, 9 and 10. Similarly, for the market of ETFs tracking international benchmarks, Figure 5.7 shows that the response of *turn* to shocks in *mret* is positive but statistically significant only in periods 4, 7 and 10. For both markets, the statistically significant, positive response of individual security turnover to shocks in market return represents the second key empirical finding of this study. Specifically, the significant, positive dependency of individual security turnover on market return provides evidence of the manifestation of investor overconfidence at the individual security level. This finding suggests that, for both markets, increased

market performance boosts investors' confidence in their trading and security valuation skills, causing investors to trade more frequently in individual ETFs.

In summary, the results of the impulse response analysis suggest that the turnover and, thus, trading activity of individual ETFs in the market of ETFs tracking international benchmarks depend on their respective historical security return. The finding that historical security returns impact the trading activity of individual ETFs is attributed to the manifestation of the disposition effect²⁶ at individual ETF level. However, the second key empirical finding of this study is represented by the significant, positive impact of lagged market returns on the current turnover of individual ETFs in both markets. This finding implies that investor overconfidence induced by market return positively impacts realised trading activity, even after controlling for the disposition effect. Overall, these findings suggest that the overconfidence effect and the disposition effect exist concurrently in the market of South African ETFs tracking international benchmarks since the overconfidence effect accounts for a portion of ETF trading activity that is not captured by the disposition effect. Similarly, Statman, *et al.* (2006), Zaiane (2013), Prosad, *et al.* (2017) and Gupta, *et al.* (2018) report that the overconfidence bias and the disposition effect manifest concurrently in the observed markets. However, in the market of South African ETFs tracking domestic benchmarks, the impulse response analysis suggests that only the overconfidence bias influences the trading activity of individual ETFs. Therefore, for both markets, the significant, positive response of current security turnover to historical market returns, even after accounting for security return, confirms that the market-wide overconfidence effect observed in Section 5.3 is not a direct summation of the disposition effect.

For completion, Appendix 8 and Appendix 9 confirm that the individual security returns and market returns of both markets are autocorrelated since they respond significantly to their own shocks. Moreover, for both markets, the bidirectional causal relationship between security returns and market returns is evident since *ret* responds significantly to shocks in *mret* and *mret* responds significantly to shocks in *ret*. On the contrary, for both markets, *ret* and *mret* respond insignificantly to shocks in *turn*.

In summary, the results obtained for research question two indicate that the overconfidence bias significantly influences the trading activities of individual ETFs in the market of South African ETFs tracking domestic benchmarks and in the market of South African ETFs tracking international benchmarks. Specifically, for the market of ETFs tracking international benchmarks, the results from the panel VAR models indicate that there is a significant positive relationship between current individual security turnover

²⁶ The disposition effect refers to the tendency of investors to hold onto losing positions whilst selling winning positions (Shefrin and Statman, 1985).

and lagged market returns. Additionally, the Granger causality tests confirm that there is a causal relationship from market return to security turnover and the impulse response functions illustrate that security turnover responds significantly and positively to shocks in market return. For the market of ETFs tracking domestic benchmarks, the results from the estimated panel VAR model suggest that there is no significant relationship between current security turnover and lagged market returns, and as such, the Granger causality test confirms that there is no causal relationship from market return to security turnover. However, the impulse response functions show that the response of security turnover to market return is positive and statistically significant. Given that VAR models do not fully capture the effect of other endogenous and exogenous variables, the results of the impulse response functions are more reliable since these functions fully capture the effects of other variables in the system (Swanson and Granger, 1997; Dewachter, *et al.*, 2015). Hence, these findings confirm that the market-wide investor overconfidence reported for research question one is not a direct aggregation of the disposition effect at the individual security level. However, it is important to note that the evidence on the presence of overconfidence in individual ETFs tracking domestic benchmarks is not strong since only the impulse response functions confirm the presence of overconfident trading whilst the VAR and Granger causality results do not confirm the presence of overconfident trading by these investors.

Given that the aggregate results for research questions one and two (discussed in Section 5.3 and Section 5.4, respectively) confirm the presence of investor overconfidence in the market of JSE-listed ETFs with South African benchmarks as well as the market of JSE-listed ETFs with non-South African benchmarks, the next section examines the effect on investor overconfidence on the volatility of the respective South African ETF market.

5.5. Empirical Results of the Effect of Investor Overconfidence on Market Volatility

The previous sections (that is, Sections 5.3 and 5.4) provide evidence of the presence of investor overconfidence at the market level and individual security level in the market of South African ETFs tracking domestic benchmarks and in the market of South African ETFs tracking international benchmarks. Hence, this section aims to examine the effect of trading volume induced by overconfident trading on the volatility of the respective market. For research question three, the effect of overconfidence on market volatility is examined using EGARCH (1,1) models, and the results are presented in the ensuing discussion.

Following Chuang and Lee (2006), the EGARCH (1,1) model is used to examine the effect of overconfident trading on the volatility of the returns from each ETF market. However, prior to the estimation of the EGARCH (1,1) model, the trading volume of the respective ETF market is decomposed into an

overconfident trading component and a non-overconfident trading component as specified in equation 4.34. The optimal lag length, J , of lagged market returns to be included in equation 4.34 is selected using the AIC, SC, and HQ information criteria. Table 5.28 presents the information criteria at each lag of market return for the market of ETFs tracking domestic benchmarks and Table 5.29 (page 126) contains these statistics for the market of ETFs tracking international benchmarks.

The information criteria statistics in Table 5.28 demonstrate that, for the market of ETFs tracking domestic benchmarks, when market turnover is the dependent variable, the AIC is minimised at 7 lags of the market return while the SC and HQ are minimised at 1 lag of the market return. Regarding the market of ETFs tracking international benchmarks, Table 5.29 shows that the AIC is minimised at 3 lags of market return, whilst the SC and HQ are minimised at 1 and 2 lags of the market return, respectively. Overall, the information criteria statistics presented in Tables 5.28 and 5.29 show that the information criteria (that is, the AIC, SC, and HQ) provide inconsistent results. Therefore, given that the AIC offers better insight into the features of the data, the optimal lag length of market returns is selected based on the AIC. Based on the AIC, equation 4.34 is estimated for the market of ETFs tracking domestic benchmarks using 7 lags of market return (refer to Appendix 12, page 155) and 3 lags of the market return for the market of ETFs tracking international benchmarks (refer to Appendix 13, page 156).

Table 5.28: Lag Order Selection Criteria for Market Return in the Market of ETFs Tracking Domestic Benchmarks

Lags	AIC	SC	HQ
1	3.857942	3.860596*	3.858873*
2	3.858310	3.862293	3.859707
3	3.858558	3.863869	3.860421
4	3.858987	3.865627	3.861317
5	3.859458	3.867428	3.862255
6	3.857566	3.866865	3.860829
7	3.857025*	3.867654	3.860754
8	3.857453	3.869413	3.861650

Notes:

1. * indicates lag order selected by the criterion
2. AIC refers to the Akaike information criterion, SC refers to the Schwarz information criterion and HQ refers to the Hannan-Quinn information criterion.

Table 5.29: Lag Order Selection Criteria for Market Return in the Market of ETFs Tracking International Benchmarks

Lags	AIC	SC	HQ
1	2.950003	2.953551*	2.951270
2	2.949017	2.954340	2.950918*
3	2.948446*	2.955544	2.950980
4	2.948572	2.957447	2.951741
5	2.949160	2.959813	2.952964
6	2.949406	2.961837	2.953845
7	2.949873	2.964084	2.954947
8	2.949204	2.965195	2.954914

Notes:

1. * indicates lag order selected by the criterion
2. AIC refers to the Akaike information criterion, SC refers to the Schwarz information criterion and HQ refers to the Hannan-Quinn information criterion.

Thereafter, as specified in equation 4.34, the constant and residuals of the estimated regressions are used to compute the series representing the non-overconfident trading component ($NONOVER_{\tau}$) as well as the series representing the overconfident trading component ($OVER_{\tau}$). However, prior to the estimation of the respective EGARCH (1,1) models, the overconfident trading and non-overconfident trading series are tested for stationary, and the results of the PP unit root test are presented in Table 5.30. In addition, Table 5.30 provides the PP unit root test results for the market return of each market.

Table 5.30: PP Unit Test Results

Variable	PP test stat.	Prob.
Over_dom	-63.031	0.0001
Nonover_dom	-83.701	0.0001
R_dom	-70.547	0.0001
Over_int	-56.380	0.0001
Nonover_int	-69.131	0.0001
R_int	-63.984	0.0001

Notes:

1. PP test stat. and *prob.* refers to the PP unit root test statistic and probability value, respectively.

The results of the PP unit root test in Table 5.30 indicate that, for each series, the probability values associated with each PP test statistic is less than 0.01. As a result, the null hypothesis of a unit root in the

series is rejected at a 1% level of significance for all series. Therefore, it is concluded that, for the market of ETFs tracking domestic benchmarks and the market of ETFs tracking international benchmarks, the daily series representing the overconfident trading component, the non-overconfident trading component and the market return are stationary at levels. Thus, all variables used in the estimation of the EGARCH models are stationary at levels.

In addition to the above, Table 5.31 shows that the correlation between $OVER_{\tau}$, $NONOVER_{\tau}$ and R_{τ} is low for the market of ETFs tracking domestic benchmarks. Except for the moderate correlation (-0.4757) between $OVER_{\tau}$ and $NONOVER_{\tau}$, the remaining correlations are low for the market of ETFs tracking international benchmarks. Overall, the variables used to estimate the respective EGARCH models do not exhibit high correlations, and thus, multicollinearity will not be a problem.

Table 5.31: Correlation Coefficients Computed Between the Respective EGARCH Variables

	R_dom	Over_dom	Nonover_dom
R_dom	1.0000		
Over_dom	0.0024	1.0000	
Nonover_dom	-0.0328	0.0001	1.0000
	R_int	Over_int	Nonover_int
R_int	1.0000		
Over_int	-0.0051	1.0000	
Nonover_int	0.0174	-0.4757	1.0000

The stationary market return, overconfident trading and non-overconfident trading series are used to estimate the respective EGARCH (1,1) model which follows the mean equation defined in equation 4.35 and the conditional variance equation specified in equation 4.37. For both markets, Table 5.32 (page 128) contains the results of the estimated EGARCH (1,1) models.

With regards to the mean equations, the results presented in Table 5.32 suggest that, for both markets, the average market return (μ) based on past information exhibits a positive relationship with the current market returns. However, this relationship is only statistically significant for the market of ETFs tracking international benchmarks. This finding suggests that the current daily returns of the market of ETFs with international benchmarks can be explained by its average daily returns.

Table 5.32: Results of the Estimated EGARCH Models

Estimated Mean Equation: $R_t = \mu_t + \eta_t$ Estimated Conditional Variance Equation: $\ln(\sigma^2) = \omega + f_1 \frac{ \eta_{t-1} }{\sqrt{\sigma_{t-1}^2}} + f_2 \frac{\eta_{t-1}}{\sqrt{\sigma_{t-1}^2}} + f_3 \ln(\sigma_{t-1}^2) + f_4 OVER_t + f_5 NONOVER_t$		
Variable	Market of ETFs with Domestic Benchmarks	Market of ETFs with International Benchmarks
μ	0.1369 (1.3866)	0.9575*** (3.4596)
ω	-10.6081*** (-416.022)	-0.4525*** (-6.4488)
f_1	-0.0449*** (-3.6714)	0.1675*** (12.8001)
f_2	-0.0031 (-0.3723)	-0.0173** (-2.2177)
f_3	-0.0678*** (-17.9346)	0.9660*** (167.214)
f_4	1.6044*** (15.3167)	0.0098** (2.4372)
f_5	-0.4264*** (-140.047)	-0.0020 (-0.4128)

Notes:

1. Values in brackets ‘()’ represent z-statistics.
2. ***, **, * indicate statistical significance at 1%, 5%, and 10% levels, respectively.

Regarding the conditional variance equations, Table 5.32 suggests that the constant term (ω) is a significant explanatory variable in the conditional variance equations of both markets since the constant terms are statistically significant at a 1% level of significance. In addition, the ARCH (f_1) and GARCH (f_3) terms are significant explanatory variables in the conditional variance equations. For both markets, f_1 is statistically significant at a 1% level of significance, suggesting that innovative residual shocks influence the current conditional volatility of the respective ETF market. From Table 5.32, it is also evident that the current conditional volatility of both markets is affected by its historical conditional volatility since f_3 is statistically significant at a 1% level of significance in both markets. Overall, the results contained in Table 5.32 suggest that the current conditional volatility of the market of JSE-listed ETFs with local benchmarks and the market of JSE-listed ETFs with international benchmarks respond significantly to innovative residual shocks and past conditional volatility.

In addition to the above, Table 5.32 illustrates that, for the market of ETFs tracking domestic benchmarks, the sum of f_1 and f_3 is less than one, implying that the conditional variance follows a stationary process. On the contrary, for the market of ETFs tracking international benchmarks, the sum of f_1 and f_3 is greater than one, suggesting an integrated process in which volatility is persistent and possibly explosive (Paul and Theodore, 2006). In other words, the conditional variance is non-stationary and shocks in volatility do not decay over time. Interestingly, Chuang and Lee (2006), Abbas (2013) and Jlassi, *et al.* (2014) also report ARCH and GARCH terms that sum to greater than one. Karanasos, Paraskevopoulos, Ali, Karoglou and Yfanti (2014) argue that the persistence of volatility in market returns may be due to the high levels of market volatility experienced during the 2008 financial crisis. This argument put forward by Karanasos, *et al.* (2014) is further supported by the results presented in Section 5.6 (refer to Table 5.34, page 132) in which the conditional variance of the market of ETFs tracking international benchmarks is only explosive during the 2008 global financial crisis and stationary before and after the crisis.

With regards to the asymmetry term (f_2), f_2 in the conditional variance regression estimated for the market of ETFs tracking international benchmarks is negative and statistically significant at a 5% level of significance. The significant, negative asymmetry term indicates that leverage effects are present. Hence, negative return shocks have a greater impact on volatility in comparison to positive return shocks of the same magnitude. Notably, the asymmetry term is also negative in the conditional variance regression estimated for the market of ETFs tracking domestic benchmarks, however, this coefficient is statistically insignificant. This being so, significant asymmetric volatility effects do not exist in the market of JSE-listed ETFs tracking South African benchmarks.

In Table 5.32, f_4 signifies the effect of overconfident trading on volatility and f_5 represents the effect of non-overconfident trading on volatility. In the conditional variance equation of both markets, f_4 is positive and statistically significant at a 1% and 5% level of significance for the market of ETFs tracking domestic and international benchmarks, respectively. The positive relationship between trading volume induced by investor overconfidence ($OVER_t$) and conditional volatility represents the third key empirical finding of this study. Specifically, this finding indicates that, for both markets, an increase (decrease) in overconfident trading leads to an increase (decrease) in the conditional volatility of the respective market return. Overall, this positive relationship between overconfident trading and market volatility suggests that high ETF market volatility may partially be due to the presence of investor overconfidence in the respective South African ETF market.

The positive relationship between overconfidence and volatility is also documented in studies conducted by Chuang and Lee (2006), Abbas (2013), and Jlassi, *et al.* (2014). Theoretical explanations for the positive association between overconfident trading and market volatility are provided by Odean (1998) and Scheinkman and Xiong (2003) as discussed in Section 2.3.5. In summary, overconfident traders generate excess market volatility by driving asset prices away from the intrinsic values since overconfident investors ignore market signals and trade excessively. Moreover, overconfident investors overestimate asset prices, subsequently, contributing to excess price volatility since these price deviations are not warranted by fundamental information. This being so, the overconfidence present in the South African ETF market may increase the volatility of the ETF market, which may contribute to the passive investment as predicted by Michael Burry. This is because, Jlassi, *et al.* (2014) note that overconfidence increases security mispricing, subsequently, fuelling the formation of price bubbles. Collectively, Daniel, *et al.* (1998), Scheinkman and Xiong (2003) and Jlassi, *et al.* (2014) support that notion that the presence of investor overconfidence generates excess market volatility which may fuel asset price bubbles. Hence, Section 5.6 examines whether investor overconfidence in the South African ETF market could have contributed to the excess market volatility which fuelled the 2008 global financial crisis.

