Cash Flow as a Predictor of Share Returns: Evidence from the Johannesburg Stock Exchange

Sergei Alexandroi

ALXSER001

Prepared under the supervision of Professor Francois Toerien and presented to the Department of Finance and Tax at the University of Cape Town in partial fulfilment of the requirements for the award of the degree of

MASTER OF COMMERCE *Specialising in* FINANCE

(*in the field of* Investment Management)

April 2019

University of Cape Town

Abstract

The existence of so-called equity market anomalies suggests that factors outside of the traditional asset-pricing models can model share returns. Despite this, there is limited empirical evidence on cash flow metrics as anomalies, and less so on cash flows as a predictor of share returns. The aim of this study is to provide a new insight into the South African equity market by investigating and comparing the extent of return predictability displayed by cash and accrual measures. This research extends the work of Foerster, Tsagarelis and Wang (2017) and investigates previously untested cash-based measures on an untested sample of shares in an emerging market. Fixed effects panel regression models are applied to a dataset consisting of 85 shares listed on the Johannesburg Stock Exchange over the period 2008 to 2018, using cash and accounting variables to test for predictive ability on six-month ahead total share returns. In contrast to the findings by Foerster, Tsagarelis and Wang (2017), the results suggest that accrual-based measures provide more explanatory power for share return variation than cash flow measures. However, using these variables for purposes of earning consistent excess returns requires further investigation. In addition, the strongest regression model consists of both bottom-line earnings and cash flow variables, suggesting that there is predictive power in a combination of traditional profitability and cash flow figures. The value of using such cash flow information in the fundamental investment process has practical implications on asset-pricing, the presence of anomalies in financial markets as well as return prediction. Underlying this research is also an inherent test of the level of market efficiency on the JSE. The resulting significance levels suggest that some variation in future returns can be explained by prior movements in company financial figures, which contributes to the understanding of how South African equity markets process and reflect financial data. The study therefore provides evidence to reject a strong-form level of market efficiency and support the argument for a semi-strong form level of market efficiency on the JSE.

Plagiarism Declaration

I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.

I have used the Harvard-UCT convention for citation and referencing. Each contribution to, and quotation in, this dissertation from the work(s) of other people has been attributed, and has been cited and referenced.

This dissertation is entirely my own work.

I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

Name: Sergei Alexandroi

Signature: Signed by candidate

Date: 20/04/2019

Acknowledgements

I would like to acknowledge Dr. Francois Toerien for his effort, support and guidance throughout this process. His expertise in the field was invaluable to my progress and the significant time allocated to supervising my research is greatly appreciated.

My family and friends are the ones who understood the greater motivations for this dissertation. I would like to express my gratitude for their understanding, patience and encouragement throughout. I share this entire undertaking with them.

Finally, I would like to acknowledge the University of Cape Town and every staff member at the Department of Finance and Tax for providing the educational environment and opportunities to all alumni.

Contents

List of Tables

List of Equations

Chapter 1: Introduction and Background

Fundamental characteristics of firms are derived from figures presented in their company financial statements. Accounting bodies outline methods to calculate, adjust and present this information to stakeholders and the investing public. These procedures serve as legal guidance on the preparation of financial statements, which are to be audited if the firm is listed on a securities exchange. The earliest accounting practises began on a purely cash-basis, but as globalisation and technology spurred trading activity, obligations and ownership interests required an accrual-based approach.

The modern accounting process requires considerable estimation, erratic recognition, valuation and interpretation of accounting standards. These issues compound when a business runs complicated activities in several jurisdictions under an intricate ownership structure. Despite common knowledge of accounting manipulation, the collapse of institutions that were 'too big to fail' and frequent charges of fraud that audit firms face, traditional accrual accounting figures still play a dominant role in financial markets. This brings into question the efficacy of relying on this information to value shares and predict returns.

Investors rely on these accrual-based figures to assess the intrinsic value of an asset and to predict the cross section of average returns. Market efficiency theories describe the degree to which financial markets absorb such information and price assets accordingly. A market that is highly efficient reflects all information quickly, making it impossible to actively select a portfolio of assets that will outperform the market. In an informationally efficient market, the Capital Asset Pricing Model (CAPM) and the Arbitrage Pricing Theory (APT) are models that reflect that assets are fairly priced. However, there has been considerable empirical research suggesting that markets are not entirely efficient – especially in emerging markets with relatively new financial systems (Lee, Lee and Lee, 2010). These deviations from the expected behaviour of share prices are known as anomalies. Such anomalies bear a significant predictive relationship with asset returns, and indicate that either the market may not be efficient, or that the asset pricing models fail to capture some risk factor. Inefficiencies create an opportunity for investors to earn abnormal returns on the market – justifying the role that active asset management plays. While various anomalies have been documented – with contradictory evidence – limited attention has been given to the cash flow component of company financial statements. Given the amount of discretion in accounting practises, the cash flows of an entity are far less susceptible to manipulation. Investors may easily be misled in this regard – a failing entity is able to reflect an accounting profit on their financial statements. Therefore, this study aims to determine whether the 'cleaner' cash flow aspect of a firm can be used to predict share returns. In addition, this research investigates whether various components of cash flows differ in their predictive ability, and whether these are superior to traditional profitability measures used in practice.

At this point, there has been limited research on the predictive ability of cash flows specifically. To the author's knowledge, this has only been tested in a developed market (the USA). There appears to be little, if any, academic research around this topic in the developing world. In South Africa, cash flow-to-price is the only related measure considered in literature, and is shown to be a significant predictor of excess returns on the JSE over the period 1985 to 2011 (Muller and Ward, 2013). Given the controversy in the audit industry, investigating a fundamental component that is devoid of manipulation is considered relevant and necessary. If financial theory implies that the price of an asset is determined by the present value of its cash flows, it is worthwhile to investigate whether these cash flows contain information about the future price behaviour of the asset.

Section 1.1 presents the motivation behind this study. Section 1.2 discusses the research objectives and contribution to academic literature. Section 1.3 outlines the organisation of this dissertation.

1.1 Motivation for the Study

Cash flow figures are mainly used to assess solvency, liquidity and value investment opportunities, but generally not to attempt to predict future equity returns. Investigating the relationship between cash flows and returns is a worthwhile task, because it directly addresses the issue of whether accounting figures provide value relevant information to investors. Inherent in this research is the comparison between accruals and cash flows. The accrual method of accounting records revenue and expenses when they are earned or incurred, regardless of whether cash is received or paid. As these are forward-looking transactions, there are often later adjustments on the amounts, tax considerations and write-offs. On the other hand, the cash flow figures provide an idea of how a firm generates value from primary and secondary business activities.

Listed South African entities are required to present business activities and undergo audit under the relevant accounting requirements – the International Financial Reporting Standards (IFRS). IFRS requires mandatory presentation of the Statement of Cash Flows. However, it is at the discretion of the manager and the nature of the business as to how the 'Cash Flow from Operating Activities' is obtained. The figure can be calculated via the direct or indirect method. Most entities choose the indirect method because it reconciles movements in asset and liability accounts to profit, arriving at a cash flow figure. The direct method provides a clearer summary of the core business activities and the extent to which these generate cash flow. If the core operations of a business provide healthy cash flow movements, the market should price such a competitive advantage accordingly.

Empirical evidence of anomalies in financial markets suggests that the two most common asset pricing models used in equity markets, the Capital Asset Pricing Model (CAPM) (Sharpe, 1964; Lintner, 1965; Mossin, 1966) and the Arbitrage Pricing Theory (APT) (Ross, 1976), may not fully capture risk factors. In the South African context, Rensburg and Robertson (2003a; 2003b) and Auret and Sinclaire (2006), have found that such anomalies exist on the Johannesburg Stock Exchange. However, there is an apparent lack of research investigating anomalies based on cash flows in its different forms. This is surprising given that Kruger (2011), when examining JSE anomalies over unstable market crisis periods, finds cash-flow-to-price to be the only variable that remains a statistically significant predictor of share returns.

At this point, there is an absence of empirical evidence on the predictive ability of cash flows on the JSE. Foerster, Tsagarelis and Wang (2017) conduct a similar study in a US GAAP environment, but this is yet to be tested on an exchange operating under IFRS, such as South Africa. This study is therefore intended to provide new insights into this topic for an emerging market with a relatively sophisticated financial system. Inherent is also a test of information efficiency of the South African equity market (the Johannesburg Stock Exchange or JSE). This research intends to provide an understanding of factors impacting share returns and consider whether cash flows provide incremental information (beyond profitability figures) on price behaviour.

1.2 Contribution and Objectives

This study aims to provide a new insight into the South African equity market, specifically by testing for the existence of a cash flow anomaly. Building on the work of Novy-Marx (2013), Ball et al. (2015), Foerster, Tsagarelis and Wang (2017) and Hou, Karolyi, and Kho (2011), it further aims to provide a novel contribution to accounting and finance research in emerging markets. Investigating whether cash flows can be used to predict future share performance has both theoretical and practical implications. Testing various cash flow measures on their predictive value for earning abnormal returns has consequences on the market efficiency of the JSE. It also contributes to existing literature on anomalies, or lack thereof, supporting the semi-strong form efficiency argument. Furthermore, the research contributes to the debate of cash vs. accrual information content, usefulness and explanatory power.

Given the surprising lack of literature around this topic, this study aims to explore the value of using cash flow information in the fundamental investment process by providing clarity on the relationship between a listed company's cash flows and its equity return behaviour. If markets overvalue earnings and bottom-line figures, exploiting mispriced assets from a cash flow perspective offers an advantage. The existence of such an anomaly has implications for active portfolio management, asset allocation and risk management. Therefore, the objectives of this study are to:

1. Investigate whether cash flows can predict future share returns for companies on the Johannesburg Stock Exchange, and to what extent investors can use this information to generate profits.

2. Consider whether different components of cash flows, such as cash flows from operating activities, cash flows from financing activities, and cash flows from investing activities, differ in their ability to predict returns.

3. Compare calculated cash flow measures to traditional profitability and free cash flow measures commonly used to model future share returns.

4. Test the level of market efficiency on the Johannesburg Stock Exchange; if cash flows have explanatory power or predictive ability, the market is either inefficient or there is some misspecification of asset-pricing models.

5. Enhance the existing international and South African empirical literature by testing previously untested cash-based measures on an untested sample of 85 shares in an emerging market.

1.3 Thesis Organisation

The remainder of the thesis will be structured as follows. Chapter 2 discusses the theories underlying market efficiency, asset pricing and accounting principles. Chapter 3 reviews empirical evidence of return predictability and examines literature on cash flows. Chapter 4 discusses the data required for this study, as well as necessary adjustments and considerations made to mitigate bias. Chapter 5 outlines the methodologies employed in this study to reach robust and accurate results. Chapter 6 presents and discusses the empirical results and Chapter 7 suggests extensions for future research, summarises the thesis and concludes.

Chapter 2: Theoretical Background

Before empirically investigating the predictive ability of cash flows, it is important to consider and contextualise the foundations of portfolio theory and investment management. Two major themes in modern finance are information efficiency and asset pricing theory. These areas have undergone considerable scrutiny over many years – academic research has had to evolve with the growing complexity of financial markets and often reconsider prior findings. As such, financial reporting standards have had to do the same. Fundamental to investment finance is the valuation of securities, and many asset pricing models aim to accurately assess and evaluate investment opportunities.

To achieve this, these models require certain assumptions of how efficiently asset prices reflect information. As to the profit-motive, the logical extension of an asset pricing model is its predictive ability. Deviations from market efficiency or a pricing model suggests the presence of an 'anomaly' or a 'style' factor. Often, these anomalies are based upon fundamental variables derived from figures found in companies' financial statements. Investigating the presence and extent of return predictability requires an understanding of market efficiency, asset pricing theories, as well as applicable financial reporting principles. This Chapter reviews the development of these theories.

Section 2.1 covers aspects of efficiency in financial markets, Section 2.2 discusses theories of asset pricing, Section 2.3 discusses relevant accounting literature with regards to accruals and cash flows, and Section 2.4 presents a summary of the chapter and concludes.

2.1 Market Efficiency

The notion that stock market movements reflect not only the past, but both the present and future, was first proposed by Bachelier (1900). This was later reviewed and developed into the efficient market hypothesis (EMH), where an 'efficient market' is one whose prices 'fully reflect' all available information relevant to itself and its constituents (Fama, 1970). The result is that a stock's current price is the best estimate of its intrinsic value and is equal to the present value of the certainty equivalent of its future cash flows (Brown, Harlow and Tinic, 1988). The random walk model, an associated concept conditional upon the EMH, suggests that the price changes are random and independent of preceding movements (Fama, 1965b). This is due to the day's new information being incorporated into the price, and given that news is inherently unpredictable, the resulting movement follows a 'random walk' with a certain drift.

The random walk model, an associated concept conditional upon the EMH, suggests that the price changes are random and independent of preceding movements (Fama, 1965b). This is due to the day's new information being incorporated into the price, and given that news is inherently unpredictable, the resulting movement follows a 'random walk' with a certain drift. The implication is that historical prices cannot be used to predict future behaviour. Fama (1970) went on to further describe three subsets of information efficiency, each with its own implications.

The weak-form efficiency of the EMH suggests that the current market price reflects all information contained in historical price movements, rendering technical analysis unreliable. Semi-strong form efficiency suggests that in addition to historical price data, current prices reflect all publicly available information. Therefore, fundamental analysis provides no predictive power for a stock's future behaviour. Strong-form efficiency, a theoretical benchmark of a fully efficient state, suggests that current market prices reflect all relevant information, whether publicly available or privately held. This implies that investors trading on proprietary information cannot consistently earn abnormal returns.

Fama (1970) explains that it is easy to determine sufficient conditions for a market in which a security fully reflects all available information. That is, a market in which (i) there are no transaction costs associated with trading, (ii) all information is available to all market participants without cost, and (iii) all market participants agree on the implications of information on current and future prices. Of course, these conditions are not descriptive of those in practice, so the extreme views on efficiency are unjustified (Fama, 1991). There is enough empirical evidence on the existence of anomalies for investors to continue the search for mispriced opportunities. Grossman and Stiglitz (1980) point out that information is not costless, and rational investors will have an incentive to uncover this information only if it is likely to generate higher investment returns. It is also clear that the degree of efficiency varies across markets and assets. For example, less sophisticated financial markets, like those in emerging economies, have less scrutiny and lower volume of trades. The same can be said about smaller capitalisation shares. Instead of asking whether markets are efficient, it is perhaps better to investigate the *extent* of efficiency in a market.

2.2 The Theory of Asset Pricing

Asset pricing theory is concerned with determining the fundamental value of assets which generate an uncertain stream of cash flows. It aims to consider risks, explain prices and find appropriate levels of return for investment opportunities. To determine this, asset pricing models have varying degrees of assumptions of the market or security in question, with market efficiency often a key consideration. This section is therefore closely linked with the EMH discussed in Section 2.1.

Academic research has demonstrated that many asset pricing theories fail to sufficiently explain asset prices. This resulted in various modifications of the pricing models, from the underlying assumptions, to the dynamic or static state of the model. It is also important to note that any test of an asset pricing model faces the joint-hypothesis problem[1]. Therefore, this section discusses the theories that are most appropriate to understanding the empirical evidence reported later in this study. Frameworks behind the Capital Asset Pricing Model (CAPM), the Arbitrage Pricing Theory (APT), the Fama and French Three-Factor Model, the Carhart Four-Factor Model and the Fama and French Five-Factor Model are reviewed to provide context.

2.2.1 The Capital Asset Pricing Model (CAPM)

The underlying principles of the CAPM are attributed to earlier works in Modern Portfolio Theory by Markowitz (1952) and in the separation theorem by Tobin (1958). Prior research derived expected returns as a function of the source and cost of financing that asset, and while there was an understanding of the value of diversification, Markowitz (1952) was the first to provide a mathematical association between risk and return in the context of portfolio formation. Markowitz argues that the characteristics of a portfolio of assets will differ from the characteristics of the individual assets within the portfolio. Given that risk is measured by the variance of an asset's expected returns, Markowitz demonstrates that an asset's contribution to the overall portfolio risk profile is more important than the individual risk characteristics of the asset. This emphasises the importance of

[1] To test for market efficiency, an asset pricing model requires the market portfolio to be represented by a proxy based on the assumptions of that particular model. If the model is found to price assets incorrectly, it is uncertain whether this is due to a market inefficiency, an erroneous asset pricing model or both. This is called the joint-hypothesis problem.

understanding the inter-relationships between securities and taking both individual risk and covariances into account.

Furthermore, investment decisions require an evaluation of the risk-return properties of a portfolio. Markowitz (1952) models this behaviour on the basis that investors will choose 'mean-variance efficient' portfolios. That is, investors prefer to minimize the variance of the portfolio return given a level of expected return, and prefer to maximize the expected return, given a level of portfolio variance. This provides a mathematical condition for including assets in portfolios, producing a frontier of mean-variance efficient portfolios. The Markowitz Efficient Frontier is a hyperbolic representation of these optimal portfolio choices.

Tobin (1958) added to the portfolio selection model by introducing the borrowing and lending of a risk-free asset, which better represents the range of liquidity preferences and investment choices available to investors. Tobin's separation theorem shows that the investment process can be broken down into two stages: first, identifying the optimal combination of risky assets for a portfolio, and second, a separate decision on the allocation of funds to that risky portfolio and a single risk-free asset. If all funds are invested in a risk-free asset, the resultant portfolio would generate the risk-free rate of return, no variance and zero correlation with all risky assets. Tobin reasoned that investors facing the same universe of assets would hold the same optimal portfolio on the efficient frontier (with the highest return per unit of risk) and cater for individual risk and liquidity preferences by borrowing or lending the risk-free asset.