For completeness, the results presented in Table 5.32 demonstrate that trading volume that is unrelated to investor overconfidence ($NONOVER_{\tau}$) exhibits a negative effect on volatility since f_5 is negative in both market's conditional variance equations. Noticeably, Abbas (2013) finds that $NONOVER_{\tau}$ has a statistically significant negative effect on market volatility for the United States, United Kingdom and Korea stock market indices. The non-overconfidence trading component is attributed to trading caused by other psychological biases, such as herd behaviour (Jlassi, *et al.*, 2014). Notably, Chen, *et al.* (2011) and Bahadar, *et al.* (2019) document the presence of herd behaviour by ETF investors. BenSaïda (2017) argues that, in the presence of herd behaviour, market volatility decreases as the average trading volume increases because informed traders engage in higher trading, thus, prices become more informed, and as a result, there is less volatility in the market. However, in this study, the negative relationship between non-overconfident trading and volatility is insignificant for the market of ETFs tracking international benchmarks but significant for the market of ETFs with domestic benchmarks at a 1% significance level. Overall, the magnitude of f_4 is greater than f_5 in the conditional variance equations of both markets, indicating that trading volume induced by overconfidence exhibits a greater contribution to conditional volatility than the trading volume that is induced by other factors (except overconfidence).

In summary, for research question three, the results obtained from the estimated EGARCH (1,1) models indicate that, for both markets, trading volume induced by investor overconfidence exhibits a significant,

positive effect on the conditional volatility of market return. Moreover, the effect of trading volume induced by overconfidence on volatility is greater than the effect of trading volume induced by non-overconfidence factors. The next section provides a discussion of the sub-period analysis of this effect.

5.6. Empirical Results for the Sub-Period Analysis of the Effect of Overconfidence on Volatility

The preceding section investigated the effect of overconfident trading on the volatility of the returns from the respective ETF market. In an attempt to answer research question four, the current section examines the effect of overconfident trading on market volatility before, during, and after the 2008 global financial crisis. To ensure consistency in the $OVER_t$ and $NONOVER_t$ values on day t , the $OVER_t$ and $NONOVER_t$ series computed for the full sample period are employed in this section, however, the full sample series is divided into subsamples as specified in Table 4.1. The EGARCH (1,1) model is estimated using the pre-, during-, and post-crisis subsamples for the market of South African ETFs tracking domestic benchmarks and the market of South African ETFs tracking international benchmarks and the results are presented in Tables 5.33 and 5.34. Table 5.33 (page 132) contains the results of the EGARCH models estimated for the market of ETFs tracking domestic benchmarks over the different subsamples and Table 5.34 (page 133) presents these results for the market of ETFs tracking international benchmarks.

The results presented in Tables 5.33 and 5.34 suggest that, for both markets, the average daily return, μ , exhibits a positive relationship with the current market return since all μ coefficients are positive. However, for the market of ETFs tracking domestic benchmarks, this positive relationship between average market returns based on past information and current market returns is statistically insignificant before, during, and after the 2008 global financial crisis since the probability values associated with the μ coefficients are greater than 0.10. This insignificant positive relationship is consistent with the results obtained for the full sample period which is presented in Table 5.32. Likewise, during the crisis, there is a statistically insignificant relationship between the average returns of the market of ETFs tracking offshore benchmarks and its current market returns. On the contrary, Table 5.34 indicates that, for the market of ETFs tracking international benchmarks, there exists a significant (at a 1% level of significance) positive relationship between average market returns and current market returns before and after the 2008 crisis. This finding is consistent with the results obtained for the full sample period (refer to Table 5.32, page 128) and suggests that the average daily returns could be used to predict the current returns before and after the 2008 global financial crisis.

Table 5.33: Results of the EGARCH Models Estimated for the Market of ETFs Tracking Domestic Benchmarks During Different Subsamples

Estimated Mean Equation: $R_\tau = \mu_\tau + \eta_\tau$ Estimated Conditional Variance Equation: $\ln(\sigma^2) = \omega + f_1 \frac{ \eta_{\tau-1} }{\sqrt{\sigma_{\tau-1}^2}} + f_2 \frac{\eta_{\tau-1}}{\sqrt{\sigma_{\tau-1}^2}} + f_3 \ln(\sigma_{\tau-1}^2) + f_4 OVER_\tau + f_5 NONOVER_\tau$			
Variable	Pre-crisis	During-crisis	Post-crisis
μ	0.0482 (0.3458)	0.3164 (1.1423)	0.0149 (0.3605)
ω	-9.1744*** (-113.5161)	-8.8882*** (-26.5547)	-15.8258*** (-88.4422)
f_1	-0.4898*** (-18.6655)	-0.2873*** (-3.6416)	0.1906*** (9.0029)
f_2	0.3050*** (12.5498)	0.2218*** (4.1677)	-0.01700 (-1.2055)
f_3	-0.1840*** (-13.3150)	0.1088*** (2.8793)	-0.3757*** (-15.3660)
f_4	2.1355*** (14.4744)	1.5166** (2.4161)	-0.2417 (-0.2050)
f_5	-0.2351*** (-43.8272)	-0.4812*** (-18.0910)	-0.4519*** (-22.0501)

Notes:

1. Values in brackets ‘()’ represent z-statistics.
2. ***, **, * indicate statistical significance at 1%, 5%, and 10% levels, respectively.

Table 5.33 and Table 5.34 show that the constant term, ω , is a significant explanatory variable in the conditional variance equations of both markets before, during and after the 2008 global financial crisis. Specifically, for the market of ETFs with domestic benchmarks, ω is negative and significant at a 1% level of significance before, during and after the crisis, although, for the market of ETFs with international benchmarks, ω is negative and significant at a 1% level only before and after the crisis, however, during the crisis, ω is significant at a 10% level of significance. Moreover, for both markets, the ARCH term (f_1) is significant at a 1% level of significance before, during and after the crisis suggesting that innovative residual shocks significantly impact the conditional volatility of both markets before, during, and after the crisis. Additionally, Tables 5.33 indicates that the current conditional volatility of the market of ETFs tracking domestic benchmarks is significantly influenced by its past conditional volatility since the GARCH term (f_1) is statistically significant at a 1% level in all three subsamples. However, Table 5.34 shows that the previous conditional volatility significantly impacts the current volatility of the market of ETFs tracking

international benchmarks only during and after the crisis whilst there is a statistically insignificant relationship between past volatility and current volatility before the crisis.

Table 5.34: Results of the EGARCH Models Estimated for the Market of ETFs Tracking International Benchmarks During Different Subsamples

Estimated Mean Equation: $R_t = \mu_t + \eta_t$ Estimated Conditional Variance Equation: $\ln(\sigma^2) = \omega + f_1 \frac{ \eta_{t-1} }{\sqrt{\sigma_{t-1}^2}} + f_2 \frac{\eta_{t-1}}{\sqrt{\sigma_{t-1}^2}} + f_3 \ln(\sigma_{t-1}^2) + f_4 OVER_t + f_5 NONOVER_t$			
Variable	Pre-crisis	During-crisis	Post-crisis
μ	2.080*** (2.7717)	0.0463 (0.0564)	1.087*** (3.3972)
ω	-10.1709*** (-5.1243)	-0.1044* (-1.8016)	-4.4945*** (-6.8103)
f_1	0.3968*** (3.9671)	0.0748*** (2.7004)	0.3379*** (10.1389)
f_2	0.1861** (2.4481)	-0.0731*** (-4.0631)	-0.0151 (-0.6398)
f_3	0.0201 (0.1031)	0.9881*** (200.3889)	0.5447*** (8.3600)
f_4	0.0315 (0.8201)	0.0095 (1.1027)	0.0860*** (6.9979)
f_5	-0.1259 (-1.4937)	0.0085 (1.0861)	0.0032 (0.1189)

Notes:

1. Values in brackets ‘()’ represent z-statistics.
2. ***, **, * indicate statistical significance at 1%, 5%, and 10% levels, respectively.

The asymmetry term (f_2) is positive in the conditional variance equations of both markets before the 2008 global financial crisis, although, this finding is significant at a 1% and 5% level of significance in the market of ETFs tracking domestic and international benchmarks, respectively. The significant, positive asymmetry terms suggest that, before the 2008 global financial crisis, a positive return shock induced an increase in volatility more than a negative return shock of the same magnitude. This finding supports the notion that overconfident traders overreact to positive returns and good news. During the global financial crisis, f_2 is positive and significant for the market of ETFs with domestic benchmarks but negative and significant for the market of ETFs with international benchmarks. The significant, negative f_2 implies that, during the crisis, negative shocks had a greater impact on volatility relative to positive return shocks of equal magnitude, thereby, suggesting that investors trading in ETFs with offshore exposures placed more weight

on negative return shocks and bad news. Interestingly, the conditional variance equations estimated using the post-crisis subsamples exhibit statistically insignificant f_2 coefficients for both markets. Therefore, after the 2008 global financial crisis, significant asymmetric volatility effects are not present in the market of JSE-listed ETFs with South African benchmarks and the market of JSE-listed ETFs with non-South African benchmarks.

Regarding the effect of trading volume induced by overconfidence on volatility (f_4), the results presented in Tables 5.33 and 5.34 illustrate that, before the 2008 global financial crisis, f_4 is positive in both markets, yet, only significant for the market of ETFs with domestic benchmarks. Likewise, during the crisis, f_4 is positive in both markets but also only significant for the market of ETFs tracking domestic benchmarks. Overall, these results suggest that overconfident trading exhibits a positive effect on the conditional volatility of both markets before and during the 2008 global financial crisis, however, these influences are only significant for the market of ETFs tracking domestic benchmarks. The insignificant overconfidence effect for the market of ETFs tracking international benchmarks may be attributed to the unpopularity of these ETFs before the crisis since there were only two ETFs tracking international benchmarks before the 2008 global financial crisis. Notably, in both markets, the magnitude of the overconfidence effect is greater before the crisis. This finding supports the theoretical assumptions that overconfidence is more pronounced in tranquil periods in which overconfident investors trade aggressively because they tend to ignore systematic market signals. As a result, overconfident investors overestimate ETF prices, resulting in an abnormal increase in ETF market volatility, which may have increased the volatility that fuelled the 2008 global financial crisis.

In addition, during the 2008 global financial crisis, f_4 is positive in both markets but only significant for the market of ETFs tracking domestic benchmarks. The significant, positive effect of overconfident trading on volatility during the crisis is indicative of the presence of an illusion of control, as discussed in section 2.3.1.1. In other words, overconfident traders believe that they can control difficult situations (in this case, the crisis), and therefore, these investors continue to trade excessively even during the crisis. As a consequence, overconfident trading may have intensified the volatility experienced during the global financial crisis. These findings suggest that if there is an ETF market crash as predicted by market experts, overconfident trading by ETF investors could increase the volatility experienced during the crash. Consistently, Jlassi, *et al.* (2014) also report that an increase in overconfident trading leads to an increase in volatility before and during the 2008 crisis. However, Abbas (2013) finds that the overconfidence bias cannot explain volatility during the global financial crisis.

The conditional variance equations estimated for the post-crisis subsamples indicate that the overconfidence effect (f_4) is statistically insignificant for the market of ETFs tracking domestic benchmarks, yet, positive and significant at a 1% level of significance for the market of ETFs tracking international benchmarks. The insignificant overconfidence effect after the crisis for the market of ETFs tracking domestic benchmarks may be attributed to its relatively low returns (illustrated in Figure 5.2, page 85) leading to low levels of investor overconfidence post the global financial crisis. Instead, after the global financial crisis, the number and popularity of ETFs tracking international ETFs increased indicating that investors are moving towards ETFs with international exposures since these ETFs provide low-cost access to international markets that generally perform better than the domestic market. Consistent with Jlassi, *et al.* (2014), this finding indicates that overconfidence is a persistent psychological bias affecting the decisions made by investors trading in South African ETFs that track international benchmarks even after the global financial crisis.

It is important to note that, for the market of ETFs with domestic benchmarks, there are other possible factors (that is, other psychological biases like herd behaviour which was discussed in Section 5.5) that could explain its current conditional volatility since f_5 is significant and negative before, during, and after the crisis, thereby, indicating that trading relating to non-overconfidence factors negatively impacts its conditional volatility. However, for the market of ETFs tracking international benchmarks, non-overconfidence factors do not significantly influence its current conditional volatility since f_5 is statistically insignificant before, during, and after the crisis. Nevertheless, Tables 5.33 and 5.34 show that, when overconfident trading is a significant explanatory variable in the conditional variance equation, the magnitude of the effect is overconfidence is greater than the magnitude of the non-overconfident trading effect.

In summary, for research question four, this study documents that, before the 2008 global financial crisis, trading volume induced by overconfident trading exhibits a positive effect on the conditional volatility of the market, however, this effect is only statistically significant for the market of ETFs tracking domestic benchmarks. Likewise, during the global financial crisis, the positive effect of overconfident trading on the market volatility is only significant for the market of ETFs tracking domestic benchmarks. On the contrary, after the 2008 global financial crisis, the positive effect of overconfident trading on market volatility is only statistically significant in the market of ETFs tracking international benchmarks. Overall, when the overconfidence effect is statistically significant, the effect of trading volume induced by overconfidence is greater than the effect of trading volume generated by non-overconfidence factors.

5.7 Summary

Chapter 5 presents and discusses the results obtained from the estimations of the different empirical methodologies. The results for each research question are summarised in Table 5.35 (page 137). In summary, this study reports four key empirical findings. Firstly, the results of the market VAR models suggest that investor overconfidence is present in the South African ETF market. Secondly, the results of the panel VAR models indicate that investor overconfidence influences the trading activities of individual ETFs in the South African ETF market. Therefore, the observed market-wide investor overconfidence is not a direct summation of the disposition effect. The third key empirical finding of this study suggests that trading volume induced by investor overconfidence contributes to excess volatility in the South African ETF market. Lastly, the sub-period analysis indicates investor overconfidence exhibits a positive effect on ETF market volatility before and during the 2008 global financial crisis, but this effect is only significant for the market of ETFs with domestic benchmarks. However, after the 2008 global financial crisis, the positive effect of investor overconfidence on ETF market volatility is only significant for the market of ETFs with international benchmarks.

Table 5.35: Summary of Results

Research Question	Finding(s)
Are investors overconfident when trading in the South African ETF market?	Estimated VAR models and their associated impulse response functions indicate that market-wide turnover significantly responds to lagged market returns in a positive manner, and thus, provides evidence in support of the overconfidence hypothesis. This relationship holds for both the market of South African ETFs tracking domestic benchmarks and the market of South African ETFs tracking international benchmarks.
Does investor overconfidence influence the trading activity of individual ETFs?	Impulse response functions associated with the estimated panel VAR models illustrate that lagged market returns also positively influences the turnover of individual ETFs in both markets. These findings provide evidence that the overconfidence bias influences the trading of individual ETFs. Therefore, the market-wide overconfidence found for research question one is not a direct summation of the disposition effect.
Does investor overconfidence exhibit a positive or negative effect on the volatility of the returns from the South African ETF market?	The estimated EGARCH (1,1) models suggest that, over the full sample period, overconfident trading exhibits a significant positive effect on the return volatility of both markets.
Does the volatility of the South African ETF market respond positively or negatively to trading volume induced by investor overconfidence before, during, and after the 2008 global financial crisis?	The sub-period analysis reveals that, for the pre-crisis period, there is a positive relationship between trading volume induced by overconfidence and volatility, but this relationship is only significant for the market of ETFs with domestic benchmarks. Similarly, during the crisis period, this positive relationship is only significant for the market of ETF with domestic benchmarks. On the contrary, after the crisis, the overconfidence effect is significant and positive only for the market of ETFs with international benchmarks.

CHAPTER 6: CONCLUSIONS AND RECOMMENDATIONS

6.1. Review of Objectives

The concept of market efficiency is based on the assumption that all traders are rational and, as such, the price at which an asset trades is an efficient price that fully incorporates all available information. On the contrary, behavioural finance theories argue that investors are not always rational because their trading decisions may be influenced by psychological and cognitive biases, subsequently, resulting in suboptimal investment choices. One of the most prominent behavioural biases is the overconfidence bias. Given the recent exponential growth in the ETF market, investment decisions made by ETF traders may not be rational, thus, posing a threat to the efficiency of ETF markets. For instance, investors may become overconfident if they recognise that ETFs are profitable investment instruments.

In 2019, the global shift of investors away from actively managed funds to passively managed funds, like ETFs, lead to several market experts questioning whether a bubble is brewing in passive investing which could subsequently lead to an ETF market crash. Therefore, the soaring popularity of ETFs could actually pose a threat to the stability of financial markets. The causes of price bubbles may be identified through an examination of the presence of behavioural biases in the respective market. Notably, overconfident investors overestimate asset prices resulting in price distortions that could trigger a market bubble. However, to the author's knowledge, there is only one existing study that covers overconfidence in ETF markets, and therefore, this lack of empirical studies that primarily focus on the overconfidence in ETF markets provides a basis for further analysis.

On this background, the objectives of this study are as follows:

- To investigate whether investors are overconfident when trading in the South African ETF market.
- To determine whether investor overconfidence influences the trading activities of individual ETFs.
- To examine the effect of investor overconfidence on the volatility of the returns from the South African ETF market.
- To assess the effect of overconfident trading on the volatility of the returns from the South African ETF market before, during, and after the 2008 global financial crisis.

To achieve the abovementioned objectives, this study employs three different empirical methodologies - the results of which are discussed in the preceding chapter. The next section provides a summary of the main findings in an attempt to address the research questions of this study.

6.2. Summary of Findings

6.2.1. Research Question One - Are Investors Overconfident When Trading in the South African ETF Market?

The results from the estimated VAR models reveal that there is a significant and positive relation between the current turnover of the market and the historical returns of the market. This finding holds for both the market of JSE-listed ETFs replicating local benchmarks and the market of JSE-listed ETFs replicating offshore benchmarks and holds even after controlling for other possible explanations of trading volume, viz. the volume-volatility relationship and portfolio rebalancing activities. Results from the Granger causality tests indicate that, for both markets, there is a unidirectional Granger causality from market return to market turnover. Moreover, the impulse response functions associated with the respective VAR models illustrate that, for both ETF markets, market turnover responds significantly and positively to shocks in market returns.

Overall, for both markets, the estimated VAR models, Granger causality tests and impulse response functions provide robust results of a significantly positive association between current market turnover and past market returns. Therefore, the results of this study support the overconfidence hypothesis which asserts that an increase(decrease) in past market returns increases(decreases) current trading because investors' confidence in their trading skills increases(decreases) subsequent to market gains(losses). Hence, it is concluded that investors are overconfident when trading in the market of South African ETFs replicating domestic benchmarks as well as in the market of South African ETFs replicating international benchmarks which, when combined, makes up the entire South African ETF market.

6.2.2. Research Question Two – Does Investor Overconfidence Influence the Trading Activity of Individual ETFs?

The results of the panel VAR models suggest that, for the market of JSE-listed ETFs replicating domestic benchmarks, there is no significant relationship between current security turnover and lagged market returns. However, for the market of JSE-listed ETFs replicating international benchmarks, the interrelation between current individual security turnover and past market returns is positive and statistically significant. Granger causality tests confirm that there is a significant causal relationship from market returns to security turnover for the market of ETFs replicating offshore benchmarks, but not for the market of ETFs replicating local benchmarks. Interestingly, the analysis of the associated impulse response functions reveals that, for both ETF markets, individual security turnover exhibits significant positive responses to shocks in the respective market return. This relation holds even after accounting for individual security return, and

therefore, suggests that investor overconfidence induced by market return positively impacts the realised trading volume of individual ETFs in both markets, even after controlling for the disposition effect. Overall, the positive response of security turnover to market return, even after accounting for security return, confirms that the market-wide overconfidence effect observed is not a simple aggregation of the disposition bias. Therefore, in the South African ETF market, the overconfidence effect accounts for a portion of ETF trading activity that is not captured by the disposition effect.

6.2.3. Research Question Three – Does Investor Overconfidence Exhibit a Positive or Negative Effect on the Volatility of the Returns From the South African ETF Market?

The results of the estimated EGARCH (1,1) models indicate that trading volume induced by investor overconfidence exhibits a significant positive relationship with the conditional variance of market returns. This relationship holds over the full sample period for both South Africa's market of ETFs replicating local benchmarks and the market of ETFs replicating offshore benchmarks. Therefore, these results suggest that high ETF market volatility may partially be due to the prevalence of investor overconfidence in the South African ETF market. Overall, this study concludes that an increase(decrease) in overconfident trading leads to an increase(decrease) in the volatility of the returns from the respective South African ETF market. Hence, the overconfident trading by investors in South Africa's ETF market may create excess volatility which could fuel the predicted passive investment bubble, and subsequently, may lead to an ETF market crash.

6.2.4. Research Question Four – Does the Volatility of the South African ETF Market Respond Positively or Negatively to Trading Volume Induced by Investor Overconfidence Before, During, and After the 2008 Global Financial Crisis?

The sub-period analysis reveals that, for both markets, overconfident trading positively impacts the conditional volatility before and during the 2008 global financial crisis, however, these influences are only significant for the market of JSE-listed ETFs replicating local benchmarks. Notably, in both markets, the magnitude of the overconfidence effect is the greater before the crisis suggesting that the excess volatility in the ETF market due to overconfident trading may have contributed to the volatility which fuelled the 2008 global financial crisis. Moreover, these findings suggest that overconfident trading could increase the volatility experienced during an ETF market crash. Regarding the post-crisis subsamples, the overconfidence effect is insignificant for the market of ETFs with local benchmarks but positive and significant for the market of ETFs with international benchmarks. Therefore, the results produced imply

that overconfidence is a persistent psychological bias and may delay economic recovery since it creates excess volatility in ETF markets.

6.3. Implications of Findings

The EMH asserts that all investors are rational, however, this study provides strong empirical evidence that the overconfidence bias influences the trading decisions of investors in the South African ETF market. These findings imply that the South African ETF market does not follow an efficient form, and thus, patterns in return, volatility, and trading activity may be better understood by accounting for behavioural biases, specifically, the overconfidence bias. Additionally, the results of this study prove that trading volume induced by investor overconfidence increases the volatility of the South African ETF market by creating price distortions which could fuel the predicted passive investment bubble. Therefore, the results of this study have important implications for various stakeholders, including the following:

- Overconfident investors should conduct regular analysis of their investments because the excessive trading by overconfident investors may actually reduce their returns. Moreover, investors trading in the South African ETF market can avoid large losses by avoiding highly traded ETFs that are subject to investor overconfidence. Therefore, it is important for investors to educate themselves on anomalous market behaviour and different pricing mechanisms so that they can calculate the ETF's true worth based on fundamental information. Given this, it is important for both overconfident investors and rational investors trading in South Africa's ETF market to seek the help of well-trained and professional financial advisors prior to making an investment decision.
- Financial advisors, fund managers, and investment management companies should mitigate biased investment decisions of their clients by giving advice that is well-educated and rational. However, in order to mitigate biased decisions, it is important for these stakeholders to first understand what biases affect investors' investment decisions. Overall, understanding what biases drive investment decisions will benefit both investment companies and individual investors by reducing their risk exposures and increasing their opportunities for generating higher returns.
- Examining ETF data from the perspective of investor overconfidence may help identify the missing link between the prices at which ETFs trade and their fundamental values. Thus, it is important for policymakers and regulators to implement policies that promote the efficiency of South Africa's ETF market by eradicating the overconfidence present in the South African ETF market. Enhancing the efficiency of the South African ETF market may help reduce the possibility of a passive

investment bubble since overconfidence creates excess market volatility and price distortions. Given this, policymakers and regulators should focus on the disclosure of reliable and unbiased information, such as ensuring that complete information about each ETF is published on an easily-accessible database, as well as improving the creation and redemption process of ETFs to limit manipulation of ETF prices. Moreover, to ensure that investors and advisors understand different behavioural biases, regulators educate investors and investment professional on the different aspects of behaviour finance by, for example, ensuring that behavioural finance is included in the curriculum of investment courses.

6.4. Recommendations for Future Studies

This study had a few limitations. Firstly, due to the minimum data requirements, the sample data did not include all South African ETFs listed on the JSE. Secondly, while there are several behavioural biases that could influence ETF investment choices, this study only focused on the overconfidence bias. Additionally, while investor overconfidence could influence different asset classes, this study only focused on the South African ETF market. Based on these limitations, the following research opportunities have been identified to contribute to existing knowledge on behavioural biases and the overconfidence bias in the South African context:

- The current study only examines the overconfidence bias, however, there are several other biases, such as anchoring and herding, that may influence the trading decisions of South African ETF investors. Therefore, future research could examine the presence and effect of other biases on ETF investment choices. Such studies could provide greater insight into the missing link between the prices at which ETFs trade and their fundamental values.
- An improvement of the current study would be to analyze primary data that is collected from investors as opposed to the secondary data and proxies used in this study. The use of primary data could improve the reliability of the statistical procedures as well as provide insight into other patterns of human behaviour.
- Several studies show that the level of overconfidence differs across changing market conditions, viz. bullish and bearish market conditions. Given this, a recommendation for future research is to investigate the lead-lag relationship between current trading activity and lagged market return under changing market conditions by employing a Markov-switching regime model.

- Since this study only evaluates overconfidence in the South African ETF market, future studies could evaluate the presence of overconfidence in other South African financial markets, such as, the equities, bonds, futures, and commodities markets.

6.5. Conclusion

The primary objective of the current study is to investigate the presence of overconfident trading by investors in South Africa's ETF market. Using a sample of JSE-listed ETFs replicating domestic benchmarks as well as a sample of JSE-listed ETFs replicating international benchmarks, the results obtained provide evidence in support of the overconfidence hypothesis which posits a positive association between the current turnover of the market and its past market returns. This relation holds for market-wide turnover and individual security turnover and even after controlling for alternative explanations of trading activity. Regarding the effect of overconfidence, this study finds that trading volume induced by investor overconfidence exhibits a significant positive relationship with the volatility of the ETF market. Moreover, using the 2008 global financial crisis as a structural break, the results presented reveal that overconfidence is a persistent bias influencing South African ETF investors -with the overconfidence effect being greater before the crisis. Overall, this study concludes that investors trading in the South African ETF market are overconfident, and such trading activities create price distortions that contribute to the excess volatility in the South African ETF market.

APPENDIX

Appendix 1: Summary of JSE-listed ETFs as at 30 August 2019

JSE Ticker	Long Name	Inception Date	Delisted Date	Type of Benchmark
Currently Listed ETFs				
STX40	Satrix 40 ETF	27 November 2000	--	Domestic
STXFIN	Satrix FINI ETF	15 February 2002	--	Domestic
STXIND	Satrix INDI ETF	15 February 2002	--	Domestic
GLD	NewGold ETF	01 November 2004	--	Domestic
SYGEU	Sygnia DJ EuroStoxx 50 ETF	10 October 2005	--	International
SYGUK	Sygnia FTSE 100 ETF	10 October 2005	--	International
STXSWX	Satrix SWIX Top 40 ETF	10 April 2006	--	Domestic
STXRES	Satrix RESI ETF	10 April 2006	--	Domestic
STXDIV	Satrix DIVI ETF	30 August 2007	--	Domestic
PTXSPY	CoreShares PropTrax SAPY ETF	25 September 2007	--	Domestic
SYGJP	Sygnia MSCI Japan ETF	01 April 2008	--	International
SYGUS	Sygnia MSCI USA ETF	01 April 2008	--	International
SYGWD	Sygnia MSCI World ETF	01 April 2008	--	International
GIVISA	NewFunds S&P GIVI SA Top 50 ETF	23 June 2008	--	Domestic
ASHT40	Ashburton Top 40 ETF	15 October 2008	--	Domestic
STXRAF	Satrix RAFI 40 ETF	16 October 2008	--	Domestic
NEWFSA	NewFunds NewSA ETF	01 December 2008	--	Domestic
NFSH40	NewFunds Shari'ah Top 40 ETF	06 April 2009	--	Domestic
ASHINF	Ashburton Inflation ETF	20 May 2009	--	Domestic
GIVFIN	NewFunds S&P GIVI SA Financial 15 ETF	15 June 2009	--	Domestic
GIVIND	NewFunds S&P GIVI SA Industrial 25 ETF	15 June 2009	--	Domestic
GIVRES	NewFunds S&P GIVI SA Resource 15 ETF	15 June 2009	--	Domestic
STAN40	Stanlib Top 40 ETF	18 October 2010	--	Domestic
STANSX	Stanlib SWIX 40 ETF	18 October 2010	--	Domestic
MAPPSG	NewFunds MAPPS Growth ETF	25 May 2011	--	Domestic
MAPPSP	NewFunds MAPPS Protect ETF	25 May 2011	--	Domestic
PTXTEN	CoreShares PropTrax Ten ETF	30 May 2011	--	Domestic
NFSWIX	NewFunds SWIX ETF	27 January 2012	--	Domestic

NFEMOM	NewFunds Equity Momentum ETF	27 January 2012	--	Domestic
NFGOVI	NewFunds GOVI ETF	27 January 2012	--	Domestic
NFILBI	NewFunds ILBI ETF	27 January 2012	--	Domestic
NFTRCI	NewFunds TRACI 3 Month ETF	27 January 2012	--	Domestic
PREFTX	CoreShares PrefTrax ETF	28 March 2012	--	Domestic
ASHMID	Ashburton Mid Cap ETF	15 August 2012	--	Domestic
STPROP	Stanlib SA Property ETF	14 February 2013	--	Domestic
NGPLT	NewPlat ETF	26 April 2013	--	Domestic
ETFPLD	AfricaPalladium ETF	26 March 2014	--	Domestic
NGPLD	NewPalladium ETF	27 March 2014	--	Domestic
ETFGLD	AfricaGold ETF	07 April 2014	--	Domestic
ETFPLT	AfricaPlatinum ETF	07 April 2014	--	Domestic
DIVTRX	CoreShares DivTrax ETF	14 April 2014	--	Domestic
CTOP50	CoreShares Top 50 ETF	03 June 2015	--	Domestic
KCCGLD	Krugerrand Custodial Certificate	17 July 2015	--	Domestic
ETFRHO	AfricaRhodium ETF	04 December 2015	--	Domestic
CSP500	CoreShares S&P 500 ETF	04 November 2016	--	International
GLPROP	Coeshares S&P Global Property ETF	04 November 2016	--	International
DCCUSD	Dollar Custodial Certificate - 10 Year	24 January 2017	--	International
STXILB	Satrix ILBI ETF	24 February 2017	--	Domestic
STXPRO	Satrix Property ETF	24 February 2017	--	Domestic
AMIB50*	AMI Big 50 (ex-SA) ETF	20 April 2017	--	International
STX500*	Satrix S&P 500 ETF	25 July 2017	--	International
STXEMG*	Satrix MSCI Emerging Markets ETF	25 July 2017	--	International
STXWDM*	Satrix MSCI World ETF	25 July 2017	--	International
ASHGEQ*	Ashburton Global 1200 Equity ETF	06 October 2017	--	International
SYG500*	Sygnia Itrix S&P 500	30 October 2017	--	International
SYGSW4*	Sygnia Itrix SWIX 40 ETF	07 November 2017	--	Domestic
SYGT40*	Sygnia Itrix Top 40 ETF	07 November 2017	--	Domestic
SYGP*	Sygnia Itrix Global Property ETF	07 November 2017	--	International
DCCUS2*	Dollar Custodial Certificate - 2 Year	14 November 2017	--	International
SYG4IR*	Sygnia 4th Industrial Revolution Global Equity ETF	06 December 2017	--	International