While the literature by Markowitz (1952) and Tobin (1958) was significant in the formation of asset pricing theory, certain aspects needed further development. Markowitz did not propose a model for expected asset returns that could be used to generate the efficient frontier. Moreover, in the context of large portfolios of assets, the calculation of the covariances between every asset pair in the portfolio was an impractical undertaking.

Sharpe (1964) and Lintner (1965) proposed the CAPM as a theoretical model of equilibrium expected returns on risky assets, addressing the prevailing shortcomings in portfolio theory at the time. The CAPM retains the following assumptions put forward by Markowitz (1952):

- All investors are 'Markowitz efficient' investors who, after including all investable assets in their approximation of the efficient frontier, will be rational mean-variance optimisers;
- All investors have a single and identical holding period;
- It is assumed that there are no taxation or transaction costs applied to traded assets;
- Investors are all price-takers;
- There is no inflation nor interest rate changes, or it is fully anticipated;
- It is assumed that capital markets are priced appropriately to their risk levels i.e. capital markets are in equilibrium

Sharpe (1964) and Lintner (1965) add two further assumptions to the Markowitz model:

- Investors are all able to borrow or lend any amount of money at a fixed, risk-free rate of return;
- All investors estimate identical probability distributions of future cash flows from the available asset universe, i.e. investors have homogenous expectations

If the portfolios of all individual investors are aggregated, all borrowing and lending of the risk-free asset will equate, and the value of this aggregate portfolio will be the aggregate wealth of the economy. That is, if investors with homogenous expectations use the same Markowitz analysis on an identical universe of assets over a matching holding period, they must all arrive at the same risky asset (Sharpe, 1964). This is the Market Portfolio, M. Under the premise that investors can and will borrow at a risk-free rate to leverage their portfolios as per Tobin (1958), M is the single best portfolio available to all investors. It is therefore the tangent point of the efficient frontier and the Capital Market Line (CML). The CML reflects the expected rates of returns for all combinations of optimal risky portfolios and the risk-free asset's return. Thus, the efficient frontier is simplified – as expected return and risk for a portfolio are linear combinations, the graph of possible returns and risks becomes a linear combination of the return on the Market Portfolio, M, and return on the risk-free asset.

All portfolios chosen by rational investors will lie on the CML, which depicts the expected return of any efficient portfolio as a function of its risk: $E(r_P) = r_f + \frac{\sigma_P}{\sigma_m}\left[E(r_m) - r_f\right]$ (2.1) where $E(r_P)$ is the expected return on a portfolio *(P)* of risky assets, r_f is the risk-free rate of return, σ_P is the risk (standard deviation) of the risky portfolio, σ_m is the risk (standard deviation) of the market portfolio *(M)* and $E(r_m)$ is the expected return on the market portfolio *(M)*.

Sharpe (1964) further demonstrates the price adjustment process when an individual asset deviates from the Market Portfolio, M, resolving an existing limitation of asset pricing theory. If an asset's risk-return profile differs from that of M, market forces will adjust the price until it is included in the optimal portfolio. This relationship highlights that the risk of a portfolio can be understood as the extent to which asset and market returns move together. Therefore, portfolio risk is the covariance of an asset with M, as opposed to the covariance amongst all combinations of assets in that portfolio. The beta of an asset, *i,* is proportional to the covariance of the asset's returns to the variance of the market's return:

$$\beta_i = \frac{Cov\ (r_i, r_m)}{\sigma_m^2} \tag{2.2}$$

where:

$$Cov\ (r_i, r_m) = \rho_{i,m}\sigma_i\sigma_m \tag{2.3}$$

Sharpe (1964) places emphasis on an individual asset and its relationship to the market; the dynamics of this relationship are established by Lintner (1965), Mossin (1966) and Black (1972). One such advancement is the understanding of how diversification affects and limits risk exposures. Sharpe (1964) suggests that the Market Portfolio (M) is the highest possibility portfolio on the CML, and contains all risky assets in proportion to their relative market value. It follows that if all assets are included, an investor can no longer diversify risk associated with the individual assets (unsystematic or specific risk). What remains is therefore the risk associated with M (systematic or non-specific risk) which cannot be diversified away – a result of macroeconomic risk factors.

The beta coefficient in Equation 2.2 is a numerical representation of the exposure to such systematic risk, affecting expected returns and, in turn, risk premiums. A central premise of CAPM is that an efficiently diversified portfolio must contain only systematic risk, and on average, investors are not compensated for taking on any unsystematic risk. A market in equilibrium results in all assets and portfolios lying on a graphical representation of the CAPM, the Security Market Line (SML).

In contrast to the CML, the SML depicts the linear relationship between the expected return of an *individual* asset and its market risk exposure, measured by beta:

$$E(r_i) = r_f + \beta_i\left[E(r_m) - r_f\right] \tag{2.4}$$

Equations 2.2 to 2.4 above demonstrate that the expected return for an asset or portfolio of assets is a function of the risk-free rate and its covariance with the market portfolio. The latter represents a major implication of CAPM – that beta is the only risk factor which explains expected returns (Fama and French, 1996). Furthermore, Fama and French (1996) argue that a positive relation between beta and expected return can be support for the CAPM if beta is indeed the *only* factor explaining expected returns.

While a simple and elegant framework for asset pricing and investor behaviour, the CAPM has various flaws in its assumptions, application and reliability. In practice, the prevailing inflation rate impacts the short-term, highly liquid risk-free asset proposed in the model, creating uncertainty in terms of a real rate of return for investors. Considering that markets will adjust any disequilibrium, expected rates of return and therefore betas will not be stable over time. Here, investors face a limitation in their estimation of beta. Since only historical price data is available, investors do not have the necessary information required to properly estimate beta, and will have an incomplete outlook on their future risk. More importantly, CAPM assumes that risk in financial markets is adequately represented by the variance of returns. This stems from the assumption of mean-variance optimisers having quadratic utility functions and homogenous expectations. Given that variance has upside potential, the risk is better represented by an asymmetric function – the probability of losing. Furthermore, the assumption of asset prices having a normal distribution has been contested, with evidence suggesting that a leptokurtic distribution is better suited.

Finally, the CAPM is conditional upon a market portfolio consisting of all available assets traded in the market. This extends beyond the equity index, as it also includes all privately held assets. As a result, the true market portfolio is both unobservable and unmeasurable, becoming problematic to employ in practice. Market indices are therefore often used as proxies to represent this portfolio. This gives rise to the joint-hypothesis problem. Any test of market efficiency, such as the one implied by CAPM, requires an asset pricing model generating expected returns to compare to real returns. When a proxy is used to determine both the prices of assets and to test the CAPM, a mispricing is difficult to attribute to either. Consequently, the CAPM becomes impractical.

2.2.2 The Arbitrage Pricing Theory (APT)

The CAPM is derived from a single-factor model of return, with the underlying assumption that since every investor, on average, owns the market portfolio, the market is efficient and maintains equilibrium. The model is simple but ignores the possibility of systematic risk arising from multiple common factors – the notion that assets with similar fundamental or industry characteristics tend to exhibit similar behaviours. Ross (1976) proposes a linear model linking risk to expected returns, whereby markets will adjust temporarily mispriced assets back to their fair market value.

Under the APT, the stochastic process underlying the generation of asset returns over time can be simplified in the form of a linear k-factor model:

$$R_i = E(R_i) + \sum_{k=1}^{K} b_{ik} f_k + \varepsilon_i \tag{2.5}$$

where R_{it} is the realised return earned by the asset i, $E(R_i)$ is the expected rate of return for asset i, f_k is the kth common risk factor that impacts asset i's returns with $E(f_k)$ = 0, b_{ik} is a factor loading coefficient for asset i on factor k and ε_i is the random error term representing the unsystematic component of returns with $E(\varepsilon_i)$ = 0. ε_i is also assumed to be uncorrelated across assets and uncorrelated with the risk factors.

Instead of specifying an unobservable market portfolio, it assumes that each investor holds a unique portfolio of assets with an array of betas specific to their asset composition. These capture the sensitivity of returns to changes in related macroeconomic variables, and any shocks to these factors would cause expected asset returns to change. In this manner, assets can be priced correctly based on the expected return derived from the model, using the sum of all future cash flows discounted at the APT rate.

The APT is far less restrictive in its assumptions of the market. Unlike the CAPM, it does not rely on mean-variance optimising behaviour, normally distributed returns, or an unobservable market portfolio. Ross (1976) puts forward three assumptions underlying the APT:

(i) Asset returns can be explained by a systematic factor model;
(ii) There are sufficient assets available for investors to diversify away unsystematic risk;
(iii) Markets do not allow for well-diversified portfolios to have persistent arbitrage opportunities.

The APT states that asset returns follow a factor structure. Applying assumption (ii) to Equation 2.5 above results in the random error component, ε_i, being diversified away for large portfolios. Like the CAPM, the APT assumes that portfolios are sufficiently diversified and the resulting contribution to total portfolio risk from unsystematic risk is zero. Expected returns on an asset can now be represented as a linear function of the asset's sensitivity and risk premia associated with the k risk factors:

$$E(R_i) = r_f + \sum_{k=1}^{K} b_{ik}\lambda_k \qquad (2.6)$$

where $E(R_i)$ is the expected rate of return for asset i, r_f is the return on the risk-free asset, b_{ik} is a factor loading coefficient for asset i on factor k and λ_k is the risk premium of factor k (expected rate of return on kth factor in excess of the risk-free rate of return).

A crucial pricing relationship proposed by Ross (1976) is that market equilibrium will not allow for arbitrage opportunities. An arbitrage opportunity exists when a mispricing occurs, and investors can exploit and profit from the mispriced asset(s) without any additional investment or risk. This gives rise to the law of one price, a related economic concept whereby two securities, commodities or assets that are identical in every respect should have the same price. Any violation of the law will result in arbitrageurs trading the asset where profit opportunity exists until it is exploited away. This mechanism ensures that assets are priced fairly and markets will be in equilibrium.

Ross (1976) asserts that expected returns have a linear relationship to multiple, unknown systematic factors underlying the asset or portfolio in question. The APT model does not make any assumptions about investor expectations or risk preference. Consequently, it does not specify, identify or number the risk factors. This is further complicated by the fact that relevant risk factors are likely to change over time.

2.2.3 The Fama and French Three-Factor Model

The CAPM offers a powerful and intuitive way to measure risk and understand the relationship between risk and expected return, but the model's empirical record raises concern over its validity. Numerous studies have shown that additional risk factors (other than the CAPM beta) provide explanatory power for share returns. Surprisingly, some of these factors had no prior role in asset pricing theory. Market capitalisation price-earnings ratio and leverage are some examples of factors shown to explain the cross section of expected returns (Basu, 1977, 1983; Bhandari, 1988). There is evidence that a company's book value of equity to market value of equity ratio (BE/ME) has a positive relationship with average returns (Rosenberg, Reid and Lanstein, 1985; Chan, Hamao and Lakonishok, 1991). Lakonishok, Sheifer and Vishy (1994) find that in addition to BE/ME, cash flow-to-price also has predictive power on average returns.

Fama and French (1992) present a key argument against the CAPM by specifying firm characteristics that proxy for exposure to systematic risk. Expanding the CAPM, the Three-Factor Model has two additional risk factors, and is one of the most prominent multifactor models in finance. Fama and French argue that smaller companies are more sensitive to macroeconomic changes and that companies with high BE/ME are more likely to experience financial distress. Thus, the model adds a size proxy, SMB (Small Minus Big), and a BE/ME proxy, HML (High Minus Low), as additional risk factors. The equation for the Fama and French (1993) Three-Factor Model is:

$$R_p - r_f = \alpha_i + \beta_{iM} R_M + \beta_{iSMB} SMB + \beta_{iHML} HML + \varepsilon_i \tag{2.7}$$

where R_p is the realised return on a portfolio *(P)* of risky assets, r_f is the risk-free rate of return, α_i is the abnormal return and intercept, R_M is the realised return on the market, SMB is the return on a portfolio of small shares in excess of the return on a portfolio of large shares, HML is the return on a portfolio of shares with a high book-to-market ratio in excess of the return on a portfolio of shares with a low book-to-market ratio, $\beta_{iM}, \beta_{iSMB}, \beta_{iHML}$ are the sensitivities associated with the corresponding risk factors and ε_i is the estimation error.

Historical observations suggest that average returns on shares of small firms and high BE/ME are in fact above that predicted by the CAPM (Fama and French, 1993). The model claims that all market returns can be explained by these new factors and that it captures risk premiums, therefore invalidating the CAPM. Several researchers have criticised the Three-Factor Model, suggesting that the results are influenced by survivorship bias in the COMPUSTAT database originally used to derive it, as well as a data snooping effect in portfolio construction (Kothari, Shanken and Sloan, 1995; Black, 1993; MacKinlay, 1995). Basiewicz and Auret (2010) investigate the Three-Factor Model's feasibility in a South African context. The model is shown to capture a substantial amount of time-series variation with low pricing errors, which supports its use in expected return estimation for companies listed on the Johannesburg Stock Exchange (Basiewicz and Auret, 2010).

2.2.4 The Carhart Four-Factor Model and the Fama and French Five-Factor Model

Carhart (1997) extends the Fama and French Three-Factor Model by including a momentum factor into the asset pricing model. Momentum captures the tendency of a share price to maintain its current trend into the future. Using data on surviving funds as well as those which have disappeared, Carhart compares the performance of the Four-Factor Model to the CAPM and the Three-Factor Model while addressing the survivorship bias issue in the US stock market. In context of evaluating mutual fund performance, Carhart finds that the momentum effect explains some of what was previously attributed to the fund alpha. The additional factor is commonly used in present day asset management and fund evaluation.

Two additional factors have recently been added to the Fama and French Three-Factor Model. Fama and French (2015) extend their previous model to include a profitability factor, representing the relationship between operating profitability and the performance of stocks, and investment, representing the relationship between asset growth and the performance of stocks. This Five-Factor Model has been criticised for its exclusion of a momentum factor, the lack of robustness of the two additional factors and poor performance at the global level.

2.3 Accounting Literature

This section covers relevant accounting frameworks and literature, specifically that of International Financial Reporting Standards (IFRS) and discusses the principles of cash and accrual accounting methodologies.

Prior to the 18th century, cash basis was the main system of accounting (Winjum, 1972). As trade routes were discovered and commerce expanded into different geographical locations the complexity of business required accounting for agency costs, credit agreements and capital formation through different periods. Thus, the purpose of preparing accounts shifted to the protection of corporate investors, income determination, and dividend payments (Chatfield, 1974).

2.3.1 International Financial Reporting Standards (IFRS)

The objective of financial reporting is to provide information about the reporting entity that is useful to existing and potential investors, lenders and other creditors making decisions about providing resources to the entity. This includes buying, selling or holding equity and debt instruments, as well as providing or settling forms of credit. These decisions are based on expected returns, and expected returns are based on the assessment of future net cash inflows to the entity.

Financial Statements are prepared to provide information about the financial position, performance and cash flows of an entity that is useful to the users. The Conceptual Framework for Financial Reporting states that information is considered useful if it meets the following qualitative characteristics:

Fundamental characteristics:

- Relevance – possesses confirmative and predictive power and can influence decision making of users;
- Faithful representation – complete, neutral and free from errors and omissions

Enhancing characteristics:

- Understandable – to persons with a reasonable knowledge of business;
- Comparable – entity to entity and period to period;

- Timely – information provided before it loses potential to influence decisions;
- Verifiable – through auditing procedures and analysis

The financial statements of listed South African companies are subject to auditing under the relevant accounting requirements – the International Financial Reporting Standards (IFRS). International Accounting Standard (IAS) 1.27 requires that an entity prepares its financial statements (except for the cash flow statement) using the accrual basis of accounting. These statements are constructed, calculated and presented considering the accrual principle in different forms. A separate standard (IAS 7) details the presentation of the Statement of Cash Flows. However, it is at the discretion of the manager and the nature of the business as to how certain figures are determined. "For operating cash flows, the direct method of presentation is encouraged, but the indirect method is acceptable" (IAS 7.18). The direct method (IAS 7.19) shows each major class of cash receipts (from customers) and cash payments (to suppliers and employees) while the indirect method (IAS 7.20) adjusts the net profit or loss figure for non-cash transactions, such as depreciation and amortisation.

A logical observation is that the direct method aggregates cash flows with similar characteristics and provides valuable insight into the primary cash-generating activities of an entity. It therefore seems reasonable to investigate whether cash flow figures influence share prices, and to what extent. Bernard and Stober (1989) argue that cash flows should not be preferred to accruals on average, and find no difference between the implications in terms of stock price behaviour. Another suggestion is that the price reactions (to accrual or cash flow data) are too contextual to be modelled parsimoniously (Bernard and Stober, 1989). Still, there is value in researching the two fundamental accounting concepts, cash and accrual, and their relationship with financial market price behaviour.

2.4 Summary and Conclusion

This chapter provides a theoretical background and context to the research question. Market efficiency, asset pricing theories, as well as relevant financial reporting principles are reviewed.

The EMH is based on the degree to which the market price of a security reflects information relevant to it. Each form of market efficiency has implications for the behaviour, predictability and trading of a security. Foundational asset pricing theories are closely linked to market efficiency, investigating a variety of market factors that influence price and attempting to model it accordingly. Because of this link, tests of pricing models are implicitly tests of market efficiency – the joint-hypothesis problem. Therefore, this study considers conclusions made in terms of market efficiency or asset-pricing within the joint-hypothesis context.