GLODIV*	CoreShares S&P Global Dividend Aristocrats ETF	22 February 2018	--	International
ASHWGB*	Ashburton World Government Bond ETF	12 March 2018	--	International
ETF500*	Stanlib S&P 500 Index Feeder ETF	15 March 2018	--	International
ETF5IT*	Stanlib S&P 500 Info Tech Index Feeder ETF	15 March 2018	--	International
ETFWLD*	Stanlib MSCI World Index Feeder ETF	15 March 2018	--	International
ETFGGB*	Stanlib Global Government Bond Index Feeder ETF	15 March 2018	--	International
ETFGRE*	Stanlib Global REIT Index Feeder ETF	15 March 2018	--	International
NFEVAL*	NewFunds Value Equity ETF	26 March 2018	--	Domestic
NFEVOL*	NewFunds Low Volatility Equity ETF	26 March 2018	--	Domestic
STXNDQ*	Satrix Nasdaq 100 ETF	10 April 2018	--	International
AMIRE*	AMI Real Estate (ex-SA) ETF	01 June 2018	--	International
STXQUA*	Satrix Quality South Africa ETF	26 September 2018	--	Domestic
STXMMT*	Satrix Momentum ETF	16 November 2018	--	Domestic
NFEHGE*	NewFunds Volatility Managed High Growth Equity ETF	25 February 2019	--	Domestic
NFEDEF*	NewFunds Volatility Managed Defensive Equity ETF	27 February 2019	--	Domestic
NFEMOD*	NewFunds Volatility Managed Moderate Equity ETF	27 February 2019	--	Domestic
ETFBND*	Stanlib SA Bond ETF	15 June 2019	--	Domestic
SMART*	Coeshares Scibeta m-fi	10 July 2019	--	Domestic
Currently Delisted ETFs				
NRD	NewFunds Rand	25 June 2003	28 February 2014	Domestic
ZRNHDG*	Zshares Randhedge	04 December 2007	24 July 2009	Domestic
ZRNPLY*	Zshares Randplay	04 December 2007	24 July 2009	Domestic
ZGOVI	Zshares Govi	21 October 2008	20 January 2014	Domestic
BBET40	BettaBeta Equally Weighted TOP40	25 March 2010	06 May 2016	Domestic
BGREEN	BettaBeta CIS BGreen Portfolio	1 December 2011	06 May 2016	Domestic
LVLTRX	CoreShares Low Volatility ETF	14 April 2014	14 August 2018	Domestic
CGREEN*	CoreShares Green ETF	09 May 2016	22 August 2017	Domestic

CSEW40	Coreshares Equally Weighted Top 40 ETF	09 May 2016	09 July 2019	Domestic
--------	--	-------------	--------------	----------

Notes:

1. * denotes ETFs not included in the sample since they have been registered for less than 30 months.
2. Source: authors own compilation using data provided by etfsa (2019).
3. Table compiled on the 30th of August 2019.

Appendix 2: VAR Model Estimates for the Market of ETFs Tracking Domestic Benchmarks

	Mturn_dom (-1)	Mturn_dom (-2)	Mturn_dom (-3)	Mturn_dom (-4)	Mturn_dom (-5)	Mturn_dom (-6)
Mret_dom	-0.0290 (-0.3486)	-0.0741 (-0.8618)	0.1393 (1.5879)	0.0252 (0.2832)	-0.0215 (-0.2484)	-0.0296 (-0.3578)
	Mret_dom (-1)	Mret_dom (-2)	Mret_dom (-3)	Mret_dom (-4)	Mret_dom (-5)	Mret_dom (-6)
Mret_dom	-0.1000 (-1.4410)	0.0912 (1.3032)	0.1670** (2.2947)	0.0653 (0.9059)	-0.0517* (-1.8961)	0.1007*** (3.4163)
	C	Msig_dom (-1)	Msig_dom (-2)	Msig_dom (-3)	Msig_dom (-4)	Msig_dom (-4)
Mret_dom	-0.0113 (-1.2235)	-0.0385 (-1.2294)	0.0058 (0.1701)	-0.0055 (-0.1651)	-0.0008 (-0.0246)	-0.0832*** (-2.6897)
	Disp_dom (-1)	Disp_dom (-2)	Disp_dom (-3)	Disp_dom (-4)	Disp_dom (-4)	
Mret_dom	4.9630*** (21.0063)	0.4009 (0.9499)	-0.1141 (-0.2809)	-1.0954** (-2.5265)	0.0900 (0.2144)	

Notes:

1. Values in brackets ‘()’ represents t-statistics.
2. ***, **, * indicate statistical significance at 1%, 5%, and 10% levels, respectively.

Appendix 3: VAR Model Estimates for the Market of ETFs Tracking International Benchmarks

	Mturn_int (-1)	Mturn_int (-2)	Mturn_int (-3)	Mret_int (-1)	Mret_int (-2)	Mret_int (-3)
Mret_int	0.0982 (0.6786)	0.1919 (1.3055)	-0.0546 (-0.4007)	-0.0583 (-0.6908)	-0.0254 (-0.2953)	0.0192 (0.2161)
	C	Msig_int	Msig_int (-1)	Msig_int (-2)	Msig_int (-3)	Msig_int (-4)
Mret_int	0.0251 (1.5717)	-0.1726 (-0.7369)	-0.3060 (-1.2903)	0.0896 (0.3746)	-0.0255 (-0.1067)	-0.0246 (-0.1000)
	Msig_int (-5)	Msig_int (-6)	Msig_int (-7)	Disp_int	Disp_int (-1)	Disp_int (-2)
Mret_int	-0.0246 (-0.4658)	0.0795 (0.3369)	0.5629** (2.4798)	1.1734 (1.1018)	-0.9653 (-0.9015)	-0.3429 (-0.3196)
	Disp_int (-3)	Disp_int (-4)	Disp_int (-5)	Disp_int (-6)	Disp_int (-7)	
Mret_int	0.4589 (0.4288)	-0.7947 (-0.7184)	-0.2925 (-0.2665)	-0.0863 (-0.0795)	-1.8840* (-1.8068)	

Notes:

1. Values in brackets () represents t-statistics.
2. ***, **, * indicate statistical significance at 1%, 5%, and 10% levels, respectively.

Appendix 4: Impulse Response Table for the VAR Model Estimated for the Market of ETFs with Domestic Benchmarks

Response of MTURN_DOM:			
Period	MTURN_DOM		MRET_DOM
1	0.038905 (0.00185)		0.000000 (0.00000)
2	0.012155 (0.00283)		0.005191 (0.00270)
3	0.013060 (0.00298)		0.007976 (0.00287)
4	0.011684 (0.00306)		0.002645 (0.00306)
5	0.006574 (0.00310)		0.000993 (0.00308)

6	0.009231 (0.00298)	0.002262 (0.00219)
7	0.010380 (0.00249)	0.001544 (0.00220)
8	0.007625 (0.00228)	0.002036 (0.00197)
9	0.007769 (0.00235)	0.002930 (0.00188)
10	0.006768 (0.00235)	0.001725 (0.00187)

Response of MRET_DOM:

Period	MTURN_DOM	MRET_DOM
1	-0.004135 (0.00307)	0.045398 (0.00216)
2	-0.000714 (0.00324)	-0.004541 (0.00316)
3	-0.003541 (0.00321)	0.004445 (0.00315)
4	0.003738 (0.00321)	0.006105 (0.00317)
5	0.000282 (0.00322)	0.002058 (0.00315)
6	0.000121 (0.00311)	-0.000533 (0.00179)
7	-0.000185 (0.00279)	0.006677 (0.00174)
8	1.63E-05 (0.00148)	-0.000994 (0.00133)
9	-0.000707 (0.00144)	0.000760 (0.00116)
10	0.000812 (0.00131)	0.001359 (0.00106)

Notes:

1. Values in brackets represent analytic standard errors.

Appendix 5: Impulse Response Table for the VAR Model Estimated for the Market of ETFs with International Benchmarks

Response of MTURN_INT:		
Period	MTURN_INT	MRET_INT
1	0.021755 (0.00122)	0.000000 (0.00000)
2	0.005960 (0.00167)	0.002247 (0.00184)
3	0.002242 (0.00158)	0.004844 (0.00196)
4	0.008459 (0.00147)	0.003591 (0.00203)
5	0.005002 (0.00133)	0.001761 (0.00108)
6	0.002674 (0.00126)	0.002408 (0.00109)
7	0.003895 (0.00125)	0.002155 (0.00107)
8	0.003121 (0.00113)	0.001382 (0.00077)
9	0.002080 (0.00102)	0.001343 (0.00071)
10	0.002106 (0.00101)	0.001231 (0.00067)

Response of MRET_INT:		
Period	MTURN_INT	MRET_INT
1	-0.001286 (0.00332)	0.041970 (0.00235)
2	0.002211 (0.00316)	-0.002445 (0.00354)
3	0.004663 (0.00299)	-0.000703 (0.00358)
4	-0.000176 (0.00272)	0.001814 (0.00363)
5	0.000870 (0.00129)	0.001025 (0.00094)
6	0.002035 (0.00120)	0.000478 (0.00074)
7	0.000617 (0.00111)	0.000359 (0.00055)

8	0.000552 (0.00074)	0.000564 (0.00040)
9	0.000899 (0.00061)	0.000385 (0.00039)
10	0.000536 (0.00057)	0.000250 (0.00031)

Notes:

1. Values in brackets represent analytic standard errors.

Appendix 6: Panel VAR Model Estimates for the Market of ETFs Tracking Domestic Benchmarks

	Turn_dom (-1)	Turn_dom (-2)	Turn_dom (-3)	Turn_dom (-4)	Turn_dom (-5)	Turn_dom (-6)
Ret_dom	0.0195 (1.0171)	0.0139 (0.7205)	-0.0092 (-0.4790)	0.0094 (0.4865)	0.0131 (0.7041)	-0.0005 (-0.0995)
Mret_dom	0.0015 (0.2919)	-0.0093* (-1.8670)	0.0071 (1.4191)	-0.0085* (-1.6864)	0.0068 (1.4081)	-8.21E-06 (-0.0064)
	Ret_dom (-1)	Ret_dom (-2)	Ret_dom (-3)	Ret_dom (-4)	Ret_dom (-5)	Ret_dom (-6)
Ret_dom	-0.4914*** (-31.2011)	-0.2028*** (-11.8709)	-0.0023 (-0.1459)	-0.0476*** (-3.9273)	-0.1204*** (-11.4495)	0.0219** (2.3238)
Mret_dom	0.0071* (1.7206)	-0.0007 (-0.1509)	0.0048 (1.1463)	-1.37E-05 (-0.0043)	-0.0066** (-2.4149)	-0.0021 (-0.8603)
	Mret_dom (-1)	Mret_dom (-2)	Mret_dom (-3)	Mret_dom (-4)	Mret_dom (-5)	Mret_dom (-6)
Ret_dom	0.1617*** (3.18540)	0.1192** (2.3381)	0.4375*** (8.7274)	0.2312*** (4.7612)	0.0575 (1.1899)	0.4792*** (10.0528)
Mret_dom	-0.1185*** (-8.9730)	0.0574*** (4.3263)	0.3304*** (25.3189)	0.1096*** (8.6704)	-0.0310** (-2.4613)	0.4531*** (36.5191)
	C	Sig_dom	Sig_dom (-1)	Sig_dom (-2)	Sig_dom (-3)	Sig_dom (-4)
Ret_dom	-0.0007 (-0.1613)	-0.2801*** (-28.7209)	0.2967*** (28.0329)	0.1569*** (15.9782)	0.0911*** (9.3732)	-0.1694*** (-17.9563)
Mret_dom	0.0034*** (2.9044)	0.0141*** (5.5592)	-0.0040 (-1.4595)	-0.0017 (-0.6508)	0.0028 (1.0949)	-0.0060** (-2.4518)

Notes:

1. Values in brackets () represents t-statistics.
2. ***, **, * indicate statistical significance at 1%, 5%, and 10% levels, respectively.

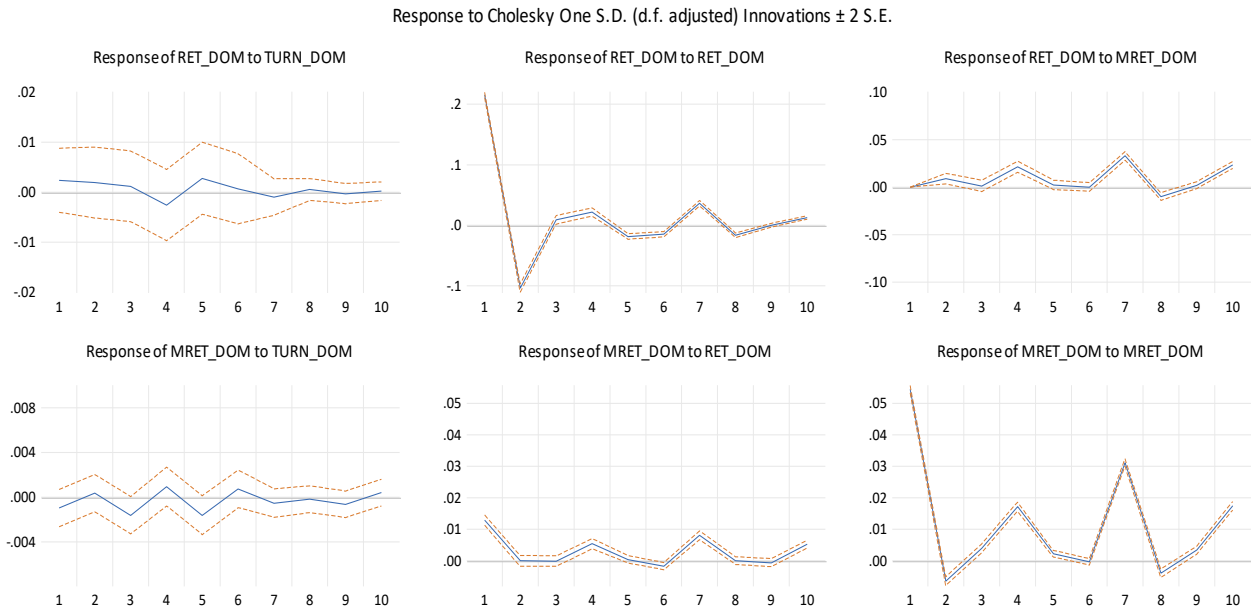
Appendix 7: Panel VAR Model Estimates for the Market of ETFs Tracking International Benchmarks

	Turn_int (-1)	Turn_int (-2)	Turn_int (-3)	Ret_int (-1)	Ret_int (-2)	Ret_int (-3)
Ret_int	0.0504 (1.3581)	0.0724* (1.9021)	-0.0308 (-0.8298)	-0.2951*** (-4.8094)	-0.0713 (-1.1277)	-0.0138 (-0.2249)
Mret_int	0.0155 (0.4728)	0.0567* (1.6868)	-0.0197 (-0.6011)	-0.0489 (-0.9011)	-0.0591 (-1.0580)	-0.0792 (-1.4626)
	Mret_int (-1)	Mret_int (-2)	Mret_int (-3)	C	Sig_int	Sig_int (-1)
Ret_int	0.2177*** (3.1697)	0.1253* (1.7866)	0.0089 (0.1290)	0.0260*** (3.9073)	-0.2050*** (-2.8263)	-0.2960*** (-4.2806)
Mret_int	-0.0652 (-1.0729)	0.0078 (0.1254)	0.0732 (1.1984)	0.0311*** (5.2929)	-0.1243* (-1.9383)	-0.3250*** (-5.3166)
	Sig_int (-2)	Sig_int (-3)	Sig_int (-4)	Sig_int (-5)	Sig_int (-6)	Sig_int (-7)
Ret_int	0.0894 (1.2698)	0.0524 (0.7413)	-0.1939*** (-2.8227)	-0.1098 (-1.6234)	0.1562* (2.3391)	0.0559 (0.8584)
Mret_int	0.0068 (0.1086)	-0.0415 (-0.6651)	-0.1325** (-2.1823)	-0.0433 (-0.7239)	0.1522** (2.5780)	0.1196** (2.0776)

Notes:

1. Values in brackets () represents t-statistics.
2. ***, **, * indicate statistical significance at 1%, 5%, and 10% levels, respectively.