The first asset pricing model was built on principles of mean-variance analysis and portfolio selection, making assumptions about investors' response when faced with risk and return options. The CAPM assumes that all investors hold some combination of the market portfolio and a risk-free asset, depending on their individual levels of risk-aversion. The APT attempts to address some of the shortcomings of the CAPM, assuming instead that each investor holds a unique portfolio with betas specific to their unique range of assets. The Fama and French (1992) Three-Factor Model adds size and value risk-factors based on empirically motivated observations that these proxies can explain returns. The commonly-used momentum factor is included in the Carhart (1997) Four-Factor Model, while the Fama and French (2015) Five-Factor Model requires more evidence to substantiate the use of five factors to price assets. As financial reporting standards prescribe how firms calculate and present financial statements, it is important to understand the principles underlying the various statements, line items and ratios. Central to this is the distinction between cash and accruals.

Finally, evidence of style anomalies imply that investors can forecast returns of the security in question. This is of course inconsistent with market efficiency theories and asset pricing models. The following chapter reviews key arguments and discusses the current status of research on anomalies, predictability of returns and cash flow variables.

Chapter 3: Literature Review

This chapter provides a critical analysis of relevant empirical findings in both the international and South African contexts, with a focus on the latter. Literature on style anomalies, return predictability and cash flows is reviewed in detail. The empirical evidence on these topics is substantial and contradictory arguments are often put forward; this chapter is intended to present the current status of research in a balanced manner, and to consider the theory and methodology underlying each conclusion.

While these anomalies have been documented, limited attention has been given to the cash flow component of company financial statements, especially within a South African context. More specifically, the predictive ability of cash flow has not, to the author's knowledge, been tested in such an environment. The lack of empirical evidence and consensus on this research question reveals the significance of this study.

Section 3.1 reviews the empirical findings on anomalies, Section 3.2 considers research of cash flows in both accounting and investment contexts, and Section 3.3 provides a summary and concludes.

3.1 Anomalies

As discussed in Section 2.5, style anomalies suggest that investors can forecast returns of the security in question, and their existence is inconsistent with efficiency theories and asset pricing models. Nevertheless, there is extensive evidence of style effects in both the international and South African financial markets. For this reason, Section 3.1 analyses the findings of the most relevant empirical literature.

3.1.1 International Anomalies

Ball (1978) suggests that price-earnings (P/E) ratios, among other variables, are useful proxies for expected stock returns. Prior to that, traditional investment strategies were roughly based on a form of the value effect (Graham and Dodd, 1940). Basu (1977) suggests that portfolios comprising stocks with low price-earnings (P/E) ratios outperform portfolios with stocks of high P/E ratios. Interestingly, this effect still holds when returns are adjusted for portfolio beta. It is argued that using CAPM to benchmark performance will always produce this result, as firms with greater risk will have lower prices (and P/E ratios), and therefore a greater expected return. Reinganum (1981) extends this research and finds evidence in support of the effect. Similar studies have found a significant and positive association between returns and the book-to-price (B/P) ratio in U.S. markets (Rosenberg, Reid and Lanstein, 1985; DeBondt and Thaler, 1985). Later, Campbell and Shiller (1988) find that earnings yield, the inverse of the P/E ratio, is able to predict market returns.

The 'size effect' is defined as the negative relationship between equity returns and the market value of the common equity of a security. Banz (1981) and Reinganum (1981) find that the size coefficient has more explanatory power than beta in the U.S. market. As discussed in Section 2.3.4, Fama and French (1992) find robust evidence of a similar relationship between size and the book-to-market ratio of stock returns. In fact, Fama and French demonstrate that the size and book-to-market effects subsume the effects of leverage, P/E ratios and earnings. Later research by Reinganum (1982), Blume and Stambaugh (1983) and Keim and Stambaugh (1986) establishes that the small firm effect occurs predominantly in January and is termed the 'January Effect'.

Fama and French (1988) and Litzenberger and Ramaswamy (1979) show that a large dividend yield (measured as dividend/price) results in large returns on the aggregate stock market. Chan, Hamao and Lakonishok (1991) find that a high ratio of cash flow-to-price is a predictor of higher returns. A similar effect is seen with the book-to-market ratio. Fama and French (1993, 1996) later adjust the returns obtained from the Three-Factor Model, and find that book-to-market, size, earnings yield, dividend yield and cash flow-to-price are no longer significant anomalies.

Technological progress has created a global financial system that is becoming increasingly interconnected. This has motivated a growing number of academic studies on the impact of global market and macroeconomic factors on the pricing of local securities. A long-standing question in international asset pricing is whether securities prices are determined by a single, cohesive and globally-integrated market, or by the characteristics of their segmented local markets. Hou, Karolyi and Kho (2011) conduct a global study of more than 27000 stocks across 49 countries to determine which firm-level characteristics have the greatest explanatory power for both time-series and cross-sectional variation in stock returns. The sample period spans from 1981 to 2003 and monthly return data is evaluated against earnings yield, dividend yield, cash flow-to-price, leverage, book-to-market equity, size and momentum. Factor-mimicking portfolios are constructed based on cash flow-to price and momentum, which are subsequently included in multifactor models. Hou, Karolyi and Kho (2011) show that there is strong evidence to suggest that a three-factor model consisting of cash flow-to-price, momentum and global market factors captures time series variation in global stock returns. This composition also produces the lowest pricing error and rejection rate of all the global multifactor models considered.

3.1.2 South African Anomalies

Evidence of asset pricing anomalies in the South African market has largely corresponded with those in international markets. In the South African context, De Villiers et al. (1986) challenge traditional pricing models with the discovery of a 'small firm' effect on the JSE. Later, Page and Palmer (1993) study a sample of 164 companies over the 1978 to 1988 period and fail to find convincing evidence of a size effect. However, these researchers do find the presence of a significant earnings (value) effect in relation to excess returns. Page (1996) supplements this finding with evidence that the earnings effect persists across various multifactor benchmark. It was common practise to omit smaller, illiquid stocks from research which resulted in generally smaller sample sizes – a major criticism of the earlier South African literature.

Van Rensburg and Slaney (1997) argue against the use of CAPM, suggesting that using the JSE All-Share Index as a market proxy results in the risk of mining shares being inflated. Van Rensburg (1999) supports van Rensburg and Slaney (1997), and confirms two sources of variation in returns on the JSE – the JSE Industrial Index and the JSE All-Gold Index. Both van Rensburg and van Rensburg and Slaney propose that a two-factor APT model using these proxies has the ability to explain how assets are priced on the JSE.

To overcome previous shortcomings in South African research, van Rensburg (2001) uses monthly dividend-adjusted share return data from 1983 to 1999 on JSE industrial shares and considers factors that have yet to be investigated on the JSE. Extending the work of Page and Palmer (1993), van Rensburg finds evidence of a size effect present on the JSE. Significant value effects for earnings yield and dividend yield are also uncovered, confirming the price-earnings effect. This gives suggests that the APT model provides a stronger representation of risk (for industrial shares) if expanded to include three factors: value, size and momentum. Fraser and Page (2000) investigate similar value and momentum strategies on industrial firms from 1973 to 1997 and find strong, independent predictive power of both strategies on one-month ahead returns on the JSE. These research outcomes illustrate either some form of market inefficiency, or model misspecification in the South African equity markets.

Further research into the size and value style characteristics is conducted by van Rensburg and Robertson (2003a), who employ a multifactor model to investigate the cross section of returns on the JSE over the 1990 to 2000 sample period. Given the prevailing problem of research excluding thinly traded shares, the study uses the largest sample size for asset pricing research on the JSE to date – an average of 336 shares per month. This more adequately represents smaller-cap shares when investigating a size effect and reduces the possibility of a 'dilution' effect on the power of statistical tests. Market capitalisation is used to represent the size effect while price-to-earnings represents the value effect, which is then compared to the traditional CAPM beta. Van Rensburg and Robertson find that price-to-net asset value, dividend yield, price-to-earnings ratio, cash flow-to-price and size function as significant explanatory variables of returns. A one-way portfolio sort shows that a small size and lower price-earnings firm earns higher returns on average, supporting evidence previously documented by van Rensburg (2001). Two-way portfolio sorts confirm that these effects operate independently of one another. Notably, van Rensburg and Robertson uncover for the first time an inverse relationship between beta and returns on the JSE.

Despite the intention of the study to improve upon previous shortcomings, it received various criticisms. Firstly, van Rensburg and Robertson (2003a) only investigate ten years of data, which they concede is not sufficient to make substantive arguments. Then, the inclusion of thinly traded shares caused the equally weighted portfolios to have a strong bias towards these smaller capitalisation companies, impacting the statistical results. This leads to the authors conceding that the negative relationship between beta and returns may be a result of including thinly traded shares in the sample.

To address the problem of a thin trading bias, Strugnell, Gilbert and Kruger (2011) extend the methodology used by van Rensburg and Robertson (2003a) and investigate a sample from 1994 to 2007 – considering a variety of holding periods – in conjunction with controls for thin trading (Scholes and Williams, 1977; Dimson, 1979). Strugnell, Gilbert and Kruger find evidence of both size and value anomalies on the JSE, as well as evidence in support of the relationship between beta and return proposed by van Rensburg and Robertson. Ward and Muller (2012) consider a sample period from 1986 to 2011 and provide additional justification against the use of a single-beta CAPM. Similar to van Rensburg (2001; 2002), these results suggest that the use of a CAPM model on the JSE, at least when the All-Share Index is used as a market proxy, is inadequate. Again, evidence suggests a market inefficiency or model misspecification on the JSE.

Another anomaly is discovered by Auret and Sinclaire (2006), who extend the work of van Rensburg and Robertson (2003a) using monthly returns on the JSE from 1990 to 2000. The five most significant factors are selected from the study by van Rensburg and Robertson, as well as adding a previously omitted book-to-market measure. Auret and Sinclaire find a significant positive relationship between share returns and the book-to-market ratio. Surprisingly, when added to a regression based on the van Rensburg and Robertson (2003) model, book-to-market subsumes the effect of both size and earnings factors, though not statistically significant at the 5% level. This may be due to correlations with other variables that include price in their calculations. Therefore, the size and earnings factors – which have low correlations with each other – are better suited to explain returns (Auret and Sinclaire, 2006).

Applying a more practical investment approach, Basiewicz and Auret (2009) investigate broader company characteristics and adjust for thin trading and transaction costs, and perform independent (as opposed to sequential) portfolio sorts. The research sample includes all firms listed on the JSE from 1989 to 2005. The most significant methodological improvements are the specific restrictions placed on price and liquidity – reflecting the kind of restrictions a portfolio manager would face in practice. Firstly, Basiewicz and Auret demonstrate that the earnings effect is the weakest, while book-to-market is the strongest, contrary to prior findings of Auret and Sinclaire (2006). Furthermore, and again contrary to Auret and Sinclaire, book-to-market is shown *not* to subsume the size effect. In fact, both the book-to-market and size effect have significant and independent ability to predict returns. The value premia for all the factors are significantly lower than any prior literature, which illustrates the impact and importance of considering liquidity and price constraints.

More recent findings by Hoffman (2012) indicate that the size effect is the greatest in micro-cap shares, while book-to-market and momentum effects are the most consistent across all portfolios. Muller and Ward (2013) conduct a comprehensive study of style anomalies on the JSE from 1985 to 2011 and find that momentum, liquidity, earnings yield, dividend yield, cash-flow-to-price, price-to-book, return on equity, return on capital and interest cover to all have significant and persistent ability to generate excess returns. Unlike previous research, there was no size effect present in the analysis. Kruger, MacDonald and Toerien (2014) find that size, book-to-market, price-to-earnings and momentum are consistent over time. However, the statistical significance of these factors depends on the sample period. Following from Kruger (2011), and upon analysing these factors over a ‘market crisis’ period, cash flow-to-price is revealed to be the only stable measure. This highlights the significance of the cash flow metric in a South African context – previously found to be significant in the global study by Hou, Karolyi and Kho (2011).

3.2 Cash Flow

Empirical research around the usefulness of cash flow data for security pricing has produced inconsistent results, which can be attributed to either the range of cash flow definitions used, or the inadequacy of models employed. The following sections discuss how the accounting and financial research has progressed, and along with it the importance of distinguishing between cash and accrual components of financial figures. In addition, evidence on the predictive ability of various traditional and cash flow measures is critically reviewed.

3.2.1 Earnings, Cash and Accruals

Ball and Brown (1968) are the first to empirically investigate the link between accounting figures and the value a market places on such information. This was an established principle in modern valuation theory, but the methodology used formed the basis for numerous future research. Earnings (defined as net income excluding extraordinary items) is shown to be a strong predictor of the cross-section of average returns in the US market. Subsequent research indicates that earnings add little incremental information over the size and book-to-market factors as found by Fama and French (1996, 2008). Ball and Brown (1968) later add to their research by investigating cash flow variables. Prior to the stagflation experienced by industrialised countries in the 1970s, accounting standards did not require mandatory disclosure of cash flows in company financial statements. As a result, academic researchers used proxies as estimates for cash flows – for example, Ball and Brown (1968) use earnings plus depreciation and amortisation. A regression model is conducted with the additional cash flow variable but is not as successful in predicting stock return residuals.

Beaver and Dukes (1972) model security returns with extensions of the earnings variable, using current earnings, earnings before deferrals and cash flows (approximated using earnings before deferrals plus depreciation and amortisation). The study is built on the work of Ball and Brown (1968) and motivated by the scrutiny over discretionary accounting estimates prescribed by Generally Accepted Accounting Principles (GAAP). However, the cash flow measure is the least significant in a test of association with the 'abnormal performance index', leading to the conclusion that it is not a significant factor for pricing stocks.

Wilson (1986) finds that total accruals and cash flows (measured by working capital from operations) have incremental information content beyond earnings, but the study is inconclusive as to whether they have incremental informational content beyond each other. Rayburn (1986) extends this study using an operating cash flow variable as well as a 20-year window and finds that cash flows and accruals do not have differential associations with returns in the U.S. Bowen, Burgstahler and Daley (1987) find evidence of a differential association using unexpected cash flows. However, these results are (1) sensitive to outliers; and (2) only significant for two years out of the ten-year sample period (Bowen et al., 1987:744).

Refining the prior research methodology, Wilson (1987) uses daily abnormal returns (as opposed to monthly data) and measures earnings announcements to the Wall Street Journal and the market's response to new information about cash and non-cash (accruals) earnings components. Evidence of association between earnings returns means that at least one of the components has some information content. The results suggest that when total accruals and cash from operations are taken together, they have incremental information content beyond earnings. A positive association between this combined measure and stock returns is also uncovered. Furthermore, Wilson finds that the total accruals component of earnings has incremental information content beyond the cash component and proposes a hypothesis: for a given amount of earnings, the market reacts more favourably the larger (smaller) are cash flows (current accruals) (Wilson, 1987).

Motivated by the hypothesis set out by Wilson (1987), Bernard and Stober (1989) extend the sample period and find evidence to the contrary. An initial hypothesis is that the relative impact of cash flows and accruals on security returns may vary depending on the state of the economy, rendering Wilson's claim circumstantially valid. However, after considering 'more contextual' macroeconomic valuation models, the decomposition of earnings into cash flow and accrual components provides no incremental information beyond earnings. In support of the findings of Rayburn (1986), Bernard and Stober conclude that cash flows should not be preferred to accruals on average. Board and Day (1989) find conflicting and complicated evidence on the relationship between earnings, cash flows and returns.

There are numerous empirical studies examining the incremental information content of cash flows in terms of earnings (Wilson, 1986; Bowen, Burgstahler and Daley, 1987) and in terms of accruals (Rayburn, 1986). This research largely finds significant incremental information content of cash flow figures. More recently, Bernard and Stober (1989) show that disaggregating net income into cash from operations and accruals does not provide additional information content beyond net income.

These studies focus on a single aspect of cash flows: cash from operations (working capital from operations is often used as a proxy). This research considers the components of cash flows as required by applicable accounting standards – cash flows from financing, investing and operating activities. Livnat and Zarowin (1990) find that separating net income into cash from operations and accruals does not provide additional association with returns beyond that of net income alone. This supports the results of Bernard and Stober (1989), who find no significant difference between the implications of cash flows and accruals, as reflected in stock price behaviour, implying that investors should not prefer cash flow to accrual numbers. However, disaggregating operating and financing cash flows into their cash and accrual components does improve differential association with annual stock returns. Consistent with traditional finance theories concerning effects of financing, investing and operating transactions, Livnat and Zarowin (1990) find that:

- cash inflows (outflows) from operations are positively (negatively) associated with returns;
- dividend payments are positively associated with returns;
- debt issuance is positively associated with returns;
- common stock issuance is positively associated with returns;

- preferred stock issuance is negatively associated with returns.

Of course, these effects are generalised. For example, the above does not address scaling of variables and in the case of common stock issuance, Bhandari (1988) demonstrates a positive relationship to the ratio of debt to equity i.e. if stock is issued, equity will increase, the debt to equity ratio will decrease, which reduces returns. Nevertheless, this supports the principle and purpose of the accounting standard in its function to signal to investors any business transactions it undertakes, and that the market prices these transactions in. Livnat and Zarowin (1990) propose that there is incremental information content available from disaggregating net income into accruals and components of cash flows from financing, investing, and operating activities. Thus, the results suggest that company financial statements in the U.S. contain more information than 'bottom-line' earnings figures.

Testing for information content of cash flow and accrual components of earnings is the main motivation behind research conducted by Wilson (1986; 1987), Rayburn (1986) and Bernard and Stober (1989). However, Dechow (1994) takes a more direct approach to evaluate whether reported earnings have greater explanatory power than realised cash flows. Previous academic literature ascribed abnormal returns to a market pricing mechanism, whereby investors re-evaluate their prior expectations of future cash flows – as both earnings and cash flow variables are adjusted to reach consensus, the returns fluctuate. Importantly, this shift brought new light to how market agents disseminate information. Actual earnings (excluding discontinued operations and certain extraordinary items) are used, as it was suggested that they would explain abnormal returns better than unexpected earnings (Ohlson, 1991). Dechow makes an additional refinement and tests accruals and cash flows for their significance of R^2 based on the univariate regression models, instead of testing the significance of response coefficients (which measures relative correlation with returns).

Dechow (1994) set out to investigate why earnings figures are the standard measure of firm performance and suggests that the importance of accruals is hypothesised to *increase*:

(i) the *shorter* the performance measurement interval;

(ii) the *greater* the volatility of the company’s working capital requirements and investing and financing activities, and;

(iii) the *longer* the company's operating cycle.

Consequently, cash flows will then have more severe timing and matching issues, reducing the ability to predict performance. The results support this hypothesis. Dechow finds that the cash flow measure is problematic when companies have larger and more volatile requirements for investing and financing activities, whereas the earnings measure is more useful. This literature provides a major contribution to the conventional theory that accrual-based earnings are more useful than cash flows in their ability to explain performance, supporting the benefits of accrual accounting and justifying the importance of earnings in finance.

According to Ali (1994), these studies provide unsatisfactory evidence regarding cash flows from operations and working capital. In addition to demonstrating a non-linear relationship between cash flow variables and returns, Ali establishes that returns are less persistent when the magnitude of changes in cash flow from operations increases. This was motivated by evidence of a non-linear relationship between abnormal returns and unexpected earnings by Freeman and Tse (1992). Ali and Pope (1995) develop this new insight further with model refinements such as linear and non-linear specifications of models, as well as using actual and changes in the explanatory variables. The results indicate that when using actual (as opposed to unexpected) explanatory variables in a non-linear model, there were significant positive response coefficients. Furthermore, every variable possessed unique incremental information content.

Sloan (1996) extends a stream of accounting literature around the use of financial statement data to predict future abnormal returns (Ou and Penman, 1989; Holthausen and Larcker, 1992; Stober, 1992). The main hypothesis considers whether share prices fully reflect information about future earnings contained in the accrual and cash flow components of current earnings. A secondary hypothesis investigates whether investors fully consider and price recent earnings announcements into future earnings i.e. post-earnings announcement drift. A comprehensive analysis is performed using all NYSE and AMEX firms' financial statement data over a 30-year period from 1962 and until 1991. A hedge portfolio is constructed which holds a long position in firms that reported low levels of accruals (high cash flows for the same level of earnings) and a short position in firms that reported high levels of accruals (Sloan, 1996). The findings suggest that prices do not fully reflect information contained in the accrual and cash flow components of earnings until it impacts future earnings. This contrasts with the findings of Bernard and Stober (1989), who find no evidence that stock prices respond in any systematic manner to information released about the cash flow and accrual components of earnings. The research uncovered the "accrual anomaly", which results in mispriced securities because investors fail to realise that accruals are less persistent than cash flows. Moreover, Sloan demonstrates that the extent to which current earnings performance persists into the future depends on the relative magnitudes of the cash and accrual components of current earnings.

Houge and Loughran (2000) follow a similar methodological process to Sloan (1996), but also include the Nasdaq in addition to the NYSE and AMEX. This increases the sample size and includes larger technology stocks like Microsoft and Intel. Houge et al. document significant excess returns from a cash flow-based trading strategy. While market reactions to accrual and cash flow components do vary across portfolio deciles, high cash flow firms significantly outperform the Fama and French (1998) three-factor benchmark, and low cash flow firms significantly lag the benchmark. Houge et al. suggest that investors underestimate the long-term persistence of cash flows and fail to appreciate the underlying quality of earnings numbers.

Garrod and Hadi (1998) look further into the information relevance of cash flow for security pricing by separating it into components and comparing with accrual variables. The regression showed that all coefficients, except for finance and tax cash flows, are significant at the 1% level. Garrod and Hadi also look at the information content of cash flow per share. However, no additional significance is found, because it represents the same information as that in cash flows. Quirin, O'Bryan and Wilcox (1999) investigate the information content of earnings and operating cash flows. Similar to Ali (1994) and extending Ali and Pope (1995) using actual operating cash flow data with a larger sample size, Quirin, O'Bryan and Wilcox find interesting results, though not robust. When earnings and cash flows are both positive, returns reacted more significantly. This does not hold when they are both negative or have opposite signs.

Livnat and Lopez-Espinosa (2008) expand prior studies – which have predominantly been done on an annual basis – by looking at quarterly accruals and net operating cash flows (on a single quarter and a rolling four quarter basis) of listed U.S. companies. This makes it more applicable to investors who must evaluate information about firms in a timely manner. The results show that operating cash flows are incrementally and significantly associated with future quarter returns (even after controlling for accruals). Accruals are not significantly associated with future returns after controlling for cash flows, a finding supported by Desai, Rajgopal and Venkatachalam (2004) when investigating firms listed on the NYSE, Amex and NASDAQ markets. These results are robust to an industry analysis, where the cash flow measure is incrementally significant for 13 out of 17 industries considered. However, when rolling four-quarter operating cash flows and accruals are used to construct portfolios that are held for a whole year, operating cash flows dominates accruals in the first three quarters, but not in the fourth quarter (Livnat and Lopez-Espinosa, 2008).

Hackel, Livnat and Rai (1994) consider an investment strategy which identifies firms with a “consistent pattern of operating and free cash flows, low financial leverage, and low free cash flow multiples” (Hackel, Livnat and Rai, 1994:21). The strategy delivers greater returns than those of the S&P 500 Index, similar size portfolios, similar beta portfolios, and similar book-to-market portfolios. Hackel, Livnat and Rai examine the returns in declining markets and specifically adjust for common anomalies, eliminating possible attribution of the results to risk or anomalies and find that a long portfolio strategy based on free cash flow consistently earns abnormal returns. Hackel, Livnat and Rai (2000) later extend this work and find that portfolios based on low free cash flow multiples can yield future abnormal returns. Similar research by Lakonishok, Vishny and Schleifer (1994) confirms that portfolios based on low cash flow multiples outperform those with high multiples.

At a firm-level, prior research indicates that accruals negatively predict returns (Sloan, 1996), and cash flows positively predict returns (Desai, Rajgopal and Venkatachalam, 2004; Pincus, Rajgopal, and Venkatachalam, 2007). Hirshleifer, Hou and Teoh (2009) extend the scope to the aggregate stock market and investigate whether the firm-level accrual and cash flow effects apply at market level. The results reflect the opposite to firm-level relationships – aggregate accruals are shown to be a strong positive time series predictor of aggregate stock returns, while cash flow is a negative predictor. Sloan (1996) offers the ‘earnings fixation hypothesis’ and Vuolteenaho (2002) claims that information about future cash flows is the dominant factor driving firm-level stock returns. Hirshleifer, Hou and Teoh argue that the aggregate-level findings are attributed to behavioural reasons: “Earnings performance attributable to an extra dollar of cash flows is more persistent than earnings performance attributable to an extra dollar of accruals, but limited attention causes investors to neglect this distinction" (Hirshleifer, Hou and Teoh, 2009:393). This has the effect of high accrual (but low cash flow) firms being associated with overvaluation and therefore earning low subsequent returns. This adds to prior literature by Lou (2008), Chan, Hamao and Lakonishok (1991) and Sloan (1996) suggesting reasons for aggregate cash flow's ability to predict returns.

3.2.2 Gross Profit, Operating Profit and Earnings in Predicting Returns

Fama and French (2008) find that there is weak evidence for a relationship between average returns and profitability – where profitability is defined as a firm's gross profits (revenue less cost of goods sold) to its assets. This is robust to controls for market capitalisation and book-to-market values. Fama and French (2006) claim that earnings have explanatory power in Fama and MacBeth (1973) cross sectional regressions, but Novy-Marx (2013) argues that gross profitability has more predictive power than earnings. The rationale being that gross-profits is a measure which is not as tainted by accounting adjustments and policies i.e. the items on an income statement between gross profit and income (before extraordinary items) are less related to "true economic profitability" (Novy-Marx, 2013).

Novy-Marx (2013) performs Fama and MacBeth regressions of returns on measures of profitability, including gross profits-to-assets, earnings-to-book equity and free cash flow-to-book equity. The sample includes stocks from the American Stock Exchange (AMEX) for the period July 1963 to December 2010. Sloan (1996) demonstrates that accruals has significant and independent predictive power in the cross section of returns. In contrast, Novy-Marx (2013) finds that gross profit-to-assets has power to predict returns which persists after controlling for accruals, which cannot be explained by the prior findings. Novy-Marx argues that despite having higher valuation ratios, profitable firms generate higher returns, are less prone to distress and have longer cash flow durations than unprofitable firms. Furthermore, free cash flow does have some predictive power, though less than that of gross profitability, and earnings are not significant in predicting future stock performance.

Gross profit scaled by book value of total assets predicts the cross-section of average returns. Despite results suggesting that gross profit has greater predictive power than net income as per Novy-Marx (2013), Ball et al. (2015) propose that this is not the case and argue that the deflator chosen to measure return prediction affects the outcome. For instance, Novy-Marx (2013) deflates the gross profit measure by the book value of total assets and deflates net income by the book value of equity. When consistent deflators are used, net income equals gross profit in predictive power, according to Ball et.al. These results are counterintuitive because gross profits are not 'claimed' by investors per se. Only after items such as research and development expenses and selling, general and administrative expenses are deducted, do investors have rights to the income. Further to that, academic research suggests that these kinds of expenses (between gross profits and net income) are able to predict future returns (Chan, Lakonishok and Sougiannis, 2001; Eisfeldt and Papanikolaou, 2013).

Ball et.al. (2015) argue that cost of goods sold (which is included in gross profit) and selling, general and administrative expenses (which are lower down on the income statement) are effectively similar in economic substance. The allocation of these expenses is not outlined in Generally Accepted Accounting Principles and is in fact at the discretion of managers. To match current income with current expenses, Ball et al. (2015) follow Novy-Marx's (2013) method of focusing on income statement items that may relate to current revenue and construct an operating profitability ratio. The results show that operating profitability explains the cross section of returns better than both the gross profitability measure of Novy-Marx (2013) and the net income (excluding extraordinary items) measure employed by Ball and Brown (1968).

Ball et al. (2016) follow a similar process to Novy-Marx (2013) and Ball et al. (2015) in constructing their sample, which includes all firms traded on the NYSE, Amex and NASDAQ from July 1963 to December 2010. An operating profitability measure is constructed following the same methodology as Ball et al. (2015) and, as a result of cash flow data constraints, a cash-based operating profitability is derived from company balance sheets. Comparing cash-based operating profitability to operating profitability using the Fama-MacBeth slope coefficients, Ball et al. (2016) find that the cash-based measure is less volatile, has a higher mean and a lower standard deviation.

Cash-based operating profitability (which excludes accruals) outperforms profitability measures such as gross profitability, operating profitability and net income, which all include accruals. This does not corroborate with the hypothesis that investors fixate on profitability put forward by Sloan (1996). Lou (2008) argues that the current reporting practice has the effect of misleading investor perceptions of a company's ability to generate cash, particularly when unusual operating cash flows are hidden in footnotes of the financial statements.

With regards to predicting the cross section of returns, Ball et al. (2016) also find that the cash-based operating profitability measure subsumes accruals, in addition to having no incremental explanatory power. When using an asset pricing model that includes a profitability variable, the presence of accruals causes the regression to separate cash-based operating profitability from the accrual-based variable, which explains the increase in the accrual anomaly (Ball et al., 2016). Moreover, the measure can explain expected returns up to ten years ahead, although it is suggested that this may be a result of cash flow information being slowly integrated into the market after an initial underreaction. In conclusion, Ball et al. suggest that investors are better-off adding only cash-based operating profitability to their investment opportunity set than by adding both profitability and accruals strategies.

Foerster, Tsagarelis and Wang (2017) extend Novy-Marx (2013), Ball et al. (2015) and Hou, Karolyi, and Kho (2011). Foerster, Tsagarelis and Wang argue that the GAAP (Generally Accepted Accounting Principles) environment in which the study is done allows companies too many alternatives in their financial statement presentations. Companies therefore over-aggregate information and are often inconsistent in their presentation methods, making it more difficult for users to understand a firm's true underlying economic activities.

The sample consists of stocks from the S&P 1500 – 500 largest stocks by market capitalisation, 400 mid-cap and 600 small-cap. Cash flow statements are transformed from the indirect method into disaggregated and direct estimates of cash flows from operations, financing activities, taxation and other sources. Portfolios are then created and tested for significance of the return differences. The results show that there is incremental information in segregating cash flow components for predicting the cross section of returns. Although traditional profitability measures have predictive power, the direct cash flow measures are even stronger predictors. In addition, Foerster, Tsagarelis and Wang (2017) find that capital expenditures and cash taxes provide negative incremental predictive power. These are found to be robust to investment horizons, across risk-factors and sector controls.

Interestingly, the best performing high-low portfolio is achieved after sorting by the ‘net cash flows from operations after financing activities less capex’ measure, also achieving the highest Information Ratio. In terms of long-short positions, Foerster, Tsagarelis and Wang (2017) suggest two reasons for the superiority of the cash flow measures: (1) the information contained in cash figures may be more valuable and (2) the cash measures capture a broader dispersion across the portfolios.

Fama and French (2006), Novy-Marx (2013) and Ball et al. (2015) suggest that accounting-based profitability measures, derived via specific methods, better reflect underlying stock value and are therefore the better predictors of returns. In contrast to cash-based operating profitability measures analysed by Ball et al. (2016), Foerster, Tsagarelis and Wang (2017) compare measures more akin to free cash flow and compare them to earnings-based profitability ratios. The findings of their study show that segregated cash flow measures are stronger predictors than both operating profitability and gross profitability to total assets (as used in Novy-Marx, 2013), because they overcome the complications created by accounting inconsistencies. The direct cash flow measures are particularly more accurate representations of the true value of an entity and are economically ‘cleaner’.

3.2.3 South African Cash Flow-related Academic Research

There is an apparent lack of robust, empirical research on return prediction based of cash flow figures in South Africa. Given the country’s relatively sophisticated financial sector and challenging business conditions, cash flows are generally studied in terms of business failure or liquidity. Jooste (2007) finds that cash flow ratios predict bankruptcy up to three years prior to the event. These companies often show a net profit and positive liquidity ratios, yet the cash flows provide early signals of financial strengths or weaknesses of a firm (Jooste, 2007).

Nyamgero (2015) finds that levels of cash holdings of South African companies are lower than they were in 1997, contrary to the criticisms major South African companies have received regarding excess cash sitting on their balance sheets. Jooste (2007) investigates failed entities and finds evidence suggesting bankrupt companies have lower cash flows and smaller reserves of liquid assets. This makes it difficult to meet debt obligations and creates further credit risk. It is unclear whether this relationship is reflected in share prices and returns on the Johannesburg Stock Exchange.

Gumbi (2013) runs a simple regression analysis of current cash flows and future share prices over a sample period of 11 years on the JSE. The results show a weak relationship in both the long and short run - average R^2 values are 0.24 and 0.33, respectively. These findings are consistent with Kim and Kross (2005) but differ in that the relationship between cash flows and share prices is growing (rather than declining) in the long run. It must be noted that the Gumbi (2013) did not address stationarity of cash flows and share prices, nor was the analysis particularly robust.

Van Niekerk (1992) shows that cash flows (from operating activities and from investing activities) have incremental information content to model share prices on the JSE. There was no such evidence for share returns. Fourie (1992) finds that cash flows are significantly associated with abnormal returns, using various cash flow variables to model twelve-month cumulative abnormal returns of 35 companies listed on the JSE over the 1987-1991 period. Wessels, Smith and Gevers (1993) propose that cash flow from operations can be considered an important indicator of the quality of income of a company. In addition, the simple linear regression model based on a smoothed cash flow beta is shown to provide significant explanatory power of the variability in market beta. Wapenaar (1996) however, does not find any significant correlation of cash flow variables with returns. De Jager (1997) conducts a factor analysis on which ratios best serve as explanatory variables in predicting corporate success or failure on the JSE. Surprisingly, out of 62 variables, De Jager finds that cash flow return on investment and financial leverage are the two most significant.

3.3 Summary and Conclusion

This chapter introduced the empirical findings on anomalies and the predictive ability of traditional accrual-based measures as well as research on cash components. Although many anomalies have been documented, limited attention has been given to cash flow variables, especially within a South African context. International research around the usefulness of cash flow data for security pricing has produced inconsistent results. This study aims to explore the value of using cash flow information in the fundamental investment process by providing clarity on the relationship between a company’s cash flows and return behaviour. The lack of empirical evidence around this topic motivates further investigation.

Chapter 4: Data

This chapter introduces the data used to answer the research problems set out in Chapter 1. The data consists of two distinct subsets: company fundamental data and share price data. The biases that may occur in financial data will be discussed, as well as adjustments that were made to the dataset to mitigate such biases. Descriptive statistics of the preliminary dataset are also presented.