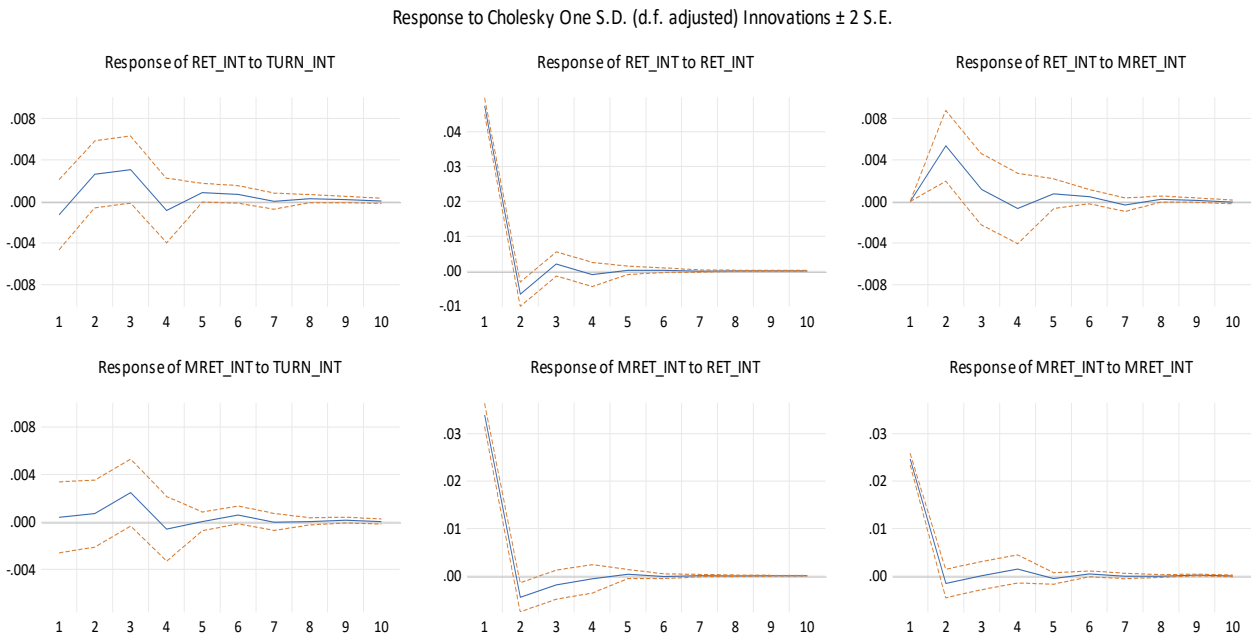
Appendix 8: Impulse Response Functions for Individual ETFs in the Market of ETFs Tracking Domestic Benchmarks



Notes:

1. The blue line represents the impulse response and the red lines represent two standard error bands.

Appendix 9: Impulse Response Functions for Individual ETFs in the Market of ETFs Tracking International Benchmarks



Notes:

1. The blue line represents the impulse response and the red lines represent two standard error bands.

Appendix 10: Impulse Response Table for Individual ETFs in the Market of ETFs Tracking Domestic Benchmarks

Period	TURN_DOM	RET_DOM	MRET_DOM
1	0.166077 (0.00174)	0.000000 (0.00000)	0.000000 (0.00000)
2	0.017346 (0.00246)	0.001502 (0.00252)	0.000491 (0.00213)
3	0.016052 (0.00246)	0.002287 (0.00246)	0.002956 (0.00213)
4	0.017310 (0.00246)	-0.000456 (0.00223)	0.002986 (0.00208)
5	0.013996 (0.00247)	-0.000422 (0.00160)	0.003668 (0.00158)
6	0.014190 (0.00236)	0.000164 (0.00152)	0.003192 (0.00154)
7	0.006516 (0.00107)	-0.000145 (0.00141)	0.001778 (0.00155)
8	0.004970 (0.00086)	0.000656 (0.00063)	0.002495 (0.00102)
9	0.004019 (0.00080)	0.000777 (0.00048)	0.003262 (0.00098)
10	0.002974 (0.00069)	0.000350 (0.00041)	0.002223 (0.00098)

Notes:

1. Values in brackets represent analytic standard errors.

Appendix 11: Impulse Response Functions for Individual ETFs in the Market of ETFs Tracking International Benchmarks

Period	TURN_INT	RET_INT	MRET_INT
1	0.042714 (0.00108)	0.000000 (0.00000)	0.000000 (0.00000)
2	0.009028 (0.00145)	0.003317 (0.00154)	0.000164 (0.00152)
3	0.004658 (0.00145)	0.003675 (0.00158)	9.40E-05 (0.00155)

4	0.011747 (0.00144)	0.003592 (0.00160)	0.007676 (0.00154)
5	0.004782 (0.00097)	0.001487 (0.00065)	0.000499 (0.00069)
6	0.003048 (0.00091)	0.000466 (0.00065)	0.000422 (0.00066)
7	0.003626 (0.00091)	0.000971 (0.00055)	0.002432 (0.00063)
8	0.001961 (0.00062)	0.000657 (0.00028)	0.000401 (0.00028)
9	0.001413 (0.00052)	0.000242 (0.00025)	0.000371 (0.00026)
10	0.001257 (0.00048)	0.000332 (0.00020)	0.000702 (0.00025)

Notes:

1. Values in brackets represent analytic standard errors.

Appendix 12: Results of the Estimated Regressions Used to Distinguish Between Overconfident Trading and Non-Overconfident Trading for the Market of ETFs Tracking Domestic Benchmarks

Dependent Variable: T_DOM

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-5.851367	0.024515	-238.6868	0.0000
R_DOM(-1)	0.546304	0.570790	0.957100	0.3386
R_DOM(-2)	-0.539452	0.570806	-0.945069	0.3447
R_DOM(-3)	-0.723920	0.570827	-1.268195	0.2048
R_DOM(-4)	-0.446923	0.570836	-0.782927	0.4337
R_DOM(-5)	0.385434	0.570776	0.675281	0.4995
R_DOM(-6)	0.027735	0.570756	0.048593	0.9612
R_DOM(-7)	-1.356004	0.570740	-2.375869	0.0175
R-squared	0.002087	Mean dependent var		-5.859493
Adjusted R-squared	0.000655	S.D. dependent var		1.663766
S.E. of regression	1.663221	Akaike info criterion		3.857025
Sum squared resid	13499.56	Schwarz criterion		3.867654
Log likelihood	-9418.568	Hannan-Quinn criter.		3.860754

F-statistic	1.457839	Durbin-Watson stat	1.472106
Prob(F-statistic)	0.177551		

Appendix 13: Results of the Estimated Regressions Used to Distinguish Between Overconfident Trading and Non-Overconfident Trading for the Market of ETFs Tracking International Benchmarks

Dependent Variable: T_INT

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-6.551814	0.018016	-363.6571	0.0000
R_INT(-1)	-1.316370	1.549337	-0.849635	0.3956
R_INT(-2)	3.309674	1.576120	2.099887	0.0358
R_INT(-3)	-0.140908	1.586370	-0.088824	0.9292
R-squared	0.001545	Mean dependent var		-6.550910
Adjusted R-squared	0.000679	S.D. dependent var		1.056591
S.E. of regression	1.056232	Akaike info criterion		2.948446
Sum squared resid	3862.295	Schwarz criterion		2.955544
Log likelihood	-5105.656	Hannan-Quinn criter.		2.950980
F-statistic	1.785344	Durbin-Watson stat		1.408970
Prob(F-statistic)	0.147721			

Appendix 14: Ethical Clearance Letter



Mr Damien Kunjal (215057603)
School Of Acc Economics&Fin
Westville

Dear Mr Damien Kunjal,

Protocol reference number: 00001799

Project title: The Presence of Investor Overconfidence in the South African Exchange Traded Fund Market

Exemption from Ethics Review

In response to your application received on 9 May 2019, your school has indicated that the protocol has been granted **EXEMPTION FROM ETHICS REVIEW**.

Any alteration/s to the exempted research protocol, e.g., Title of the Project, Location of the Study, Research Approach and Methods must be reviewed and approved through an amendment/modification prior to its implementation. The original exemption number must be cited.

For any changes that could result in potential risk, an ethics application including the proposed amendments must be submitted to the relevant UKZN Research Ethics Committee. The original exemption number must be cited.

In case you have further queries, please quote the above reference number.

PLEASE NOTE:

Research data should be securely stored in the discipline/department for a period of 5 years.

I take this opportunity of wishing you everything of the best with your study.

Yours sincerely,



Prof Josue Mbonigaba
Academic Leader Research
School Of Acc Economics&Fin

UKZN Research Ethics Office
Westville Campus, Govan Mbeki Building
Postal Address: Private Bag X54001, Durban 4000
Website: <http://research.ukzn.ac.za/Research-Ethics/>

Founding Campuses: ■ Edgewood ■ Howard College ■ Medical School ■ Pietermaritzburg ■ Westville

INSPIRING GREATNESS

Appendix 15: Amended Ethical Clearance Letter for Change in Title



30 March 2020

Mr Damien Kunjal (215057603)
School of Acc, Economics and Finance
Westville

Dear Mr Damien Kunjal,

System Nr: 00001799

Project title: Investor Overconfidence in the South African Exchange Traded Fund Market.

Approval Notification – Amendment Application

This letter serves to notify you that your application and request for an amendment received on 27 March 2020 has now been approved as follows:

- Change in Title

Any alterations to the approved research protocol i.e. Questionnaire/Interview Schedule, Informed Consent Form; Title of the Project, Location of the Study must be reviewed and approved through an amendment /modification prior to its implementation. In case you have further queries, please quote the above reference number.

PLEASE NOTE: Research data should be securely stored in the discipline/department for a period of 5 years.

Best wishes for the successful completion of your research protocol.

Yours faithfully

Prof Josue Mbonigaba

30 March 2020

.....
ACADEMIC LEADER RESEARCH

/___

Humanities & Social Sciences Research Ethics Committee
UKZN Research Ethics Office Westville Campus, Govan Mbeki Building
Postal Address: Private Bag X54001, Durban 4000
Tel: +27 31 260 8350 / 4557 / 3587
Website: <http://research.ukzn.ac.za/Research-Ethics/>

Founding Campuses: Edgewood Howard College Medical School Pietermaritzburg Westville

INSPIRING GREATNESS

BIBLIOGRAPHY

- Abbes, M.B. (2013) Does overconfidence bias explain volatility during the global financial crisis?. *Transition Studies Review*, 19(3): 291-312.
- Abdellaoui, M., Bleichrodt, H., and l'Haridon, O. (2008) A tractable method to measure utility and loss aversion under prospect theory. *Journal of Risk and Uncertainty*, 36(3): 245-266.
- Abrigo, M.R. and Love, I. (2016) Estimation of panel vector autoregression in Stata. *The Stata Journal*, 16(3): 778-804.
- Adu, G., Alagidede, P., and Karimu, A. (2015) Stock return distribution in the BRICS. *Review of Development Finance*, 5(2): 98-109.
- Aharon, D.Y. and Qadan, M. (2018) What drives the demand for information in the commodity market?. *Resources Policy*, 59: 532-543.
- Akaike, H. (1973) Information theory and an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory*, ed. B. N. Petrov and F. Csaki, 267-281. Budapest: Akademiai Kiado.
- Aldridge, I. (2016) ETFs, high-frequency trading, and flash crashes. *Journal of Portfolio Management*, 43(1): 17-28.
- Alexakis, C., Patra, T., and Poshakwale, S. (2010) Predictability of stock returns using financial statement information: evidence on semi-strong efficiency of emerging Greek stock market. *Applied Financial Economics*, 20(16): 1321-1326.
- Alicke, M.D. (1985) Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology*, 49(6): 1621-1630.
- Allison, P. (2012) When can you safely ignore multicollinearity?. *Statistical Horizons*, 10 September 2012, <https://statisticalhorizons.com/multicollinearity> [Accessed 27 November 2019].

Alpert, M. and Raiffa, H. (1982) *Judgement under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

Alsabban, S. and Alarfaj, O. (2020) An empirical analysis of behavioral finance in the Saudi stock market: evidence of overconfidence behavior. *International Journal of Economics and Financial Issues*, 10(1): 73-86.

Amenc, N., Goltz, F., and Le Sourd, V. (2017) The EDHEC European ETF and Smart Beta Survey 2016. *EDHEC-Risk Institute*, May 2017, https://risk.edhec.edu/sites/risk/files/pdf/edhec_publication_the_edhec_european_etf_and_smart_beta_survey_2016_f.pdf [Accessed 18 January 2020].

Atkinson, A.C. (1986) Diagnostic tests for transformations. *Technometrics*, 28(1): 29-37.

Atoi, N.V. (2014) Testing volatility in Nigeria stock market using GARCH models. *CBN Journal of Applied Statistics*, 5(2): 65-93.

Attari, M.I.J., Rafiq, S., and Awan, H.M. (2012) The dynamic relationship between stock volatility and trading volume. *Asian Economic and Financial Review*, 2(8): 1085-1097.

Bahadar, S., Mahmood, H., and Zaman, R. (2019) The herds of bulls and bears in leveraged ETF market. *Journal of Behavioral Finance*, 20(4): 408-423.

Bailey, W., Kumar, A., and Ng, D. (2011) Behavioral biases of mutual fund investors. *Journal of Financial Economics*, 102(1): 1-27.

Baker, H.K., Kumar, S., Goyal, N., and Gaur, V. (2019) How financial literacy and demographic variables relate to behavioral biases. *Managerial Finance*, 45(1): 124-146.

Baltagi, B.H. and Pesaran, H.M. (2007) Heterogeneity and cross section dependence in panel data models: theory and applications introduction. *Journal of Applied Econometrics*, 22(2): 229-232.

Baltagi, B.H., Bresson, G., and Pirotte, A. (2007) Panel unit root tests and spatial dependence. *Journal of Applied Econometrics*, 22(2): 339-360.

Barber, B.M. and Odean, T. (1999) The courage of misguided convictions. *Financial Analysts Journal*, 55(6): 41-55.

Barber, B.M. and Odean, T. (2000) Trading is hazardous to your wealth: the common stock investment performance of individual investors. *The Journal of Finance*, 55(2): 773-806.

Barber, B.M. and Odean, T. (2001) Boys will be boys: gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics*, 116(1): 261-292.

Barberis, N. and Huang, M. (2001) Mental accounting, loss aversion, and individual stock returns. *The Journal of Finance*, 56(4): 1247-1292.

Beard, E., Marsden, J., Brown, J., Tombor, I., Stapleton, J., Michie, S., and West, R. (2019) Understanding and using time series analyses in addiction research. *Addiction*, 114(10): 1866-1884.