Section 4.1 discusses the sample collected and used in the study. Section 4.2 covers share data and adjustments in terms of completeness, comparability, liquidity and outliers. Section 4.3 considers the possibility of bias and adjustments with regards to data snooping, look ahead bias and survivorship bias. Section 4.4 analyses some descriptive statistics of the dataset and Section 4.6 presents a summary and conclusion.

4.1 Data

All market and fundamental data were obtained from *Datastream*, accessed from the Business corner in The Research Wing at The Chancellor Oppenheimer Library at the University of Cape Town. Together with Microsoft Excel, *EViews* was used perform the data analyses in this study.

Share price, return and financial statement information was collected for constituents of the Johannesburg Stock Exchange All-Share Index (JALSH) for the period June 2007 to August 2018. However, the analysis was conducted for the period March 2008 to March 2018 due to trailing and lagging values employed in the methodology. The data is also subject to constraints discussed in Section 4.3 and Section 4.4 below.

The major fundamental and market variables are outlined below. The study will refer to variables by their ‘Code’, while ‘Datatype’ refers to the *Datastream* mnemonic. Detailed definitions are provided in Appendix A.

Table 4.1 Fundamental and Market Data

Financial Statement Data	Code	Datatype
Statement of Profit/Loss and Other Comprehensive Income		
Operating Profit	OP	SOPI
Net Income Before Extraordinary Activities	NIBX	NIBX
Net Income	NI	NINC
Statement of Financial Position		
Total Assets	TA	DWTA
Total Equity	BVE	QTLE
Shares Outstanding	NOSH	NOSH
Statement of Cash Flows		
Net Cash Flows from Operating Activities	CFOA	OTLO
Cash Flows from Financing Activities	Fin Act	FTLF
Cash Flows from Investing Activities	Inv Act	ITLI
Free Cash Flow	FCF	FCF
Net Change in Cash	NCIC	SNCC
Market Data		
Ticker Name		
Industry/Sector		
Market Value of Equity	MVE	MV
Price	P	P
Total Return Index	TR	RI

The dataset in Table 4.1 ensures that each share has sufficient cash flow and profitability variables on an interim basis. Though 'Gross Profit 'and 'Capital Expenditure' were initially included, there was significant data missing from the sample. This is attributed to a limitation of *Datastream,* specifically the lack of single-period semi-annual cash flow figures for 'Capital Expenditure' and semi-annual 'Gross Profit' figures. Although annual figures were available, it was considered inaccurate to assume an even split of year-end results. Therefore, these measures were excluded from the analysis.

4.2 Returns and Variable Adjustments

Data was collected for selected constituents on the Johannesburg Stock Exchange's All-Share Index (JALSH) for the period June 2007 to August 2018.

Returns and log returns were calculated for six-month holding periods, using the Total Return Index with gross dividends reinvested (defined in Appendix A). Histograms were plotted on the return data, showing a relatively normal distribution (see Appendix E). It is important to note that major outliers occurred because of the sample period commencing during the 2008 mortgage crisis in the U.S. Shares with the most exposure in this regard were Lonmin PLC and Northam Platinum Ltd, with returns of -73% and -68% respectively. The data was winsorised at 1% and 99% and the remaining extreme values were left in for purposes of measuring cash flow and profitability impact.

4.2.1 Completeness

Empirical research requires the dataset to be complete in all material respects. A method employed to ensure completeness was to iterate all data mining at each month-end in the sample period using the constituent list at that point in time. Consequently, all data in this study is therefore subject to the level of completeness of the data on *Datastream.* Missing returns and fundamental data that arose from delisted companies was left in the sample. A consideration was made with respect to the approach taken by Haugen and Baker (1996) in dealing with incomplete return data – the method of filling missing data-points with the population mean. It is the author's opinion that this strategy will materially affect the analysis as well as bias the results. The sample size and the nature of the methodology carried out in this study allows for missing values in the data to be left blank.

As per JSE listing requirements 3.15 (a)-(c), entities are required to release interim or quarterly financial reports to shareholders (Johannesburg Stock Exchange, 2018). As this study aims to investigate the effects of accounting data on return predictability, the interim figures were incorporated into the dataset. *Datastream* provides semi-annual single-period figures for the Statement of Cash Flows and Income Statement. The sample period was therefore further split into periods of six-months and all fundamental data was collected in this format. In terms of returns, consideration was given to trading costs playing a major role for a one-month holding period. Thus, a six-month holding period (lagged by three months) was used in this study. Treatment of companies with changes in their reporting dates are discussed in Section 4.2.2.

4.2.2 Comparability

Due consideration was given to the comparability of the data. All firms in the sample adhere to the International Financial Reporting Standards (IFRS) as per listing requirements on the Johannesburg Stock Exchange. The financial statement information and return data for all shares in the sample are denominated in South African Rand (ZAR). Where financial data is reported in another currency, *Datastream* converts items on historical income statements and cash flow items using the average monthly exchange rate during the fiscal year, while balance sheet items are converted using the fiscal year end exchange rate.

As per Fama and French (1992) and common practise in research using accounting metrics, this analysis excluded financial companies, including banks, insurance companies, investment holding companies, asset managers and Real Estate Investment Trusts (approximately 34% of the sample after the adjustment for survivorship bias in Section 4.3.3). This is due to significant regulatory and structural differences in the South African context. Companies with negative earnings and cash flows were included. The industry classification of the shares in the sample is presented in Table 4.2 below.

Table 4.2 Industry Classification Benchmark (ICB) of Sample

ICB Industry Name	No. of Companies
Basic Materials	22
Consumer Goods	12
Consumer Services	23
Health Care	5
Industrials	17
Oil & Gas	1
Technology	3
Telecommunications	4

With respect to companies that have elected to issue quarterly reports, a conversion to a semi-annual format was applied by adding the first two and last two quarters, respectively.

Approximately 9% of companies in the sample had changes in their financial year end. These were categorised and split into their year-end sub-group for purposes of matching the deflators and return periods correctly – reflected in Appendix B.

4.2.3 Liquidity

The liquidity and tradability of shares have a material impact on empirical evidence. This is particularly important in an emerging market. Due consideration and appropriate adjustments were applied to thinly traded and illiquid shares in the sample (Bowie, 1994; Atchinson, Butler and Simonds, 1987). The full All-Share Index constituent lists was obtained at the start, middle and end of the sample period (2008, 2013 and 2018). The constituents were ranked by (historical) market value of equity at each stage, and three lists of the top 100 were combined to ensure a sample of liquid shares.

4.2.4 Outliers

Errors or abnormal events during the sample period may generate some outliers in the dataset. First, visible errors were manually removed from the return and fundamental data. A two-tailed winsorisation procedure was then applied at 1% and 99% using *EViews*. Winsorisation replaces outliers with the 1st and 99th percentile values instead of removing data, allowing for a more complete dataset. The six-month return periods were matched with respective year-end dates and winsorised thereafter. Log returns were also calculated. Histograms of all measures and returns were then plotted to consider the distributions, shown in Appendix E. Outliers are still present but represent significant transactions and movements in cash flows, which may provide a useful comparison to returns in those periods. The preliminary correlations in Chapter 5 were performed before and after winsorisation to provide some additional information on the data.

4.3 Bias and Adjustments

Empirical studies conducting quantitative financial analyses may reach flawed conclusions because of bias present in the dataset. This study aimed to anticipate possible sources of bias and provide adjustments to mitigate the effects, ensuring validity and robustness of the results and interpretations drawn. These are discussed below.

4.3.1 Data Snooping

According to Haugen and Baker (1996), this bias occurs when researchers:
(a) examine the properties of a database or the results of other studies of a database;
(b) build predictive models employing promising factors based in the previous results, and;
(c) test the power of their models on the same database.
The data snooping issue, also raised by Black (1993), can be addressed either by using data from markets that have not been researched extensively, or predicting by using time periods that are new to analysis (Haugen and Baker, 1996). To mitigate this concern, the research was conducted on a current sample period using a unique research procedure. The methodology employed in this study, to the author's knowledge, has never been carried out in a South African context.

4.3.2 Look-Ahead Bias

Particularly relevant to academic research using accounting and price data, look-ahead bias occurs when information is used that would have been either unknown or unavailable during the sample period analysed (Banz and Breen, 1986; Haugen and Baker, 2006). This leads to results that are inaccurate and conclusions that are invalid.

To address this bias, this study only used data that would have been available at the time. The approach taken was informed by the methodology of Fama and French (1992). Given that *Datastream* updates its database when relevant company and market data becomes public knowledge, trailing three-month values for the fundamental variables were created. These were then updated with each interim-period's new financial statement information, as South African firms release interim results (i.e. in the middle of the company's financial year). This was then combined with the respective return data to conduct the statistical tests described in Chapter 5.

4.3.3 Survivorship Bias

Survivorship bias occurs when results obtained from existing firms are incorrectly interpreted as representative of the entire sample (Kothari, Shanken and Sloan, 1995). Haugen and Baker (1996) point out that if a performance analysis is carried out on firms that have remained listed on an exchange, their performance will likely exceed the market. Omitting delisted firms, especially in a return prediction study, can lead to incorrect levels of significance and predictive ability.

The dataset in this study was adjusted for survivorship bias such that delisted shares were included in the sample during the period of their listing on the JSE. The liquidity adjustment in Section 4.2.3 also provides an opportunity to address this bias. The historically-ranked constituent lists were compared such that shares appearing at least once in the top 100 are included in the final constituent list. As per Section 4.2.2, financial companies were removed after adjusting for survivorship bias.

4.4 Descriptive Statistics

To gain some initial insight into the dataset, an analysis of various descriptive statistics was carried out. This informed the general nature and characteristics of the data in the study before the methodology in Chapter 5 was applied.

The analysis was carried out for the 85 constituents listed on the JSE from June 2007 to August 2018. The maximum, minimum, mean, standard deviation, skewness and kurtosis were calculated for each profitability and cash flow measure. The set of summary statistics was calculated before and after winsorisation at 1% and 99% and before any standardisation procedures, is presented in Table 4.4 below.

Table 4.4 Descriptive Statistics Before and After Winsorisation

Before Winsorisation

Measure	Mean	Median	Min	Max	Std. Dev.	Kurtosis	Skew	Obs.
CFOA/TA	0.0445	0.0396	-0.2260	0.4916	0.0625	5.4498	0.9756	1696
FCF/TA	0.0198	0.0169	-0.2337	0.3046	0.0599	3.0822	0.5721	1696
NCIC/TA	0.0038	0.0019	-0.2452	0.3892	0.0544	5.3694	0.5026	1694
Fin Act/TA	-0.0037	-0.0050	-0.9000	0.5580	0.0665	30.5943	-0.1932	1677
Inv Act/TA	-0.0368	-0.0292	-0.5631	0.8108	0.0558	45.4196	0.5078	1696
OP/TA	0.0618	0.0550	-1.0028	0.5198	0.0700	46.7907	-1.7325	1696
NIBX/TA	0.0403	0.0355	-0.9090	0.6438	0.0561	69.3554	-2.6365	1696
NI/TA	0.0411	0.0353	-0.9090	1.7415	0.0707	223.6135[1]	6.8374[1]	1696
CFOA/MVE	0.0504	0.0360	-1.0752	1.8042	0.1179	42.9731	2.1369	1696
FCF/MVE	0.0148	0.0137	-2.1216	1.8042	0.1230	95.3977	-2.6468	1696
NCIC/MVE	0.0063	0.0021	-0.8712	1.8409	0.1014	75.8854	3.8600	1696
Fin Act/MVE	-0.0004	-0.0051	-2.3944	1.5261	0.1394	88.1654	-0.2072	1678
Inv Act/MVE	-0.0442	-0.0282	-1.6791	2.6633	0.1412	111.1577	3.0773	1696
OP /MVE	0.0541	0.0516	-3.7471	1.4877	0.1577	237.7787	-9.6084	1696
NIBX/MVE	0.0278	0.0327	-3.3965	1.9998	0.1553	194.6092	-8.0502	1696
NI/MVE	0.0205	0.0327	-7.6223	2.0052	0.2502	553.0781[1]	-19.6591[1]	1696

After Winsorisation

Measure	Mean	Median	Min	Max	Std. Dev.	Kurtosis	Skew	Obs.
CFOA/TA	0.0441	0.0396	-0.1068	0.2465	0.0581	1.4705	0.5302	1696
FCF/TA	0.0197	0.0169	-0.1295	0.2100	0.0569	1.3672	0.4965	1696
NCIC/TA	0.0036	0.0019	-0.1433	0.1657	0.0500	1.6592	0.2310	1694
Fin Act/TA	-0.0034	-0.0050	-0.1523	0.2447	0.0555	5.4675	1.3186	1677
Inv Act/TA	-0.0371	-0.0292	-0.2438	0.0809	0.0436	6.6448	-1.7422	1696
OP/TA	0.0623	0.0550	-0.0882	0.2955	0.0548	3.5061	1.0868	1696
NIBX/TA	0.0409	0.0355	-0.0788	0.1946	0.0416	2.5301	0.7873	1696
NI/TA	0.0410	0.0353	-0.0875	0.2083	0.0436	2.8767	0.8003	1696
CFOA/MVE	0.0503	0.0360	-0.2443	0.4487	0.0941	4.7980	1.0479	1696
FCF/MVE	0.0162	0.0137	-0.3097	0.3282	0.0846	4.5137	-0.0682	1696
NCIC/MVE	0.0055	0.0021	-0.2762	0.3401	0.0775	6.1082	0.6958	1696
Fin Act/MVE	-0.0009	-0.0051	-0.2680	0.4104	0.0815	8.9835	1.5314	1678
Inv Act/MVE	-0.0457	-0.0282	-0.4349	0.2066	0.0778	9.5698	-2.1161	1696
OP/MVE	0.0567	0.0516	-0.3173	0.3835	0.0760	10.4669	-0.4467	1696
NIBX/MVE	0.0297	0.0327	-0.3728	0.2707	0.0674	16.7535	-2.4909	1696
NI/MVE	0.0273	0.0327	-0.4288	0.2538	0.0729	19.1383	-3.1583	1696

[1] Extreme values for NI/TA before winsorisation are a result of Bidvest Corporation's discontinuation of operations for the year ended June 2016. This transaction, amounting to R79.216 million, made up 97% of the Net Income figure for this period.

4.5 Control Variables

Similar to Ball et al. (2015) and Foerster, Tsagarelis and Wang (2017), three control variables were used in this study to isolate the impact of cash flows and profits on returns.

Log(BVE/MVE) is the natural logarithm of the ratio of book value of equity to market value of equity. BVE/MVE is also referred to as Book-to-Market (BTM), a ratio of the firm's book value of equity to market value of equity, and functions as a proxy for risk (Auret and Sinclaire, 2006; Chen and Zhang, 1998). Fama and French (1992) suggest that this measure has a significant role in returns. Auret and Sinclaire (2006) explain that when comparing two firms with similar book values of equity, the firm with a higher perceived risk will have a lower market value of equity. As the marginal utility of risk is negative, the firm with less certainty will have higher risk and thus a lower BTM (Markowitz, 1959; Auret and Sinclaire, 2006). Thus, this study controls for the possible influence on returns for every data point.

Due consideration was also given to the size effect, in addition to the normalisation procedures carried out in Section 5.2. Log(MVE) is the natural logarithm of market value of equity, a proxy for the size of a company (Foerster, Tsagarelis and Wang, 2017). As per Fama and French (1992), this study controls for the possibility of stocks with a small market capitalisation outperforming those with a larger market capitalisation.

Lastly, $r_{6,6}$ is the share's six-month-prior return. This variable aims to capture the price momentum anomaly that Page, Britten and Auret (2013) found to be present on the JSE. Every stock's six-month return was also lagged by 6-months and included as a control in every regression performed. Descriptive statistics on the control variables are shown below.

Table 4.5 Descriptive Statistics for Control Variables

Measure	Mean	Median	Min	Max	Std. Dev.	Kurtosis	Skew	Obs.
Log(BVE/MVE)	-0.8569	-0.8347	-4.6296	1.9512	0.8403	0.1311	-0.0298	1559
Log(ME)	9.8369	9.6498	5.8237	14.4789	1.4604	0.3281	0.5321	1563
$r_{6,6}$	0.0686	0.0559	-0.6845	0.9041	0.2468	1.2386	0.3755	1550

4.6 Summary and Conclusion

This chapter presented the preliminary data that was used in this study. The companies' fundamental and share price information were collected from *Datastream* for all constituents listed on Johannesburg Stock Exchange's All-Share Index (JALSH) for the period June 2008 to June 2018. Due consideration was given to completeness, comparability, liquidity and outliers. The data was winsorised and adjustments were made accordingly.

Pertinent biases in the data, as well as strategies to mitigate their impact, was also discussed. Return and financial statement data was collected every six-months and lagged by three months, to reflect the average time it takes for financial results to be released in South Africa and to reduce the effect of any look-ahead bias. To address survivorship bias, constituent lists were obtained and ranked by historical market capitalisation, such that delisted shares are not overlooked**.**

The following chapter will discuss the methodological process that was followed in order to test the predictive ability of cash flows, investigate whether different categories of cash flows vary in their predictive ability, and compare this with traditional measures employed in academia and industry.

Chapter 5: Methodology

This chapter discusses the research procedures and aspects of the methodology required to determine whether cash flows are better share return predictors that profits. Empirical evidence in this regard is limited globally, and to the author's knowledge, non-existent in a South African context. As discussed in previous chapters, style anomalies have been identified on the JSE, and mixed results have been found on profitability as a predictor of share returns. Foerster, Tsagarelis and Wang (2017) provided evidence in the USA that cash flows do possess predictive power superior to that of traditional profitability measures. The methodology employed to investigate the predictive ability of cash flows on the JSE will be outlined in this chapter.