Beaver, W.H. (1968) The information content of annual earnings announcements. *Journal of Accounting Research*, 6: 67-92.

Ben-David, I., Franzoni, F., and Moussawi, R. (2018) Do ETFs increase volatility?. *The Journal of Finance*, 73(6): 2471-2535.

Benigno, P. and Karantounias, A.G. (2019) Overconfidence, subjective perception and pricing behavior. *Journal of Economic Behavior and Organization*, 164: 107-132.

Benos, A.V. (1998) Aggressiveness and survival of overconfident traders. *Journal of Financial Markets*, 1: 353-383.

BenSaïda, A. (2017) Herding effect on idiosyncratic volatility in US industries. *Finance Research Letters*, 23: 121-132.

Berg, N. and Lein, D. (2005) Does society benefit from investor overconfidence in the ability of financial market experts?. *Journal of Economic Behavior and Organization*, 58(1): 95-116.

Bernstein, P.J. (2002) A primer on exchange-traded funds. *Journal of Accountancy*, 193(1): 38-41.

Bhattacharya, A. and O'Hara, M. (2020). ETFs and systemic risks. *CFA Institute*, 22 January 2020, <https://www.cfainstitute.org/en/research/foundation/2020/etfs-and-systemic-risks> [Accessed 28 February 2020].

Bidian, F. (2015) Portfolio constraints, differences in beliefs and bubbles. *Journal of Mathematical Economics*, 61: 317-326.

Blanco, F., Matute, H., and Vadillo, M.A. (2011) Making the uncontrollable seem controllable: the role of action in the illusion of control. *The Quarterly Journal of Experimental Psychology*, 64(7): 1290-1304.

Blau, B.M. and Smith, J.M. (2014) Autocorrelation in daily short-sale volume. *The Quarterly Review of Economics and Finance*, 54(1): 31-41.

Blitz, D. and Huij, J. (2012) Evaluating the performance of global emerging markets equity exchange-traded funds. *Emerging Markets Review*, 13(2): 149-158.

Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3): 307-327.

Breusch, T.S. (1978) Testing for Autocorrelation in Dynamic Linear Models. *Australian Economic Papers*, 17(31): 334-355.

Broman, M.S. (2016) Liquidity, style investing and excess comovement of exchange-traded fund returns. *Journal of Financial Markets*, 30: 27-53.

Broms, T.J. and Gastineau, G.L. (2007) The development of improved exchange-traded funds (ETFs) in the United States. *ETFs and Indexing*, Fall 2007(1): 16-26.

Brooks, C. (2002) *Introductory Econometrics for Finance*. 1st ed. Cambridge: Cambridge University Press.

Brooks, C. (2008) *Introductory Econometrics for Finance*. 2nd ed. Cambridge: Cambridge University Press.

Brown, M. (2019) Index tracking ETFs continue to dominate South Africa's investment performance surveys over multiple time periods. *etfSA.co.za*, February 2019, http://etfSA.co.za/docs/perfsurvey/index_tracking_etfs_feb19.pdf [Accessed 14 March 2020].

Brown, M. (2020) Market capitalisation data for the South African exchange traded industry for the period ended 30 June 2020. *etfSA.co.za*, July 2020, <https://www.etfSA.co.za/docs/perfsurvey/market%20cap%20-%20july2020.pdf> [Accessed 25 July 2020].

Buehler, R., Griffin, D., and Ross, M. (1994) Exploring the planning fallacy: why people underestimate their task completion times. *Journal of Personality and Social Psychology*, 67(3): 366-381.

Burger, J.M. (1991) The effects of desire for control in situations with chance-determined outcomes: gambling behavior in lotto and bingo players. *Journal of Research in Personality*, 25(2): 196-204.

Canova, F. and Ciccarelli, M. (2013) Panel vector autoregressive models: a survey. *ECB Working paper*, No. 1507, <https://www.econstor.eu/handle/10419/153940>.

Charemza, W.W. and Deadman, D.F. (1997) *Econometrics practice: general to specific modelling, cointegration and vector autoregression*. 2nd ed. United Kingdom: Edward Elgar.

Charteris, A. (2013) The price efficiency of South African exchange traded funds. *Investment Analysts Journal*, 42(78): 1-11.

Charteris, A. and Musadziruma, A. (2017) Feedback trading in stock index futures: evidence from South Africa. *Research in International Business and Finance*, 42: 1289-1297.

Charteris, A. and Rupande, L. (2017) Feedback trading on the JSE. *Investment Analysts Journal*, 46(3): 235-248.

Charteris, A., Chau, F., Gavriilidis, K., and Kallinterakis, V. (2014) Premiums, discounts and feedback trading: evidence from emerging markets' ETFs. *International Review of Financial Analysis*, 35: 80-89.

Chen, C.D., Demirer, R., and Jategaonkar, S.P. (2015) Risk and return in the Chinese stock market: does equity return dispersion proxy risk?. *Pacific-Basin Finance Journal*, 33: 23-37.

Chen, G., Kim, K.A., Nofsinger, J.R., and Rui, O.M. (2007) Trading performance, disposition effect, overconfidence, representativeness bias, and experience of emerging market investors. *Journal of Behavioral Decision Making*, 20(4): 425-451.

Chen, H.C., Ho, J.K., Lai, C.W., and Morales-Camargo, E. (2011) Institutional trading behavior in the ETF market. In *Midwest Finance Association 2012 Annual Meetings Paper*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1934056.

Chen, H.S. and Sabherwal, S. (2019) Overconfidence among option traders. *Review of Financial Economics*, 37(1): 61-91.

Chipunza, K.J. and McCullough, K. (2018) The impact of internationalisation on stock liquidity and volatility: evidence from the Johannesburg Stock Exchange. *Journal of Economic and Financial Sciences*, 11(1): 1-11.

Choe, H. and Eom, Y. (2009) The disposition effect and investment performance in the futures market. *Journal of Futures Markets*, 29(6): 496-522.

Choi, I. (2001) Unit root tests for panel data. *Journal of International Money and Finance*, 20(2): 249-272.

Choi, I. (2006) Combination unit root tests for cross-sectionally correlated panels. *Econometric Theory and Practice: Frontiers of Analysis and Applied Research*. Cambridge: Cambridge University Press, 311-333.

Chordia, T., Huh, S.W., and Subrahmanyam, A. (2006) The cross-section of expected trading activity. *The Review of Financial Studies*, 20(3): 709-740.

Chou, R.K. and Wang, Y.Y. (2011) A test of the different implications of the overconfidence and disposition hypotheses. *Journal of Banking and Finance*, 35(8): 2037-2046.

Chuang, W.I. and Lee, B.S. (2006) An empirical evaluation of the overconfidence hypothesis. *Journal of Banking and Finance*, 30(9): 2489-2515.

Chuang, W.I. and Susmel, R. (2011) Who is the more overconfident trader? Individual vs. institutional investors. *Journal of Banking and Finance*, 35(7): 1626-1644.

- Chung, K.H., Li, M. and McInish, T.H. (2005) Information-based trading, price impact of trades, and trade autocorrelation. *Journal of Banking and Finance*, 29(7): 1645-1669.
- Clark, P. (1973) A subordinated stochastic process model with finite variance for speculative prices, *Econometrica: Journal of the Econometric Society*, 41: 135-155
- Cohen, B.H. and Shin, H.S. (2013) Positive feedback trading under stress: evidence from the US treasury securities market. *Global Economic Review*, 42(4): 314-345.
- Colegrave, A. (2008) The public and collective investment schemes. *Without Prejudice*, 8(6): 40.
- Conrad, J. and Kaul, G. (1988) Time-variation in expected returns. *Journal of Business*, 61: 409-425.
- Cook, J. and Lewandowsky, S. (2016) Rational irrationality: Modeling climate change belief polarization using Bayesian networks. *Topics in Cognitive Science*, 8(1): 160-179.
- Copeland, T. (1976) A model of asset trading under the assumption of sequential information arrival. *Journal of Finance*, 31: 1149-1168.
- Corbet, S. and Gurdgiev, C. (2019) What the hack: systematic risk contagion from cyber events. *International Review of Financial Analysis*, 65.
- Covrig, V. and Ng, L. (2004) Volume autocorrelation, information, and investor trading. *Journal of Banking and Finance*, 28(9): 2155-2174.
- Da, Z. and Shive, S. (2018) Exchange traded funds and asset return correlations. *European Financial Management*, 24(1): 136-168.
- Da Dalt, C., Feldman, D., Garvey, G., and Westerholm, P.J. (2019) Contrarians or momentum chasers? Individual investors' behavior when trading exchange-traded funds. *Journal of Futures Markets*, 39(5): 553-578.
- Daniel, K. and Hirshleifer, D. (2015) Overconfident investors, predictable returns, and excessive trading. *Journal of Economic Perspectives*, 29(4): 61-88.

Daniel, K., Hirshleifer, D. and Subrahmanyam, A. (1998) Investor psychology and security market under- and overreactions. *Journal of Finance*, 53(6): 1839-1885.

Darrat, A.F., Rahman, S., and Zhong, M. (2003) Intraday trading volume and return volatility of the DJIA stocks: A note. *Journal of Banking and Finance*, 27(10): 2035-2043.

Darwish, M. (2012) Testing the contemporaneous and causal relationship between trading volume and return in the Palestine exchange. *International Journal of Economics and Finance*, 4(4): 182-192.

Davis, R.A., Zang, P., and Zheng, T. (2016) Sparse vector autoregressive modeling. *Journal of Computational and Graphical Statistics*, 25(4): 1077-1096.

De Grauwe, P. (2012) Booms and busts in economic activity: A behavioral explanation. *Journal of Economic Behavior and Organization*, 83(3): 484-501.

De Long, J.B., Shleifer, A., Summers, L.H., and Waldmann, R.J. (1990) Noise trader risk in financial markets. *Journal of Political Economy*, 98(4): 703-738.

Deakin, R.E. and Kildea, D.G. (1999) A note on standard deviation and RMS. *Australian surveyor*, 44(1): 74-79.

Deaves, R., Lei, J., and Schröder, M. (2019) Forecaster overconfidence and market survey performance. *Journal of Behavioral Finance*, 20(2): 173-194.

Deev, O. and Linnertová, D. (2014) The determinants of ETFs short selling activity. *Procedia - Social and Behavioral Sciences*, 109: 669-673.

Del Brio, E.B. Miguel, A., and Perote, J. (2002) An investigation of insider trading profits in the Spanish stock market. *The Quarterly Review of Economics and Finance*, 42(1): 73-94.

Demos, A. and Sentana, E. (1998) Testing for GARCH effects: a one-sided approach. *Journal of Econometrics*, 86(1): 97-127.

Dewachter, H., Iania, L., Lyrio, M. and de Sola P M. (2015) A macro-financial analysis of the euro area sovereign bond market. *Journal of Banking and Finance*, 50: 308-325.

Dickason, Z. and Ferreira, S. (2018) Establishing a link between risk tolerance, investor personality and behavioural finance in South Africa. *Cogent Economics and Finance*, 6(1); p. 1519898.

Divine, J. (2019) Has passive investing become fraught with risk? *U.S. News*, 14 November 2019, <https://money.usnews.com/investing/funds/articles/do-index-funds-etfs-quietly-pose-a-systemic-risk-michael-burry-thinks-so> [Accessed 14 March 2020].

Domowitz, I., Glen, J., and Madhavan, A. (2001) Liquidity, volatility and equity trading costs across countries and over time. *International Finance*, 4(2): 221-255.

Dovern, J. and van Roye, B. (2014) International transmission and business-cycle effects of financial stress. *Journal of Financial Stability*, 13: 1-17.

Dowie, G. and Willows, G. (2016) An investigation of investors' estimates of returns earned and the effect of anchoring on these estimations. *South African Journal of Accounting Research*, 30(1): 29-40.

Dreger, C. and Wolters, J. (2010) Investigating M3 money demand in the euro area. *Journal of International Money and Finance*, 29(1): 111-122.

Economou, F., Hassapis, C., and Philippas, N. (2018) Investors' fear and herding in the stock market. *Applied Economics*, 50(34): 3654-3663.

Elton, E.J., Gruber, M.J., and de Souza, A. (2019) Passive mutual funds and ETFs: performance and comparison. *Journal of Banking and Finance*, 106: 265-275.

Engle, R.F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 50(4): 987-1007.

Escobari, D., Garcia, S., and Mellado, C. (2017) Identifying bubbles in Latin American equity markets: Phillips-Perron-based tests and linkages. *Emerging Markets Review*, 33: 90-101.

- Etfgi (2020) ETFGI ETF/ETP growth charts. *etfgi*, <https://etfgi.com/> [Accessed 25 July 2020].
- EtfSA (2019) Exchange Traded Funds. *etfSA*, <http://www.etfsa.co.za/ETFs.htm> [Accessed 31 August 2019].
- Eviews (2019) User's Guide: Unit Root Testing. *Eviews*, 4 April 2019, http://www.eviews.com/help/helpintro.html#page/content%2Fadvtimeser-Unit_Root_Testing.html%23ww184965 [Accessed 6 June 2019].
- Fama, E.F. (1965) Random walks in stock market prices. *Financial Analysts Journal*, 21(5): 55-59.
- Fama, E.F. (1970) Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2): 383-417.
- Fama, E.F. (1991) Efficient capital markets: II. *The Journal of Finance*, 46(5): 1575-1617.
- Farinella, J. and Kubicki, R. (2018) The performance of exchange traded funds and mutual funds. *Journal of Accounting and Finance*, 18(4): 2158-3625.
- Farrar, D.E. and Glauber, R.R. (1967) Multicollinearity in regression analysis: the problem revisited. *The Review of Economic and Statistics*, 49(1): 92-107.
- Fellner, G. (2009) Illusion of control as a source of poor diversification: experimental evidence. *The Journal of Behavioral Finance*, 10(1): 55-67.
- Festinger, L., Riecken, H.W., and Schachter, S. (1956) When prophecy fails: A social and finance. *International Journal of Economics and Financial Issues*, 7(1): 448-459.
- Fink, F. and Schüler, Y.S. (2015) The transmission of US systemic financial stress: evidence for emerging market economies. *Journal of International Money and Finance*, 55: 6-26.
- Flood, C. (2020) Coronavirus puts an end to ETFs' decade-long spree. *Financial Times*, 2020, <https://www.ft.com/content/2f4c96b2-63c8-11ea-abcc-910c5b38d9ed> [Accessed 1 April 2020].

Frankfurter, G.M., McGoun, E.G., and Allen, D.E. (2004) The prescriptive turn in behavioral finance. *The Journal of Socio-Economics*, 33(4): 449-468.

Fremault, A. (1991) Stock index futures and index arbitrage in a rational expectations model. *Journal of Business*, 64(4): 523-547.

French, K.R., Schwert, G.W., and Stambaugh, R.F. (1987) Expected stock returns and volatility. *Journal of Financial Economics*, 19(1): 3-29.

Fuller, W.A. (1976) *Introduction to Statistical Time Series*. 1st ed. New York: Wiley.

García, D., Sangiorgi, F., and Urošević, B. (2007) Overconfidence and market efficiency with heterogeneous agents. *Economic Theory*, 30(2): 313-336.

Gastineau, G.L. (2001) An introduction to exchange-traded funds (ETFs). *Journal of Portfolio Management and Economics*, 27(3): 88-96.

Gerlach, S. (2011) ECB repo rate setting during the financial crisis. *Economics Letters*, 112(2): 186-188.

Gervais, S. and Odean, T. (2001) Learning to be overconfident. *The Review of Financial Studies*, 14(1): 1-27.

Gilbert, E. and Strugnell, D. (2010) Does survivorship bias really matter? An empirical investigation into its effects on the mean reversion of share returns on the JSE (1984–2007). *Investment Analysts Journal*, 39(72): 31-42.