5.1 Trailing and Lagging Variables

The stock returns were collected for every six-month period and lagged by three months to reflect the time it takes for financial results to be released in South Africa. The statistical tests were conducted after these values were combined with results released at each interim-period and combined with the return data. For purposes of correlation, the data was compiled into a table with columns counting six-month periods (regardless of dates). This was done considering different reporting dates of companies. Care was taken to allocate companies into their correct groups and adjust for changes in year-ends. As mentioned in Section 4.3.2 with regards to look-ahead bias, the accounting data used in return prediction was taken from an interim or financial year-end three months prior to the date of measurement.

5.2 Normalisation of Variables

Empirical studies using cross-sectional regressions must give due consideration to the choice of deflator(s) used to normalise measures. This is done to avoid a few larger firms driving the regression results of the entire sample (Easton and Sommers, 2003). These differences in scale can lead to incorrect results through coefficient bias, heteroscedasticity and falsely increasing the R^2 metric. As a solution, prior literature suggests either deflating the regression equations by a proxy of the scale, or to include a scale proxy as an independent variable (Christie, 1987; Novy-Marx, 2013; Ball et al., 2015). This study used the latter approach.

According to Christie (1987), cross-sectional return studies on accounting variables ought to use market value of equity as the deflator. This allows for results to be correctly attributed to the mismeasurement of future cash flows by market agents. Novy-Marx (2013) deflates the gross profit measure by the book value of total assets and deflates net income by the book value of equity. Ball et al. (2015) criticise this methodology and argue that the deflator chosen to measure return prediction affects the outcome.

Prior research by Fama and French and Novy-Marx (2013) suggests that when constructing a profitability measure, the numerator and denominator should match with respect to cash flow rights. Ball et al. (2015) explains that if the profit measure in the numerator represents a flow to equity holders (net income), then the denominator should represent an equity-holder claim (market of value equity). Likewise, if the profit measure in the numerator represents flows to both equity- and debt-holders (gross profit or operating profit), then the denominator should be total assets (Ball et al., 2015).

To ensure a robust methodology, this study initially deflated measures with both market value of equity (outstanding shares multiplied by month-end share price) and total assets (as stated on the interim Statement of Financial Position). The market value of equity was lagged by a three-month period to reflect a more realistic period for markets to receive information.

5.3 Correlations

The cash flow and accounting metrics in Table 4.2 were deflated by both market value of equity (MVE) – lagged by three months – and total assets (TA) for purposes of the correlation. As stated by Ball et al (2015), market value of equity fluctuates more often and by greater magnitudes, whereas total assets experiences slower changes over time. A comparison will therefore provide useful supplementary evidence for choice of deflator on the JSE. All measures were then winsorised at 1% and 99% to remove the effects of outliers. Next, a panelled correlation of the measures for every data point in the time series allowed for preliminary results. This was done prior to any standardisation procedures and helps assess the behaviour of the cash and accounting measures as well as the impact of both deflators. Thereafter, an additional correlation table with the final variables to be used in the regressions is presented.

A correlation matrix was used to investigate the eight profitability and cash flow measures across the 85 shares in this study. Correlation coefficients are calculated for every pair throughout the sample period for a preliminary insight into the relationships and are presented in Table 5.3.1 and Table 5.3.2.

Table 5.3.1 Correlations

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	CFOA/TA	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	FCF/TA	0.91	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	NCIC/TA	0.55	0.58	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-
4	Fin Act/TA	-0.38	-0.38	0.19	1.00	-	-	-	-	-	-	-	-	-	-	-	-
5	Inv Act/TA	-0.20	-0.06	0.15	-0.45	1.00	-	-	-	-	-	-	-	-	-	-	-
6	OP/TA	0.43	0.36	0.05	-0.27	-0.18	1.00	-	-	-	-	-	-	-	-	-	-
7	NIBX/TA	0.37	0.33	0.06	-0.28	-0.10	0.89	1.00	-	-	-	-	-	-	-	-	-
8	NI/TA	0.36	0.31	0.06	-0.27	-0.10	0.86	0.96	1.00	-	-	-	-	-	-	-	-
9	CFOA/MVE	0.57	0.52	0.37	-0.19	-0.08	-0.06	-0.08	-0.10	1.00	-	-	-	-	-	-	-
10	FCF/MVE	0.62	0.70	0.43	-0.23	-0.04	0.09	0.11	0.09	0.82	1.00	-	-	-	-	-	-
11	NCIC/MVE	0.38	0.40	0.72	0.14	0.13	0.00	0.00	0.01	0.45	0.52	1.00	-	-	-	-	-
12	Fin Act/MVE	-0.25	-0.24	0.13	0.72	-0.38	-0.11	-0.11	-0.11	-0.33	-0.31	0.20	1.00	-	-	-	-
13	Inv Act/MVE	-0.04	0.03	0.10	-0.37	0.59	0.18	0.23	0.25	-0.37	-0.12	0.13	-0.39	1.00	-	-	-
14	OP/MVE	0.09	0.10	0.04	-0.06	0.01	0.37	0.29	0.26	0.19	0.16	0.05	-0.10	-0.08	1.00	-	-
15	NIBX/MVE	0.11	0.13	0.04	-0.10	0.01	0.36	0.49	0.46	0.11	0.22	0.05	-0.11	0.07	0.75	1.00	-
16	NI/MVE	0.10	0.11	0.03	-0.09	0.00	0.35	0.47	0.52	0.00	0.16	0.02	-0.11	0.15	0.59	0.87	1.00

Interestingly, the measures with MVE in the denominator display significantly lower correlations amongst themselves, as well as with TA-deflated variables. This result is consistent with Ball et al. (2015), implying the choice of deflator does impact comparisons across companies. In line with empirical evidence of predictability by Ball et al. (2015, 2016) and Foerster, Tsagarelis and Wang (2017), only measures with total assets as the deflator were used for regressions.

CFOA/TA and FCF/TA show an extremely high correlation of 0.91, which drops to 0.62 when deflated by MVE. In light of this, an additional correlation was calculated before any winsorisation, showing 0.90 and 0.46 respectively (refer to Appendix D). This is attributed to the fact that FCF uses CFOA in its calculation in combination with any *Datastream* adjustments on Capex in the FCF figures provided. FCF was subsequently removed for purposes of regressions. Similarly, NIBX and NI were initially included to assess the impact of unusual once-off transactions on return predictability. However, the correlations between them are practically identical across all variables. NIBX was therefore also removed as an independent variable on the basis of redundancy.

Correlations between profitability measures are relatively high. NI/TA, NIBX/TA and OP/TA all show correlations higher than 0.85, and approximately 0.5 when deflated by MVE. The correlations imply that accounting-based measures tend to behave in a similar fashion, while the cash measures display weaker relationships. NCIC and Inv Act show the weakest relationships, with the latter a surprise given the importance investors place on scaling operations through asset purchases. Inv Act/MVE has a weak negative relationship to Fin Act/TA and Fin Act/MVE, with -0.37 and -0.39 respectively. Similar to Foerster, Tsagarelis and Wang (2017), Fin Act has a consistently negative pattern of relatively low correlations and CFOA and FCF have the highest positive correlations.

The most important finding is that correlations between profitability and cash measures are relatively weak. CFOA/TA and OP/TA is the highest pair with a correlation of only 0.43, while the correlation between NI/MVE and CFOA/MVE is effectively 0. For the remainder of this study, the cash and accrual measures used are ones deflated by total assets (TA), unless otherwise stated.

Table 5.3.2 Correlations for Final Variables

	Y	CFOA /TA	NCIC /TA	Fin Act /TA	Inv Act /TA	OP /TA	NI /TA	Log(BVE /MVE)	Log(MVE)	r6,6
Y	1.00									
CFOA/TA	0.11	1.00								
NCIC/TA	0.08	0.53	1.00							
Fin Act/TA	-0.04	-0.40	0.21	1.00						
Inv Act/TA	-0.01	-0.18	0.16	-0.45	1.00					
OP/TA	0.17	0.41	0.02	-0.28	-0.16	1.00				
NI/TA	0.20	0.33	0.03	-0.27	-0.08	0.85	1.00			
Log(BVE/MVE)	-0.08	-0.22	0.02	0.10	0.20	-0.54	-0.53	1.00		
Log(MVE)	-0.05	0.11	-0.02	-0.11	-0.04	-0.01	0.09	-0.21	1.00	
r6,6	0.00	0.04	-0.02	-0.06	-0.01	0.17	0.20	-0.20	0.07	1.00

Table 5.3.2 presents the variables that will be used in the panel regression procedures. Based on Table 5.3.1, redundant variables have been removed and total assets (TA) has been applied as the deflator. This is in line with the methodology applied by Foerster, Tsagarelis and Wang (2017). The six-month ahead returns (*Y*) and three control variables (defined in Section 4.5) are added to the correlation. Log(BVE/MVE) appears to have the strongest correlations amongst the control variables, particularly to the OP and NI accrual measures. r6,6 (six-month ahead return lagged by six months) displays similar behaviour, although unlike Log(BVE/MVE), it is positive. Interestingly, the strongest relationships

with the dependent return variable are that of NI, OP and CFOA. These are all relatively low but positive. In terms of return predictability, the disparity seen between cash and accrual variables prompts a more in-depth investigation.

5.4 Panel Data

This study made use of panel data to investigate the predictability of returns based on cash flow and accrual measures. Panel data refers to the process of pooling data on a cross-section of a characteristic over several time periods and running a regression over these two dimensions (Baltagi, 2005). The combination allows for the measure of effects which cannot be identified in pure time-series and cross-sectional data (De Jager, 2008). Panel data provides an improved efficiency of the regression estimates. Hsiao (2007) explains that a larger data set allows for greater variability, less collinearity and more degrees of freedom. Arellano (2003) describes a major difficulty of cross-sectional analysis – the correlation between unobserved independent variables and observed ones. The main advantage is the control for individual heterogeneity – addressing the unobserved individual-specific effects in the model which reduces possible bias in the estimates (Baltagi and Song, 2006). Three panel data models were considered, and tests were conducted to choose the most appropriate model.

5.4.1 Pooled OLS

This method pools all of the data into an ordinary least squares regression. Assumptions outlined in Section 5.7 apply for this model to be most appropriate. The pooled OLS equation is:

$$\gamma_{it} = \alpha + X'_{it}\beta + \varepsilon_{it}\,(u_{it} = 0) \tag{5.4.1}$$

where γ_{it} is the dependent variable for cross-sectional unit *i* at time *t*, α is the intercept, X'_{it} is the independent variable in the model with β representing the slope vectors and ε_{it} is the estimation error. u_{it} represents the individual effect.

Greene (2008) states that when no individual effects are present, the pooled OLS method leads to estimates that are both efficient and consistent. Individual effects in the data cause several OLS assumptions to be violated, resulting in OLS estimates that are inferior to those of random or fixed effects models (Park, 2011).

5.4.2 Random Effects

The random effects model assumes that individual effects are not correlated with any regressors, the intercept is constant and error variances are distributed randomly across sections and time periods (Greene, 2008). The random effects model is:

$$\gamma_{it} = \alpha + X'_{it}\beta + (u_{it} + v_{it}) \qquad (5.4.2)$$

where γ_{it} is the dependent variable for cross-sectional unit *i* at time *t*, α is the intercept, X'_{it} is the independent variable in the model with β representing the slope vectors and ε_{it} is the estimation error. u_{it} represents the individual specific random heterogeneity effect.

The random effects model allows for inferences to be made about the entire population, as the levels are a random sample from a larger population of possible levels. It allows for the inclusion of time invariant variables in the model, with a reduction of the number of parameters to be estimated. However, Greene (2008) argues that it will produce inconsistent estimates if u_{it} is in fact correlated with regressors in the model.

5.4.3 Fixed Effects

The fixed effects model assumes some correlation between the error term and predictor variables. This approach removes the effect of time-invariant characteristics, as they are assumed to be unique to the individual observation and thus uncorrelated with other observations. The fixed effects model is:

$$\gamma_{it} = (\alpha + u_{it}) + X'_{it}\beta + v_{it} \qquad (5.4.3)$$

where γ_{it} is the dependent variable for cross-sectional unit i at time t, α is the intercept, X'_{it} is the independent variables in the model with β representing their slope vectors and ε_{it} is the estimation error. u_{it} represents the unobserved individual effect.

The model is able to investigate the changes within a group of observations. Baltagi (2005) states that this is useful when the statistical inference is limited to the observations under analysis, and thus most appropriate for most accounting research. The drawback is that fixed effects models cannot cater for time-invariant characteristics and the resulting slopes of regressors are sample-dependent (Clark and Linzer, 2012).

5.5 Regression Model Selection

To determine whether these effects exist in the cash and accrual data collected from the JSE, Park (2011) and Greene (2008) suggest that statistical tests be carried out. These were used in the selection of the appropriate method to model the sample.

5.5.1 Fixed vs Pooled

Park (2011) recommends using the F-test in deciding whether individual fixed effects are present in the data. The null hypothesis is that fixed effects, u_{it} in Equation 5.4.3, are equal to zero. In other words, there is no individual heterogeneity and a pooled OLS regression is better suited. The alternative hypothesis states that there is some unobserved individual heterogeneity and a fixed effects model is better suited to the data (Park, 2011). According to the results presented in Table G1 (Appendix G), the null hypothesis is rejected at the one percent level for all fourteen regressions. The fixed effects model is therefore preferred.

5.5.2 Random vs Pooled

Breusch and Pagan (1980) suggest applying the Lagrange multiplier (LM) in choosing the appropriate model. The LM identifies random effects by examining whether the "individual specific variance components are zero" (Park, 2011). The resulting test statistic follows a chi-squared distribution. The null hypothesis suggests that pooled OLS is favoured, while the alternative hypothesis states that the random effects model is better suited to handle the heterogeneity. According to Table G1 (Appendix G), none of the models are statistically significant at the ten percent level. The null hypothesis is not rejected and a pooled OLS model is found to be better suited.

5.5.3 Fixed vs Random

Greene (2008) and Wooldridge (2013) recommend the Hausman Test when deciding between a fixed or random effects model. The comparison is done under the null hypothesis of individual effects being uncorrelated with regressors in the model – the random effects model is preferred in this case. Park (2011) argues that if the null hypothesis is not rejected, both model's estimators are consistent, but the fixed effects model is more inefficient. The null is therefore rejected if the estimates between the two models are sufficiently different and the fixed effect estimators are accurate and consistent. As shown in Table G1, all fourteen models are statistically significant at the one percent level. Therefore, the fixed effects model provides significantly better estimates for the regression models used in this study.

5.6 Regression Models

The study made use of fourteen regression models to investigate the predictive ability of cash and accrual variables. The independent variables are regressed both individually and in various combinations to assess the impact on significance levels.

Model 1

$$Y_{i,t} = \beta_0 + CFOA_{i,t-3}\beta_1 + Log(BVE/MVE)_{i,t-3}\beta_2 + Log(MVE)_{i,t-3}\beta_3 + r_{6,6} + u_{i,t} + \varepsilon_{i,t} \tag{5.6.1}$$

Model 2

$$Y_{i,t} = \beta_0 + NCIC_{i,t-3}\beta_1 + Log(BVE/MVE)_{i,t-3}\beta_2 + Log(MVE)_{i,t-3}\beta_3 + r_{6,6} + u_{i,t} + \varepsilon_{i,t} \tag{5.6.2}$$

Model 3

$$Y_{i,t} = \beta_0 + OP_{i,t-3}\beta_1 + + Log(BVE/MVE)_{i,t-3}\beta_2 + Log(MVE)_{i,t-3}\beta_3 + r_{6,6} + u_{i,t} + \varepsilon_{i,t} \tag{5.6.3}$$

Model 4

$$Y_{i,t} = \beta_0 + NI_{i,t-3}\beta_1 + Log(BVE/MVE)_{i,t-3}\beta_2 + Log(MVE)_{i,t-3}\beta_3 + r_{6,6} + u_{i,t} + \varepsilon_{i,t} \tag{5.6.4}$$

Model 5

$$Y_{i,t} = \beta_0 + CFOA_{i,t-3}\beta_1 + OP_{i,t-3}\beta_2 + Log(BVE/MVE)_{i,t-3}\beta_3 + Log(MVE)_{i,t-3}\beta_4 + r_{6,6} + u_{i,t} + \varepsilon_{i,t} \tag{5.6.5}$$

Model 6

$$Y_{i,t} = \beta_0 + CFOA_{i,t-3}\beta_1 + NI_{i,t-3}\beta_2 + Log(BVE/MVE)_{i,t-3}\beta_3 + Log(MVE)_{i,t-3}\beta_4 + r_{6,6} + u_{i,t} + \varepsilon_{i,t} \tag{5.6.6}$$

Model 7

$$Y_{i,t} = \beta_0 + CFOA_{i,t-3}\beta_1 + NCIC_{i,t-3}\beta_1 + Log(BVE/MVE)_{i,t-3}\beta_3 + Log(MVE)_{i,t-3}\beta_4 + r_{6,6} + u_{i,t} + \varepsilon_{i,t} \quad (5.6.7)$$

Model 8

$$Y_{i,t} = \beta_0 + NCIC_{i,t-3}\beta_1 + NI_{i,t-3}\beta_2 + Log(BVE/MVE)_{i,t-3}\beta_3 + Log(MVE)_{i,t-3}\beta_4 + r_{6,6} + u_{i,t} + \varepsilon_{i,t} \quad (5.6.8)$$

Model 9

$$Y_{i,t} = \beta_0 + CFOA_{i,t-3}\beta_1 + FinAct_{i,t-3}\beta_2 + InvAct_{i,t-3}\beta_3 + Log(BVE/MVE)_{i,t-3}\beta_4 + Log(MVE)_{i,t-3}\beta_5 + r_{6,6} + u_{i,t} + \varepsilon_{i,t} \quad (5.6.9)$$

Model 10

$$Y_{i,t} = \beta_0 + NCIC_{i,t-3}\beta_1 + FinAct_{i,t-3}\beta_2 + InvAct_{i,t-3}\beta_3 + Log(BVE/MVE)_{i,t-3}\beta_4 + Log(MVE)_{i,t-3}\beta_5 + r_{6,6} + u_{i,t} + \varepsilon_{i,t} \quad (5.6.10)$$

Model 11

$$Y_{i,t} = \beta_0 + OP_{i,t-3}\beta_1 + FinAct_{i,t-3}\beta_2 + InvAct_{i,t-3}\beta_3 + Log(BVE/MVE)_{i,t-3}\beta_4 + Log(MVE)_{i,t-3}\beta_5 + r_{6,6} + u_{i,t} + \varepsilon_{i,t} \quad (5.6.11)$$

Model 12

$$Y_{i,t} = \beta_0 + NI_{i,t-3}\beta_1 + FinAct_{i,t-3}\beta_2 + InvAct_{i,t-3}\beta_3 + Log(BVE/MVE)_{i,t-3}\beta_4 + Log(MVE)_{i,t-3}\beta_5 + r_{6,6} + u_{i,t} + \varepsilon_{i,t} \quad (5.6.12)$$

Model 13

$$Y_{i,t} = \beta_0 + CFOA_{i,t-3}\beta_1 + NCIC_{i,t-3}\beta_1 + OP_{i,t-3}\beta_1 + NI_{i,t-3}\beta_1 + FinAct_{i,t-3}\beta_2 + InvAct_{i,t-3}\beta_3 + Log(BVE/MVE)_{i,t-3}\beta_4 + Log(MVE)_{i,t-3}\beta_5 + r_{6,6} + u_{i,t} + \varepsilon_{i,t} \quad (5.6.13)$$

Model 14

$$Y_{i,t} = \beta_0 + CFOA_{i,t-3}\beta_1 + NI_{i,t-3}\beta_1 + FinAct_{i,t-3}\beta_2 + InvAct_{i,t-3}\beta_3 + Log(BVE/MVE)_{i,t-3}\beta_4 + Log(MVE)_{i,t-3}\beta_5 + r_{6,6} + u_{i,t} + \varepsilon_{i,t} \quad (5.6.14)$$

where $Y_{i,t}$ is the six-month ahead return for share i, β_0 is the intercept, β_i is the slope coefficient for each respective factor, $CFOA$ is Cash Flow from Operating Activities deflated by total assets, $NCIC$ is the Net Change in Cash figure deflated by total assets, OP is Operating Profit deflated by total assets, NI is Net Income deflated by total assets, $FinAct$ is Cash Flow from Financing Activities deflated by total assets, $InvAct$ is Cash Flow from Investing Activities deflated by total assets, $Log(BVE/MVE)$ is the logarithm of the ratio of book value of equity to market capitalisation of equity, $Log(MVE)$ is the logarithm of the market capitalisation of equity, $r_{6,6}$ is the six-month return of the share lagged by six months, $u_{i,t}$ is the unobserved individual effect and $\varepsilon_{i,t}$ is the error term.

5.7 Model Diagnostics

The regression models in Section 5.6 rely upon underlying properties and distributions of the data. The Ordinary Least Squares (OLS) assumptions are discussed below in context of the results in Appendix G.

5.7.1 Multicollinearity

Multicollinearity occurs when high correlations between independent variables are present. This may increase the variance of the coefficient estimates and result in sampling errors (Keller and Warrack, 2003). The resulting coefficients of the independent variables lead to inaccurate and redundant inferences in the OLS model. As per Cohen et al. (2013), this study used the variance inflation factor (VIF) to detect multicollinearity. The VIF quantifies the severity of collinearity and can detect if a regressor is correlated with a linear combination of other regressors. The results in Table G3 (Appendix G) suggest that VIF values are less than 10 for the majority of models, indicating that multicollinearity is not a pervasive issue in the regressions (Lin, 2008). As supplementary evidence, the correlation matrix in Table 5.3 was used to detect similar variables which provided no additional insight into the comparison of cash and accrual measures.

Greene (2008) suggests that solutions often include dropping variables until no further multicollinearity occurs. However, this comes at the risk of specification errors in the model. Due to the elimination of redundant variables in Section 5.3 and evidence by Grewal et al. (2004), suggesting that adverse effects of multicollinearity occur only at extreme levels, this assumption was deemed not to be violated.

5.7.2 Autocorrelation

This refers to the OLS assumption that regression errors are independently distributed. Autocorrelation occurs when disturbances are correlated across time, and results in inefficient coefficient estimates in addition to biased standard errors (De Jager, 2008). The Durbin-Watson test was used to detect the presence of autocorrelation in the residuals. Table G1 in Appendix G suggests that autocorrelation is not present in any of the models.

5.7.3 Homoscedasticity

OLS regressions assume homoscedasticity of the underlying data. This suggests that dependent variables display an equal level of variability across values of an independent variable. Coefficient estimates become inefficient when the variance of the error term is non-constant, referred to as heteroscedasticity (Gurujarati and Porter, 2010). This study used a Breusch-Pagan test to detect heteroscedasticity, shown in Table G1. None of the models showed significant test statistics – the null hypothesis of homoscedasticity is not rejected and heteroscedasticity therefore not a concern for the underlying data.

5.7.4 Normality

OLS regressions require the error residuals to be normally distributed. According to Razali and Wah (2011), the Shapiro-Wilk test provides the most power over a range of distributions and sample sizes, as compared to the Anderson-Darling, the Lilliefors and the Kolmogorov-Smirnov tests. The results from the Shapiro-Wilk test for normality are shown in Table G2 in the Appendix. All the variables show significant test statistics, resulting in the null hypothesis of a normally distributed population to be rejected. According to Grajales et al. (2013), if the other OLS assumptions are not violated, variables that are not normally distributed can still provide unbiased and consistent estimators. Furthermore, the non-normality issue is mitigated if the sample size is large enough to apply the Central Limit Theorem. This means the distribution of the sample means approach normality as the sample size increases, regardless of what the underlying distribution is. As this study uses a sample size of over 1500, non-normality is not considered to be a material issue.

5.8 Summary and Conclusion

This chapter discussed the methodology and research procedures employed in the study. Adjustments with regards to trailing and lagging variables were made, the use of appropriate deflators was discussed, and redundant measures were removed based on a correlation matrix.

The results of relevant statistical tests informed the use of a fixed-effects panel regression in the study. Fourteen regression models use varying combinations of cash, accrual and control variables to assess predictive ability. As regression models rely on underlying properties and distributions of the data, additional statistical tests were carried out and adjustments were made accordingly. The following chapter will discuss and compare the results from these regressions and consider the outcomes in context of the research objectives.

Chapter 6: Results

This chapter presents the results found when investigating the ability of cash flow and accrual measures to predict share returns in the South African market. This is followed by a discussion on the extent of predictive ability and a comparison of various measures used in the regression. The results of 85 JSE-listed companies are analysed using the *EViews* statistical package, with fundamental and return data obtained from *Datastream* after applying adjustments outlined in Chapter 4 and Chapter 5. The sample period in this study is June 2008 to June 2018.

Section 6.1 presents and analyses the results of fourteen regressions in context of the research objectives in Chapter 1. Section 6.2 summarises the findings and concludes.

6.1 Empirical Results

The fixed effects panel regression results were obtained by regressing six-month ahead returns against cash flow and accounting variables (deflated by total assets). The regressions were conducted using semi-annual (six monthly) data for 85 JSE-listed shares over the ten-year sample period. Table 6.1.1 shows the regression results for Regression Models 1 to 4, Table 6.1.2 shows the results for Models 5 to 8, Table 6.1.3 shows the results for Models 9 to 12, and Table 6.1.3 shows the results for Models 13 and 14. The coefficient, t-statistic and level of significance are shown, together with summary statistics for each panel regression.

6.1.1 Regression Models 1 to 4

Table 6.1.1 (see next page) shows the results for regressions performed on each independent variable individually.

When regressed against returns individually, the accounting measures display higher significance than the cash flow measures, although all have significant positive coefficients. The positive CFOA coefficient supports the findings by Linvat and Zarowin (1990) where the cash component of operating cash flows is found to improve association with annual returns in the U.S., and implies that cash inflows from operations are positively associated with returns on the JSE. In a South African context, CFOA has been shown to provide information content to model share prices on the JSE, but not returns (van Niekerk, 1992). Net Income (NI) and Operating Profit (OP) are the most significant independent variables, and show the strongest association with six-month ahead returns, with higher R^2

Table 6.1.1 Results for Regression Models 1 to 4

Fixed Effects Panel	Regression Models			
	(1)	**(2)**	**(3)**	**(4)**
Variables				
CFOA	0.3562***			
	(2.8345)			
NCIC		0.2755**		
		(2.2282)		
OP			1.0073***	
			(3.9546)	
NI				1.3481***
				(4.8062)
Controls				
Log(BVE/MVE)	-0.0598***	-0.0685***	-0.0226	-0.0281
	(-2.9678)	(-3.5544)	(-0.8495)	(-1.1091)
Log(MVE)	-0.0879***	-0.0923***	-0.0764***	-0.0854***
	(-7.3253)	(-7.6692)	(-5.5186)	(-6.7773)
$R_{6,6}$	-0.0696**	-0.0687*	-0.0811**	-0.0856**
	(-1.9746)	(-1.9553)	(-2.3495)	(-2.5169)
Summary Stats				
R^2	0.1124	0.1106	0.1239	0.1301
Adjusted R^2	0.0588	0.0568	0.0710	0.0775
F-Statistic	2.0970***	2.0563***	2.3418***	2.4751***
No. obs.	1546	1544	1546	1546

* Statistical significance <0.10, ** Statistical significance <0.05, *** Statistical significance <0.01

and adjusted R^2 values for Regressions 3 and 4. Sloan (1996) argues that accruals negatively predict returns, but both accrual variables displayed positive coefficients in the panel regressions.

NI was able to explain the most variation in share returns on the JSE, compared to CFOA, NCIC and OP. This supports the finding by Fama and French (2006) that earnings contain information that is able to explain returns. Furthermore, Chan, Lokonishok and Sougiannis

(2001) and Eisfeldt and Papanikolaou (2013) argue that items in the Income Statements themselves have explanatory power. That is, incomes and expenses from non-operating business activities have standalone predictive power, and investors incorporate these ancillary items into their assessment of the final 'claim' on the business. In contrast, Ball et al. (2015) find that operating profitability explains the cross section of returns better than "bottom-line" net income. Ball et al. (2015) claim that the operating profit measure has more information on the quality of the company's primary business activities, thus impacting share price adjustments and valuations more. The results suggest that this is not the case on the JSE. Given that NI was more significant than OP, this may indicate that the entirety of a company's business activities contains more information about future returns than operations alone.

Net Change in Cash (NCIC) shows the lowest significance level of the group. As an aggregate cash flow measure that summarises operating, investing and financing activities, it is limited in its predictive ability for six-month returns on the JSE. The sign of the coefficient suggests a direct relationship with cash and returns. A positive change in cash at period or interim-end results in positive returns, and vice-versa. The cash management of a company therefore has a degree of impact on market pricing of the share.

6.1.2 Regression Models 5 to 8

The models in Table 6.1.2 (see next page) were constructed to assess the impact of combining cash and accrual variables and the extent to which predictive power is affected.

CFOA shows a lower level of significance when the OP variable is added to the regression. Unlike Model 1, which only includes CFOA, Model 5 results in a non-significant Log(BVE/MVE) value. This suggests that the operating profit metric captures some of the 'value perception' that the cash flow from operating activities metric does not – that is, the value investors place on a firm and the perception of its ability to generate further value from those assets.

The model with the second strongest predictive power is Model 6, which includes a combination of the strongest individual variable, NI, and CFOA. Compared to Model 1 in Table 6.1.1, the combined model is able to explain a greater portion of returns – showing a higher R^2 and a more significant F-statistic. This outcome is contrary to Ball et al. (2016), who find that stripping accruals from the cash profitability measure results in greater prediction in the cross section of average returns. It appears that combining cash and accrual metrics, as opposed to stripping them out, adds to the predictive ability of regression models.

Table 6.1.2 Results for Regression Models 5 to 8

Fixed Effects Panel	Regression Models			
	(5)	**(6)**	**(7)**	**(8)**
Variables				
CFOA	0.2486*	0.2645**	0.2930*	
	(1.9361)	(2.0681)	(1.8260)	
NCIC			0.0968	0.2639**
			(0.5989)	(2.0758)
OP	0.9409***			
	(3.6360)			
NI		1.2911***		1.3397***
		(4.6169)		(4.8575)
Controls				
Log(BVE/MVE)	-0.0209	-0.0248	-0.0618***	-0.0303
	(-0.7909)	(-0.9770)	(-2.9765)	(-1.2069)
Log(MVE)	-0.0744***	-0.0825***	-0.0887***	-0.0856***
	(-5.5010)	(-6.7021)	(-7.4331)	(-6.8784)
$R_{6,6}$	-0.0799**	-0.0845**	-0.0690**	-0.0841**
	(-2.3127)	(-2.4797)	(-1.9663)	(-2.4867)
Summary Stats				
R^2	0.1262	0.1326	0.1125	0.1327
Adjusted R^2	0.0727	0.0795	0.0582	0.0796
F-Statistic	2.3618***	2.5015***	2.0714***	2.4999***
No. obs.	1546	1546	1544	1544

* Statistical significance <0.10, ** Statistical significance <0.05, *** Statistical significance <0.01

The South African financial markets may therefore gather information from both cash and accounting values, which are then reflected in share price movements. The two 'bottom-line' cash and accounting variables, NCIC and NI, resulted in a regression with the strongest explanatory power amongst the models in Table 6.1.2. Although the NI measure had the strongest predictive power out of the independent variables when regressed individually, Model 8 suggests that there is additional value in adding an aggregate cash flow figure.

6.1.3 Regression Models 9 to 12

All the models in Table 6.1.3 (see next page) include financing and investing activities as independent variables, in line with the methodology of Foerster, Tsagarelis and Wang (2017). The four key measures are again regressed individually with the addition of these two cash flow measures.

The addition of the two cash flow metrics resulted in all the measures becoming more statistically significant. Again, the NI variable is the most statistically significant in the group. The CFOA model is marginally better at explaining share return variation despite including all the major sections from a cash flow statement in a regression. The improvement suggests that combining multiple cash flow measures can better predict share price behaviour. A similar finding by Van Niekerk (1992) demonstrates that both cash flow from operating activities and cash flow from investing activities have incremental information content to model share prices on the JSE, but not returns.

NCIC is an aggregate cash flow figure already including the net effect of cash flow from investing activities and cash flow from financing activities. Consequently, Model 10 had the lowest R^2 and adjusted R^2 values amongst the models in Table 6.1.3. Neither financing activities nor investing activities are significant in any of the regressions performed, except financing activities (Fin Act) in Model 10. A similar result by Bernard and Stober (1989) suggests that separating cash flows from operating and cash flows from financing activities improved association with annual returns. Linvat and Zarowin (1990) also find evidence of significantly improved association with these metrics, and no association with cash flows from investing activities. However, the result for Fin Act in Model 10 is not extremely significant and is not persistent across models. The interpretation is also limited because of the nature of the NCIC measure.

Table 6.1.3 Results for Regression Models 9 to 12

Fixed Effects Panel	Regression Models			
	(9)	**(10)**	**(11)**	**(12)**
Variables				
CFOA	0.4370***			
	(3.4129)			
NCIC		0.3917***		
		(2.9241)		
OP			1.1127***	
			(4.4899)	
NI				1.4517***
				(5.1583)
Fin Act	0.1044	-0.2735*	0.0825	0.0754
	(0.7031)	(-1.8284)	(0.5740)	(0.5057)
Inv Act	0.1377	-0.2585	0.0751	0.0279
	(0.6140)	(-1.0869)	(0.3392)	(0.1178)
Controls				
Log(BVE/MVE)	-0.0616***	-0.0628***	-0.0182	-0.0244
	(-3.0415)	(-3.1077)	(-0.7057)	(-0.9636)
Log(MVE)	-0.0896***	-0.090***	-0.0764***	-0.0863***
	(-7.6778)	(-7.7724)	(-5.6966)	(-7.0126)
$R_{6,6}$	-0.0696*	-0.0691*	-0.0799**	-0.0847
	(-1.9489)	(-1.9333)	(-2.2765)	(-2.4549)
Summary Stats				
R^2	0.1199	0.1185	0.1318	0.1381
Adjusted R^2	0.0645	0.0629	0.0772	0.0838
F-Statistic	2.1653***	2.1344***	2.4142***	2.5466***
No. obs.	1522	1520	1522	1522

* Statistical significance <0.10, ** Statistical significance <0.05, *** Statistical significance <0.01

6.1.4 Regression Models 13 and 14

Regression Model 13 in Table 6.1.4 (see next page) combines all the independent variables in the study, and Model 14 includes the combination of a cash and accounting variable in an attempt to find the strongest regression model for predicting share returns. Both regressions display higher R^2 and adjusted R^2 values, with more significant F-statistics than any of the fourteen regressions performed. Combining all variables results in non-significant values in all but two variables – CFOA and NI. This effect is similar to Model 6 in Table 6.1.2, where the combination of these two measures results in the most explanatory power amongst those regressions.