Glaser, M. and Weber, M. (2007) Overconfidence and trading volume. *The Geneva Risk and Insurance Review*, 32(1): 1-36.

Glaser, M. and Weber, M. (2009) Which past returns affect trading volume?. *Journal of Financial Markets*, 12(1): 1-31.

Glosten, L., Nallareddy, S., and Zou, Y. (2016) ETF trading and informational efficiency of underlying securities. *Working paper*, Columbia University.

Godfrey, L.G. (1978) Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables. *Econometrica: Journal of the Econometric Society*, 46: 1303-1310.

Gorton, G. and Pennacchi, G. (1991) Security baskets and index-linked securities. *NBER Working paper*, No: 3711.

Granger, C.W. (1969) Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 37(3): 424-438.

Griffin, J.M., Nardari, F., and Stulz, R.M. (2007) Do investors trade more when stocks have performed well? Evidence from 46 countries. *The Review of Financial Studies*, 20(3).

Guedj, I., Li, G., and McCann, C. (2011) Futures-based commodities ETFs. *The Journal of Index Investing*, 2(1): 14-24.

Guney, Y. and Komba, G. (2016) Testing for the weak-form market efficiency of the Dar es Salaam stock exchange. *Handbook of Frontier Markets*, 1: 3-26.

Gupta, S., Goyal, V., Kalakbandi, V.K., and Basu, S. (2018) Overconfidence, trading volume and liquidity effect in Asia's giants: evidence from pre-, during-and post-global recession. *Decision*, 45(3): 235-257.

Gurdus, L. (2020) Volatility spikes have brought about 'a great victory for etfs as a structure,' CIO says. *CNBC*, 13 March 2020, <https://www.cnbc.com/2020/03/13/corona-virus-stock-market-etfs-holding-up-against-volatility-says-cio.html> [Accessed 1 April 2020].

Gürkaynak, R.S. (2008) Econometric tests of asset price bubbles: taking stock. *Journal of Economic Surveys*, 22(1): 166-186.

Gwilym, O.A and Sutcliffe, C. (2012) Problems encountered when using high frequency financial market data: suggested solutions. *Journal of Financial Management and Analysis*, 25(2): 15-28.

Halkos, G.E. and Tsirivis, A.S. (2019) Effective energy commodity risk management: econometric modeling of price volatility. *Economic Analysis and Policy*, 63: 234-250.

Hamm, S.J.W. (2014) The effect of ETFs on stock liquidity. *Working paper*, The Ohio State University.

Hannan, E.J. and Quinn, B.G. (1979) The determination of the order of an autoregression. *Journal of the Royal Statistical Society: Series B (Methodological)*, 41(2): 190-195.

Harvey, C.R. (1995) Predictable risk and returns in emerging markets. *The Review of Financial Studies*, 8(3): 773-816.

Hatemi, A. (2014) Asymmetric generalized impulse responses with an application in finance. *Economic Modelling*, 36: 18-22.

Hatemi, A. and Hacker, S.R. (2009) Can the LR test be helpful in choosing the optimal lag order in the VAR model when information criteria suggest different lag orders?. *Applied Economics*, 41(9): 1121-1125.

He, H. and Wang, J. (1995) Differential information and dynamic behavior of stock trading volume. *The Review of Financial Studies*, 8(4): 919-972.

Hill, J.M. (2016) The evolution and success of index strategies in ETFs. *Financial Analysts Journal*, 72(5): 8-13.

Horta, P., Lagoa, S., and Martins, L. (2016) Unveiling investor-induced channels of financial contagion in the 2008 financial crisis using copulas. *Quantitative Finance*, 16(4): 625-637.

Hsiao, C. (2005) Why panel data?. *The Singapore Economic Review*, 50(2): 143-154.

Hsiao, C. (2007) Panel data analysis—advantages and challenges. *Test*, 16(1): 1-22.

Huang, M.Y. and Lin, J.B. (2011) Do ETFs provide effective international diversification?. *Research in International Business and Finance*, 25(3): 335-344.

Hurlin, C. (2010) What would Nelson and Plosser find had they used panel unit root tests?. *Applied Economics*, 42(12): 1515-1531.

Ihantola, E.M. and Kihn, L.A. (2011) Threats to validity and reliability in mixed methods accounting research. *Qualitative Research in Accounting and Management*, 8(1): 39-58.

Im, K.S., Pesaran, M.H., and Shin, S. (1995) Testing for Unit Roots in Heterogeneous Panels. *Working paper*, University of Cambridge.

IOL (2020) Exchange Traded Funds are still popular trading instruments after 20 years. *IOL*, 17 July 2020, <https://www.iol.co.za/personal-finance/investments/exchange-traded-funds-are-still-popular-trading-instruments-after-20-years-51115810> [Accessed 25 July 2020].

Israeli, D., Lee, C.M., and Sridharan, S.A. (2017) Is there a dark side to exchange traded funds? An information perspective. *Review of Accounting Studies*, 22(3): 1048-1083.

Janssen, D.J., Füllbrunn, S., and Weitzel, U. (2019) Individual speculative behavior and overpricing in experimental asset markets. *Experimental Economics*, 22(3): 653-675.

Jarque, C.M. and Bera, A.K. (1980) Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3): 255-259.

Jiang, X. (2010) Return dispersion and expected returns. *Financial Markets and Portfolio Management*, 24(2): 107-135.

Jlassi, M., Naoui, K., and Mansour, W. (2014) Overconfidence behavior and dynamic market volatility: evidence from international data. *Procedia Economics and Finance*, 13: 128-142.

Johnson, W.F. (2009) Tracking errors of exchange traded funds. *Journal of Asset Management*, 10(4): 253-262.

Johnston, R., Jones, K., and Manley, D. (2018) Confounding and collinearity in regression analysis: a cautionary tale and an alternative procedure, illustrated by studies of British voting behaviour. *Quality and Quantity*, 52(4): 1957-1976.

Joo, B.A. and Durri, K. (2015) Comprehensive review of literature on behavioural finance. *Indian Journal of Commerce and Management Studies*, 6(2): 11-19.

Juselius, K. (2015) Haavelmo's probability approach and the cointegrated VAR. *Econometric Theory*, 31(2): 213-232.

Kahneman, D. and Tversky, A. (1979) Prospect theory: an analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 47(2): 263-292.

Kamesaka, A., Nofsinger, J.R., and Kawakita, H. (2003) Investment patterns and performance of investor groups in Japan. *Pacific-Basin Finance Journal*, 11(1): 1-22.

Kapoor, S. and Prosad, J.M. (2017) Behavioural finance: a review. *Procedia Computer Science*, 122: 50-54.

Karanasos, M., Paraskevopoulos, A.G., Ali, F.M., Karoglou, M., and Yfanti, S. (2014) Modelling stock volatilities during financial crises: a time varying coefficient approach. *Journal of Empirical Finance*, 29: 113-128.

Karpoff, J. (1987) The relation between price changes and trading volume: a survey. *Journal of Financial and Quantitative Analysis*, 22: 109-126.

Kilian, L. and Demiroglu, U. (2000) Residual-based tests for normality in autoregressions: asymptotic theory and simulation evidence. *Journal of Business and Economic Statistics*, 18(1): 40-50.

Kim, H. and Cho, M. (2019) The big short's Michael Burry sees a bubble in passive investing. *Bloomberg*, 28 August 2019, <https://www.bloomberg.com/news/articles/2019-08-28/the-big-short-s-michael-burry-sees-a-bubble-in-passive-investing> [Accessed 14 March 2020].

Kim, K.A. and Nofsinger, J.R. (2007) The behavior of Japanese individual investors during bull and bear markets. *The Journal of Behavioral Finance*, 8(3): 138-153.

Kim, T.H. and White, H. (2004) On more robust estimation of skewness and kurtosis. *Finance Research Letters*, 1(1): 56-73.

Kleine, J., Wagner, N., and Weller, T. (2016). Openness endangers your wealth: noise trading and the big five. *Finance Research Letters*, 16: 239-247.

Ko, K.J. and Huang, Z.J. (2007) Arrogance can be a virtue: overconfidence, information acquisition, and market efficiency. *Journal of Financial Economics*, 84(2): 529-560.

Kroner, K.F. and Lastrapes, W.D. (1993) The impact of exchange rate volatility on international trade: reduced form estimates using the GARCH-in-mean model. *Journal of International Money and Finance*, 12(3): 298-318.

Kyle, A.S. and Wang, F.A. (1997) Speculation duopoly with agreement to disagree: can overconfidence survive the market test?. *Journal of Finance*, 52(5): 2073-2090.

Kyrychenko, V. and Shum, P. (2009) Who holds foreign stocks and bonds? Characteristics of active investors in foreign securities. *Financial Services Review*, 18(1): 1-21.

Labuschagne, C.C., Oberholzer, N., and Venter, P.J. (2017a) Univariate and multivariate GARCH models applied to the CARBS indices. *International Conference on Applied Economics*. Cham: Springer, 69-83.

Labuschagne, C.C., Oberholzer, N., and Venter, P.J. (2017b) Univariate GARCH model generated volatility skews for the CIVETS stock indices. *Advances in Applied Economic Research*. Cham: Springer, 333-347.

Langer, E. (1975) The Illusion of Control. *Journal of Personality and Social Psychology*, 32(2): 311-328.

Ledoit, O. and Wolf, M. (2008) Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, 15(5): 850-859.

Lerskullawat, P. and Ungphakorn, T. (2018) Does overreaction still exist in Thailand?. *Kasetsart Journal of Social Sciences*, <https://doi.org/10.1016/j.kjss.2018.02.001>.

Lettau, M. and Madhavan, A. (2018) Exchange-traded funds 101 for economists. *Journal of Economic Perspectives*, 32(1): 135-54.

Levin, A. and Lin, C.F. (1992) Unit root tests in panel data: asymptotic and finite-sample properties. UC San Diego Working paper, No: 92-23.

Li, F.W. and Zhu, Q. (2018) Short Selling ETFs. *Finance Down Under 2019 Building on the Best from the Cellars of Finance*, <http://dx.doi.org/10.2139/ssrn.2836518>.

Li, Y. (2019) Michael Burry of ‘the big short’ says he has found the next market bubble. *CNBC*, 5 September 2019, <https://www.cnbc.com/2019/09/04/the-big-short-michael-burry-says-he-has-found-the-next-market-bubble.html> [Accessed 9 January 2020].

Liddle, B. and Lung, S. (2013) The long-run causal relationship between transport energy consumption and GDP: evidence from heterogeneous panel methods robust to cross-sectional dependence. *Economics Letters*, 121(3): 524-527.

Lin, C.Y., Rahman, H., and Yung, K. (2010) Investor overconfidence in REIT stock trading. *Journal of Real Estate Portfolio Management*, 16(1): 39-57.

Lin, W.L., Law, S.H., Ho, J.A., and Sambasivan, M. (2019) The causality direction of the corporate social responsibility - corporate financial performance nexus: application of panel vector autoregression approach. *The North American Journal of Economics and Finance*, 48: 401-418.

List, J.A. (2004) Neoclassical theory versus prospect theory: evidence from the marketplace. *Econometrica: Journal of the Econometric Society*, 72(2): 615-625.

Liu, S. (2015) Investor sentiment and stock market liquidity. *Journal of Behavioral Finance*, 16(1): 51-67.

Liu, W. and Strong, N. (2008) Biases in decomposing holding-period portfolio returns. *The Review of Financial Studies*, 21(5): 2243-2274.

Lo, A.W. and MacKinlay, A.C. (1990) An econometric analysis of nonsynchronous trading. *Journal of Econometrics*, 45: 181-211.

Lo, A.W., and Wang, J. (2000) Trading volume: definitions, data analysis, and implications of portfolio theory. *The Review of Financial Studies*, 13(2): 257-300.

Lowies, G.A., Hall, J.H., and Cloete, C.E. (2016) Heuristic-driven bias in property investment decision-making in South Africa. *Journal of Property Investment and Finance*, 34(1): 51-67.

- Lundbergh, S. and Teräsvirta, T. (2002) Evaluating GARCH models. *Journal of Econometrics*, 110(2): 417-435.
- Ma, J.Z., Ho, K.C., Yang, L., and Chu, C.C. (2018) Market sentiment and investor overreaction: evidence from New York listed Asian country Exchange Traded Funds. *Emerging Markets Finance and Trade*, 54(11): 2455-2471.
- Madhavan, A. (2014) Exchange-traded funds: an overview of institutions, trading, and impacts. *Annual Review of Financial Economics*, 6: 311-341.
- Madhavan, A. and Sobczyk, A. (2016) Price dynamics and liquidity of exchange-traded funds. *Journal of Investment Management*, 14(2): 1-17.
- Madura, J. and Ngo, T. (2008) Impact of ETF inception on the valuation and trading of component stocks. *Applied Financial Economics*, 18(12): 995-1007.
- Madura, J. and Richie, N. (2004) Overreaction of exchange-traded funds during the bubble of 1998-2002. *The Journal of Behavioral Finance*, 5(2): 91-104.
- Mak, M.K. and Ip, W.H. (2017) An exploratory study of investment behaviour of investors. *International Journal of Engineering Business Management*, 9: 1-12.
- Malkiel, B.G. (1973) *A random walk down wall street*. 6th ed. New York: W.W. Norton and Company.
- Mantalos, P. and Shukur, G. (2005) The effect of the GARCH (1,1) on autocorrelation tests in dynamic systems of equations. *Applied Economics*, 37(16): 1907-1913.
- Markowitz, H. (1952) Portfolio selection. *The Journal of Finance*, 7(1): 77-91.
- Masry, M. (2017) The impact of technical analysis on stock returns in an emerging capital markets (ECM's) country: theoretical and empirical study. *International Journal of Economics and Finance*, 9(3): 91-107.
- McKenzie, M.D. and Faff, R.W. (2003) The determinants of conditional autocorrelation in stock returns. *Journal of Financial Research*, 26(2): 259-274.

Meier, C. (2018) Aggregate investor confidence in the stock market. *Journal of Behavioral Finance*, 19(4): 421-433.

Merton, R.C. (1987) A simple model of capital market equilibrium with incomplete information. *The Journal of Finance*, 42(3): 483-510.

Metwally, A.H. and Darwish, O. (2015) Evidence of the overconfidence bias in the Egyptian stock market in different market states. *The Business and Management Review*, 6(4): 178-198.

Michailova, J. and Schmidt, U. (2016) Overconfidence and bubbles in experimental asset markets. *Journal of Behavioral Finance*, 17(3): 280-292.

Mohamad, A., Jaafar, A., and Goddard, J. (2016) Short selling and exchange-traded funds returns: evidence from the London stock exchange. *Applied Economics*, 48(2): 152-164.

Moore, D.A. and Healy, P. (2008) The trouble with overconfidence. *Psychological Review*, 115(2): 502-517.

Moore, D.A. and Schatz, D. (2017) The three faces of overconfidence. *Social and Personality Psychology Compass*, 11.

Moore, D.A., Carter, A.B., and Yang, H.H. (2015) Wide of the mark: evidence on the underlying causes of overprecision in judgment. *Organizational Behavior and Human Decision Processes*, 131: 110-120.

Müller, S. and Weber, M. (2010) Financial literacy and mutual fund investments: who buys actively managed funds?. *Schmalenbach Business Review*, 62(2): 126-153.

Mushinada, V.N.C. (2020) Are individual investors irrational or adaptive to market dynamics?. *Journal of Behavioral and Experimental Finance*, 25.

Mushinada, V.N.C. and Veluri, V.S.S. (2018) Investors overconfidence behaviour at Bombay stock exchange. *International Journal of Managerial Finance*, 14(5): 613-632.

Narayan, P.K., Narayan, S. and Prasad, A. (2008) Understanding the oil price-exchange rate nexus for the Fiji islands. *Energy Economics*, 30(5): 2686-2696.

Naseer, M. and Tariq, Y.B. (2015) The efficient market hypothesis: a critical review of the literature. *IUP Journal of Financial Risk Management*, 12(4): 48-63.

Naumenko, K. and Chystiakova, O. (2015) An empirical study on the differences between synthetic and physical ETFs. *International Journal of Economics and Finance*, 7(3): 24-35.

Nelson, D.B. (1991) Conditional heteroskedasticity in asset returns: a new approach. *Econometrica: Journal of the Econometric Society*, 59(2): 347-370.

Neves, M.E.D., Fernandes, C.M., and Martins, P.C. (2019) Are ETFs good vehicles for diversification? New evidence for critical investment periods. *Borsa Istanbul Review*, 19(2): 149-157.

Newey, W.K. and West, K.D. (1987) A simple-positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica: Journal of the Econometric Society*, 55: 703-708.

Ngene, G., Tah, K.A., and Darrat, A.F. (2017) The random-walk hypothesis revisited: new evidence on multiple structural breaks in emerging markets. *Macroeconomics and Finance in Emerging Market Economies*, 10(1): 88-106.

Ning, C., Xu, D., and Wirjanto, T.S. (2015) Is volatility clustering of asset returns asymmetric?. *Journal of Banking and Finance*, 52: 62-76.

O'Brien, R.M. (2017) Dropping highly collinear variables from a model: why it typically is not a good idea. *Social Science Quarterly*, 98(1): 360-375.

O'Hara, M. and Ye, M. (2011) Is market fragmentation harming market quality?. *Journal of Financial Economics*, 100(3): 459-474.

Odean, T. (1998) Volume, volatility, price, and profit when all traders are above average. *The Journal of Finance*, 53(6): 1887-1934.

Odean, T. (1999) Do investors trade too much?. *American Economic Review*, 89(5): 1279-1298.

Padungsaksawasdi, C. and Daigler, R.T. (2014) The return-implied volatility relation for commodity ETFs. *Journal of Futures Markets*, 34(3): 261-281.

Pak, T.Y. and Chatterjee, S. (2016) Aging, overconfidence, and portfolio choice. *Journal of Behavioral and Experimental Finance*, 12: 112-122.

Paul, A. and Theodore, P. (2006) Calendar anomalies in an emerging African market: evidence from the Ghana stock exchange. *Journal of Economics*, 52(12): 58-96.

Pawley, M.G. (2006) The impact of survivorship bias on South African unit trust performance: 1972–2004. *Investment Analysts Journal*, 35(64): 21-26.

Petajisto, A. (2017) Inefficiencies in the pricing of exchange-traded funds. *Financial Analysts Journal*, 73(1): 24-54.

Pezzo, S.P., Pezzo, M.V., and Stone, E.R. (2006) The social implications of planning: how public predictions bias future plans. *Journal of Experimental Social Psychology*, 42(2): 221-227.

Phillips, P.C. and Perron, P. (1988) Testing for a unit root in time series regression. *Biometrika*, 75(2): 335-346.

Phillips, P.C. and Yu, J. (2011) Dating the timeline of financial bubbles during the subprime crisis. *Quantitative Economics*, 2(3): 455-491.

Pikulina, E., Renneboog, L., and Tobler, P.N. (2017) Overconfidence and investment: an experimental approach. *Journal of Corporate Finance*, 43: 175-192.

Prosad, J.M., Kapoor, S., Sengupta, J., and Roychoudhary, S. (2017) Overconfidence and disposition effect in Indian equity market: an empirical evidence. *Global Business Review*, 19(5): 1303-1321.

Rahman, M.M. and Shahbaz, M. (2013) Do imports and foreign capital inflows lead economic growth? Cointegration and causality analysis in Pakistan. *South Asia Economic Journal*, 14(1): 59-81.

- Rajdev, A.A. and Raninga, A.M. (2016) Gender and heuristic driven biases: A review of literature. *International Journal of Commerce, Business and Management*, 5(3): 35-38.
- Ramiah, V. and Davidson, S. (2007) Information-adjusted noise model: evidence of inefficiency on the Australian stock market. *Journal of Behavioral Finance*, 8(4): 209-224.
- Raut, R.K., Das, N., and Kumar, R. (2018) Extending the theory of planned behaviour: impact of past behavioural biases on the investment decision of Indian investors. *Asian Journal of Business and Accounting*, 11(1): 265-291.
- Redl, C. (2018) Macroeconomic uncertainty in South Africa. *South African Journal of Economics*, 86(3): 361-380.
- Robinson, J. (2005) Stock price behaviour in emerging markets: tests for weak form market efficiency on the Jamaica stock exchange. *Social and Economic Studies*, 54(2): 51-69.
- Rodgers, J.L and Nicewander, W.A. (1988) Thirteen ways to look at the correlation coefficient. *The American Statistician*, 42(1): 59-66.
- Rompotis, G.G. (2009) Interfamily competition on index tracking: The case of the vanguard ETFs and index funds. *Journal of Asset Management*, 10(4): 263-278.
- Rompotis, G.G. (2013) Actively vs. passively managed exchange traded funds. *Aestimatio: The IEB International Journal of Finance*, 6: 116-135.
- Rompotis, G.G. (2014) On the trading behavior of emerging market ETFs. *The Journal of Trading*, 10(1): 56-86.
- Rubinstein, M. (1975) Securities market efficiency in an Arrow-Debreu economy. *The American Economic Review*, 65(5): 812-824.
- Ruppert, D. (1987) What is kurtosis? An influence function approach. *The American Statistician*, 41(1): 1-5.

Salahuddin, M., Vink, N., Ralph, N., and Gow, J. (2019) Globalisation, poverty and corruption: retarding progress in South Africa. *Development Southern Africa*.

Sale, J.E., Lohfeld, L.H., and Brazil, K. (2002) Revisiting the quantitative-qualitative debate: implications for mixed-methods research. *Quality and Quantity*, 36(1): 43-53.

Samouilhan, N.L. (2006) The relationship between international equity market behaviour and the JSE. *South African Journal of Economics*, 74(2): 248-260.

Scheinkman, J.A. and Xiong, W. (2003) Overconfidence and speculative bubbles. *Journal of Political Economy*, 111(6): 1183-1220.

Schmidt, U. and Zank, H. (2012) A genuine foundation for prospect theory. *Journal of Risk and Uncertainty*, 45(2): 97-113.

Schwarz, G. (1978) Estimating the dimension of a model. *The Annals of Statistics*, 6(2): 461-464.

Sehgal, S. and Garg, V. (2016) Cross-sectional volatility and stock returns: evidence for emerging markets. *The Journal for Decision Makers*, 41(3): 234–246.

Sentana, E. and Wadhwani, S. (1992) Feedback traders and stock return autocorrelations: evidence from a century of daily data. *The Economic Journal*, 102(411): 415-425.

Serletis, A. (2007) *Quantitative and empirical analysis of energy markets*. New Jersey: World Scientific.

Shah, S.Z.A., Ahmad, M., and Mahmood, F. (2018) Heuristic biases in investment decision-making and perceived market efficiency. *Qualitative Research in Financial Markets*, 10(1): 85-110.

Shefrin, H. (2000) *Beyond greed and fear: understanding behavioral finance and the psychology of investing*. Cambridge: Harvard Business School Press.

Shefrin, H. and Statman, M. (1985) The disposition to sell winners too early and ride losers too long: theory and evidence. *Journal of Finance*, 40: 777-790.

Shefrin, H. and Statman, M. (1994) Behavioral capital asset pricing theory. *Journal of Financial and Quantitative Analysis*, 29(3): 323-349.

Shefrin, H. and Statman, M. (2000) Behavioral portfolio theory. *Journal of Financial and Quantitative Analysis*, 35(2): 127-151.

Sheikh, M.F. and Riaz, K. (2012) Overconfidence bias, trading volume and returns volatility: evidence from Pakistan. *World Applied Science Journal*, 18(12): 1737-1748.

Shepperd, J.A., Klein, W.M., Waters, E.A., and Weinstein, N.D. (2013) Taking stock of unrealistic optimism. *Perspectives on Psychological Science*, 8(4): 395-411.

Shieh, G. (2010) On the misconception of multicollinearity in detection of moderating effects: multicollinearity is not always detrimental. *Multivariate Behavioral Research*, 45(3): 483-507.

Simo-Kengne, B.D., Balcilar, M., Gupta, R., Reid, M., and Aye, G.C. (2013) Is the relationship between monetary policy and house prices asymmetric across bull and bear markets in South Africa? Evidence from a markov-switching vector autoregressive model. *Economic Modelling*, 32: 161-171.

Simon, H.A. (1955) A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1): 99-118.

Sims, C.A. (1980) Macroeconomics and reality. *Econometrica: Journal of the Econometric Society*, 48(1): 1-48.

Sloof, R. and von Siemens, F.A. (2017) Illusion of control and the pursuit of authority. *Experimental Economics*, 20(3): 556-573.

Smith, A. (1759) *The theory of moral sentiments*. London: A. Millar.

Staer, A. (2017) Fund flows and underlying returns: the case of ETFs. *International Journal of Business*, 22(4): 276-304.

Statman, M., Thorley, S., and Vorkink, K. (2006) Investor overconfidence and trading volume. *The Review of Financial Studies*, 19(4): 1531-1565.

Steyn, J. (2019) The tracking efficiency of exchange-traded funds listed on the Johannesburg stock exchange. *Journal of the Southern African Institute for Management Scientists*, 28(4): 2-14.

Strydom, B., Charteris, A., and McCullough, K. (2015) The relative tracking ability of South African exchange traded funds and index funds. *Investment Analysts Journal*, 44(2): 117-133.

Subrahmanyam, A. (1991) Risk aversion, market liquidity, and price efficiency. *The Review of Financial Studies*, 4(3): 417-441.

Swanson, N.R. and Granger, C.W. (1997) Impulse response functions based on a causal approach to residual orthogonalization in vector autoregressions. *Journal of the American Statistical Association*, 92(437): 357-367.

Tapa, A. and Hussin, M. (2016) The relationship between stock return and trading volume in Malaysian ACE market. *International Journal of Economics and Financial Issues*, 6(7S): 271-278.

Taylor, S.E. and Brown, J.D. (1988) Illusion and well-being: a social psychological perspective on mental health. *Psychological Bulletin*, 103(2): 193-210.

Thanasegaran, G. (2009) Reliability and Validity Issues in Research. *Integration and Dissemination*, 4: 35-40.

Thompson, C.G., Kim, R.S., Aloe, A.M., and Becker, B.J. (2017) Extracting the variance inflation factor and other multicollinearity diagnostics from typical regression results. *Basic and Applied Social Psychology*, 39(2): 81-90.

Tokic, D. (2019) The passive investment bubble. *Journal of Corporate Accounting and Finance*, 31(1): 7-11.

Tran, V.L. and Leirvik, T. (2019) A simple but powerful measure of market efficiency. *Finance Research Letters*, 29: 141-151.

Trehan, B. and Sinha, A.K. (2011) A study of existence of overconfidence biases among investors and its impact on investment decision. *Asia Pacific Journal*, 11: 1-15.

Trejos, C., van Deemen, A., Rodríguez, Y.E., and Gómez, J.M. (2019) Overconfidence and disposition effect in the stock market: A micro world based setting. *Journal of Behavioral and Experimental Finance*, 21: 61-69.

Trichet, J.C. (2010) State of the union: the financial crisis and the ECB's response between 2007 and 2009. *Journal of Common Market Studies*, 48: 7-19.

Trinugroho, I. and Sembel, R. (2011) Overconfidence and excessive trading behavior: an experimental study. *International Journal of Business and Management*, 6(7): 147-152.

Tsaur, R.C. (2013) Fuzzy portfolio model with different investor risk attitudes. *European Journal of Operational Research*, 227(2): 385-390.

Tsay, R. (2005) *Analysis of financial time series*. USA: John Wiley and Sons.

Tversky, A. and Kahneman, D. (1991) Loss aversion in riskless choice: a reference-dependent model. *The Quarterly Journal of Economics*, 106(4): 1039-1061.

Venkataraman, R. and Venkatesan, T. (2016) Evaluation of growth of mutual funds and exchange traded funds in India. *SDMIMD Journal of Management*, 7(1): 41-47.

Verousis, T. and Voukelatos, N. (2018) Cross-sectional dispersion and expected returns. *Quantitative Finance*, 18(5): 813-826.

Vidal-Tomás, D. and Ibañez, A. (2018) Semi-strong efficiency of bitcoin. *Finance Research Letters*, 27: 259-265.

Virigineni, M. and Rao, M.B. (2017) Contemporary developments in behavioural finance. *International Journal of Economics and Financial Issues*, 7(1): 448-459.

Volkwyn, B. (2019) Is the huge growth of ETFs a risk to investors?. *IOL*, 25 September 2019, <https://www.iol.co.za/personal-finance/investments/is-the-huge-growth-of-etfs-a-risk-to-investors-33527662> [Accessed 9 January 2020].

Watanapalachaikul, S. and Islam, S.M. (2007) Rational speculative bubbles in the Thai stock market: econometric tests and implications. *Review of Pacific Basin Financial Markets and Policies*, 10(1): 1-13.

Weinstein, N.D. (1980) Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, 39(5): 806-820.

Weisberg, S. (1983) Discussion: some principles for regression diagnostics and influence analysis. *Technometrics*, 25(3): 240-244.

West, R.R. (1975) On the difference between internal and external market efficiency. *Financial Analysts Journal*, 31(6): 30-34.

White, H. (1980) A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 48(4): 817-838.

Willows, G. and West, D. (2015) Differential investment performance in South Africa based on gender. *International Business and Economics Research Journal*, 14(1): 221-236.

Wojcik, S.P. and Ditto, P.H. (2014) Motivated happiness: self-enhancement inflates self-reported subjective well-being. *Social Psychological and Personality Science*, 5(7): 825-834.

Wolde-Rufael, Y. (2007) Another look at the relationship between telecommunications investment and economic activity in the United States. *International Economic Journal*, 21(2): 199-205.

Wolfe, C.R., Reyna, V.F., and Smith, R.J. (2018) On judgments of approximately equal. *Journal of Behavioral Decision Making*, 31(1): 151-163.

Wooldridge, J.M. (2019) Correlated random effects models with unbalanced panels. *Journal of Econometrics*, 211(1): 137-150.

Yeh, C.H. and Yang, C.Y. (2011) Examining the effects of traders' overconfidence on market behavior. *Agent-Based Approaches in Economic and Social Complex Systems VI*, pp. 19-31. Tokyo: Springer.

Yeh, C.H. and Yang, C.Y. (2014) How does overconfidence affect asset pricing, volatility, and volume?. *Advances in Computational Social Science*, pp. 107-122. Tokyo: Springer.

Yu, C.F. (2014) CEO overconfidence and overinvestment under product market competition. *Managerial and Decision Economics*, 35(8): 574-579.

Yung, K. and Liu, Y.C. (2009) Implications of futures trading volume: hedgers versus speculators. *Journal of Asset Management*, 10(5): 318-337.

Zaiane, S. (2013) Overconfidence, trading volume and the disposition effect: evidence from the Shenzhen stock market of China. *Issues in Business Management and Economics*, 1(7): 163-175

Zaiane, S. and Abaoub, E. (2009) Investor overconfidence and trading volume: the case of an emergent market. *International Review of Business Research Papers*, 5(2): 213-222.

Zell, E. and Alicke, M.D. (2011) Age and the better-than-average effect. *Journal of Applied Social Psychology*, 41(5): 1175-1188.

Zia, L., Sindhu, L.M., and Hashmi, H.S. (2017) Testing overconfidence bias in Pakistani stock market. *Cogent Economics and Finance*, 5(1).