Table 6.1.4 Results for Regression Models 13 and 14

Fixed Effects Panel	Regression Models	
	(13)	**(14)**
Variables		
CFOA	0.6950*	0.3725***
	(1.8330)	(2.8631)
NCIC	-0.3544	
	(-0.8979)	
OP	0.2746	
	(0.7894)	
NI	1.1914***	1.4032***
	(2.8479)	(5.0794)
Fin Act	0.5554	0.2265
	(1.4435)	(1.4716)
Inv Act	0.5154	0.1707
	(1.1875)	(0.7317)
Controls		
Log(BVE/MVE)	-0.0166	-0.0226
	(-0.6390)	(-0.9060)
Log(MVE)	-0.0807***	-0.0839***
	(-6.3309)	(-7.0168)
$R_{6,6}$	-0.0836**	-0.0833**
	(-2.4292)	(-2.4193)
Summary Stats		
R^2	0.1434	0.1457
Adjusted R^2	0.0876	0.0880
F-Statistic	2.5676***	2.6131***
No. obs.	1520	1522

* Statistical significance <0.10, ** Statistical significance <0.05, *** Statistical significance <0.01

In addition, the inclusion of Fin Act and Inv Act resulted in higher R^2 and adjusted R^2 values in Section 6.1.3. For this reason, Model 14 attempts to create the panel regression with the most potential to predict future returns. In terms of Model 8 appearing to be the strongest regression in Table 6.1.2, Model 14 indicates that Inv Act, Fin Act and CFOA contain more information on share returns when separated and combined with the NI variable. Contrary to Linvat and Zarowin (1990), separating net income and cash from operations does contribute to future share returns, and more so than net income alone. Of the fourteen regression models, the independent variables in Model 14 explain the largest proportion of six-month ahead returns on the JSE. Surprisingly, the NI is consistently the most significant measure in this regard**.** Fama and French (2006) demonstrate that current earnings can be used as a proxy for future profitability. The findings across all regressions suggest that this is also the case in an emerging market. A similar result in a less developed market context is obtained by Hunjra et al. (2014), where profit after tax displayed a significant impact of share performance in the Pakistan financial market.

Dechow (1994) argues that cash flow metrics are flawed when companies have larger cash requirements with more volatility, which reduces their predictive ability. The strength of the NI variable supports the conclusion by Rayburn (1986) that cash flows should not be preferred to accruals. The ‘accrual anomaly’ researched by Sloan (1996) may therefore be present in the South African equity market. The consistently high significance of the NI measure suggests that investors do in fact fixate on profitability, and future share price movements reflect this.

Similar to the outcome in Section 6.1.2, the cash flow metric reflecting the business activities of the company and a ‘bottom-line’ accounting measure explains variation in share returns better than any other model in the study. Both NI and CFOA in Model 14 are highly significant and show positive coefficients. The results corroborate evidence by Quirin, O’Bryan and Wilcox (1999), who use U.S. company financials from 1972 to 1981 to demonstrate that when cash flows and earnings are positive, returns react more significantly.

6.1.5 Control Variables

The control variables displayed negative coefficients in all panel regressions carried out. The most significant control variable is Log(MVE), which attempts to capture the size effect. Contrary to Muller and Ward (2013), the 85 shares in the sample showed a significant size effect. Similar to van Rensburg and Robertson (2003), there is evidence of an anomaly related to the market capitalisation of a company. The valuation and trading volume of a share had a significant effect on share returns in this study, supporting the inclusion of a size control variable, as there appears to be a substantial negative relationship between the size of a firm and the six-month returns.

Following the methodology of Foerster, Tsagarelis and Wang (2017) and Fama and French (1992), the Log(BVE/MVE) control variable was included in all panels. Auret and Sinclaire (2006) also add BTM to the size and PE effects in their investigation of JSE anomalies. The variable was shown to subsume the effect of size but was not significant enough to explain returns in the stock market. The Log(BVE/MVE) control variable is significant where cash metrics are included as variables, and is no longer significant if an accrual variable is added to the regression model. Log(BVE/MVE) displayed a high level of significance under the cash flow variables in Models 1 and 2. Interestingly, it was no longer significant when regressed with OP and NI. This suggests that the cash flow measures fail to capture some information regarding the perceived risk of a share. Further evidence of this is the high significance obtained in Model 7, where only CFOA and NCIC are included as independent variables, and Models 8 and 9, which only comprise cash metrics. The accrual variables appear to contain information which this control consistently captured in regressions where cash flow measures were included.

The control variable for momentum, $r_{6,6}$, showed a relatively high level of significance in almost all regression models. This supports the findings by Page, Britten and Auret (2013), suggesting the presence of a momentum anomaly on the JSE and van Rensburg (2015), who finds that momentum forms some representation of style-based risk in South African equity markets. The control variable was only non-significant in Model 12, where NI is combined with Fin Act and Inv Act.

6.2 Summary and Conclusion

This chapter presented the empirical results for cash flow and accrual variables and investigated their ability to predict returns. Results from the fourteen panel regressions showed that traditional profitability measures provided more predictive ability than cash flow measures.

In all regressions, CFOA, NCIC, OP and NI displayed positive coefficients, suggesting a direct relationship with returns and an increase in profit or cash, as represented by the explanatory variables. Three control variables were included in all regressions, and these all displayed negative coefficients. Log(MVE) displayed the highest level of significance throughout, indicating the existence of a size effect during the sample period. The presence of a momentum effect was also found to be significant, though to a lesser extent. The Log(BVE/MVE) control variable was only significant if regressions excluded accrual measures, which may suggest that cash flow figures fail to capture information regarding the riskiness of a share.

Overall, the accrual-based measures were stronger predictors than the cash flow measures, with the aggregated cash figure, NCIC, displaying the lowest significance. The more direct cash flow measure, CFOA, was found to be superior, as it contains movements related to a stock's primary income-generating activities. Interestingly, the accrual variables suggest the opposite – the NI figure was found to be more significant than OP. While NI was the most significant variable individually, a combination of NI and CFOA in a regression model revealed the strongest statistical indicators. Furthermore, adding Fin Act and Inv Act as supplementary independent variables provided incremental predictive power. The most significant regression model consisted of three cash flow variables and one accrual variable. This model is the therefore the strongest candidate for predicting returns, suggesting some degree of inefficiency in the JSE market.

Therefore, the results of the study suggest the existence of both cash and accrual anomalies, in addition to confirming the 'size' and 'momentum' anomalies that have been previously uncovered in South African financial literature.

Chapter 7: Summary and Conclusions

Given the limited existing research on the predictive ability of cash flows with regards to equity returns, the objective of this study was to investigate whether cash flows can predict future share returns for companies on the JSE, and to what extent investors can use this information to generate profits. A further objective was to determine, within the context of the JSE, whether different components of cash flows, such as cash flows from operating activities, cash flows from financing activities, and cash flows from investing activities, differ in their ability to predict returns. This was then extended to the comparison of calculated cash flow measures to traditional profitability and free cash flow measures commonly used to model future share returns.

Underlying this research is an inherent test of the level of market efficiency on the JSE. In other words, if cash flows were to have predictive ability, then the market would either be inefficient, or there would be a prevailing model misspecification. Finally, this research set out to enhance existing international and South African literature by testing previously untested cash-based measures on an untested sample of shares in an emerging market.

Section 7.1 presents a summary of the empirical results in context of the study's research objectives, Section 7.2 suggests extensions for future research, and Section 7.3 concludes.

7.1 Summary of Results

The dataset used to investigate the research objectives consists of company financial statement figures and total returns. The data was collected from *Datastream* on a semi-annual basis over an eleven-year period from June 2007 to August 2018, and the analysis was conducted from March 2008 to March 2018 due to the trailing and lagging variables used.

In order to achieve robust results, the sample data was adjusted for comparability by aligning interim periods of companies with different year-end dates and adjusting for any changes in financial year-ends. Liquidity was addressed by gathering constituent lists at different points in the sample period and ranking by market value of equity at each point. A two-tailed winsorisation at 1% and 99% was carried out to address the effects of outliers in the data.

This study considered possible sources of bias and adjusted the data to mitigate the effects. Due to the unique data and research procedure being carried out on a current sample period, data-snooping was mitigated. Look-ahead bias was addressed with the use of trailing three-month fundamental values – updated as company interim results were released. Finally, survivorship bias is often omitted from research methodologies, leading to inaccurate interpretations of results. The data in this study was adjusted for survivorship bias such that delisted shares were included in the sample during the period of their listing. The historically-ranked constituent lists were compared such that constituents appearing at least once in the top 100 were included in the final constituent list.

The sample used to investigate the predictability of share returns comprises 85 shares listed on the JSE. Testing procedures were conducted to inform the use of a fixed-effects panel regression, and the following OLS assumptions were considered and adjusted for: multicollinearity, autocorrelation, homoscedasticity and normality. A total of fourteen regressions were carried out on six independent variables and three control variables.

In terms of the first objective, accrual-based measures were superior to cash flow measures in predicting returns. Initial regressions used individual variables, and all were significant at the 5% level. Net Income (NI) was consistently the strongest and most significant independent variable in all the panel regressions, while Net Change in Cash (NCIC) had the least predictive power. Cash Flow from Operations (CFOA) was significant at the 1% level, however inferior to the equivalent accrual measure, Operating Profit (OP), which displayed a higher t-statistic. Furthermore, OP subsumed much of the predictive ability of CFOA when regressed together. All four of the explanatory variables displayed positive coefficients, suggesting that an increase in profitability or cash has an impact on future share prices and returns. It appears that cash flows may indicate the health of a business, but net income represents the claim that investors can expect to have on the firm’s profits – driving the prices to adjust.

Regarding the second objective, Cash Flow from Financing Activities (Fin Act) and Cash Flow from Investing Activities (Inv Act) were added to regression models for the comparison of different cash flow measures. The results showed that CFOA was consistently a stronger predictor of share returns. Fin Act and Inv Act were non-significant, suggesting that cash movements from business operations contains more incremental information able to explain returns, when compared to the financing and investing cash movements.

In response to the third objective, traditional accrual-based measures provided more predictive power than cash flow measures. The more direct cash flow measure, CFOA, was found to be superior to NCIC, Inv Act and Fin Act, as it contains movements related to primary income-generating activities of a business. Interestingly, the accrual variables displayed the opposite – the NI figure was found to be more significant than OP. The addition of two cash flow metrics (Fin Act and Inv Act), as per the methodology of Foerster, Tsagarelis, and Wang (2017), improved the significance levels of all measures. An interesting outcome was that although profitability measures were more significant, the model able to explain the most variation in share return on the JSE consisted of NI and CFOA, with the addition of Inv Act and Fin Act. In this combination, all variables were more significance and provided the highest R^2 value (15%) and most significant F-statistic (at the 1% level). The results suggest that implementing measures representing an aggregate picture of a business, in cash *and* accrual terms, can explain the greatest proportion of share return variation for a six-month holding period.

In terms of the level of market efficiency on the JSE, there is evidence from the analysis to support the existence of cash and accrual anomalies. The significance levels and predictive ability suggests that variation in future returns can be explained by prior movements in company financial figures. Therefore, there is evidence to reject a strong-form level of market efficiency and support the argument for semi-strong form market efficiency of the JSE. The analysis on control variables also supports evidence of a size effect, represented by Log(MVE), and a momentum effect, represented by $r_{6,6}$.

The final objective was achieved by carrying out a unique research methodology on a previously untested sample of shares and variables. The thesis provided new insights into a limited research area and contributed to literature around the return predictability of cash flows.

7.2 Suggestions for Extension

The topic could be further investigated by improving aspects of the sample, accounting variables, adjustments to data, and methodologies employed – all limitations of this study. More robust results could be achieved by extending the sample period from to 20 years, or 40 six-month periods. Kruger (2011) argues that evidence of return predictability is generally not robust to changes in market regime, and therefore investigates 20 firm characteristics on the JSE over a market crisis period 2002 to 2009. The results show that the cash flow-to-price ratio is the only consistent predictor of returns on the All-Share Index. This study was limited by the sample period, and investigating the cash flow predictability for an out-of-sample period in addition to a comparison of different holding periods would be valuable to further research.

Though financial and technology companies were largely excluded, their inclusion could provide an interesting insight into return predictability given the structural and regulatory differences impacting cash flows. As a result, a comparison of findings by industry could provide additional insight. *Datastream* converts financial data reported in another currency using average monthly and period-end exchange rates. More accurate currency conversion processes (specifically the dates) would lead to a more robust set of results. While the study made a deliberate effort to consider survivorship bias using historical constituent lists, this should be further revised by capturing members on an annual or monthly basis.

In terms of the variables used, additional profitability and cash flow measures may provide interesting results. Cash taxes paid, Novy-Marx's (2013) free cash flow measure (net income plus depreciation minus working capital change minus capital expenditures), research and development expenses and cash vs. credit sales are considered valuable decision-making items to investors and may provide further insight into return behaviour. The inclusion of interaction terms is also recommended. Although firms listed on the JSE are required to adhere to IFRS, many back-end calculations on *Datastream* use GAAP adjustments. Many functions for IFRS-adjusted data also return incomplete or missing values. In addition to addressing this, updating the analysis with the ongoing revisions to IFRS standards is also suggested.

As argued by Sloan (1996), the cash-based component of earnings is of a higher quality and contains more information that the accruals component. It is of interest to the author to investigate which specific components of cash have the strongest predictive ability. IFRS allows for presentation of Cash Flow from Operating Activities via the Direct or Indirect method. The option is intended to cater towards businesses with varying levels of cash flows, assets and transactions. This study may be extended by creating a direct method template to organise business activities into sections such that the value-generating aspects can be isolated and tested. In addition, different methodologies should be used in assessing this anomaly. Portfolio-sorted regressions (including a long-short portfolio), the Fama and French (1993) Three-Factor Model, Carhart's (1997) Four-Factor Model, the Fama and French (2015) Five-Factor Model and a sector neutral analysis, are suggested as alternative approaches to the research question. Investigation into the topic could also make use of more advanced modelling techniques such as LSTM (long short-term memory) neural networks. Another fascinating extension to the return predictability of cash flows would be an analysis of emerging economies and a comparison to more developed (and informationally efficient) financial systems.

7.3 Conclusion

This study adds to the existing body of knowledge around asset pricing, the presence of anomalies in financial markets, return prediction, and the value of using cash flow information in the fundamental investment process. The research extends the work of Foerster, Tsagarelis and Wang (2017) and applies a similar investigation to a developing market. It contributes to the understanding of how equity markets process financial statement figures and reflect that information in price-adjustments. Accrual-based measures were able to provide more explanatory power than cash flow measures, but there was evidence of additional predictive power in a combination of traditional profitability as well as cash flow figures. These results serve as evidence for return prediction and in particular, the value that cash flow-based anomalies add to predicting price movements and stock returns. As one of the few research studies to investigate cash flow's predictive ability in a less developed market, and to the author's knowledge, the first on the JSE, the study opens a new direction for empirical research. Finally, the findings suggest that some variation in future returns can be explained by prior movements in financial statement figures, which provides evidence to reject a strong-form level of market efficiency and support the argument for a semi-strong form level of market efficiency on the JSE.

References

Ali, A. & Pope, P. (1995). The Incremental Information Content of Earnings, Fund Flow and Cash Flow: UK Evidence. *Journal of Business Finance and Accounting.* 22(1)*:*19-34.

Ali, A. (1994). The Incremental Information Content of Earnings, Working Capital from Operations and Cash Flows. *Journal of Accounting Research.* 32(1):61-74.

Arellano, M. (2003). *Panel Data Econometrics*. Oxford University Press.

Atchison, M.D, Butler, K. & Simonds, R. (1987). Nonsynchronous Trading Activity and Market Index Autocorrelation. *Journal of Finance*. 62:111-118.

Auret, C. J. & Sinclaire, R. (2006). Book-to-Market Ratio and Returns on the JSE. *Investment Analysts Journal.* 63:31-38.

Bachelier, L. (1900). Théorie de la speculation. *Annales Scientifiques de l'École Normale Supérieure*. 3(17):21-86.

Baltagi, B. H. (2005). *Econometric Analysis of Panel Data.* 3rd ed. Wiley. New York.

Baltagi, B. H. & Song, S. H. (2006). Unbalanced Panel Data: A Survey. *Statistical Papers.* 47:493-523.

Ball, R. & Brown, P. (1968). An Empirical Evaluation of Accounting Income Numbers. *Journal of Accounting Research.* 7:300-323.

Ball, R. (1978). Anomalies in Relationships Between Securities' Yields and Yield Surrogates. *Journal of Financial Economics. 6(2):*103-126.

Ball, R., Gerakos, J., Linnainmaa, J. & Nikolaev, V. (2015). Deflating Profitability. *Journal of Financial Economics*. 117(2):225-248.

Ball, R., Gerakos, J., Linnainmaa, J. & Nikolaev, V. (2016). Accruals, Cash Flows, and Operating Profitability in the Cross Section of Stock Returns. *Journal of Financial Economics.* 121(1):28-45.

Banz, R. W. & Breen, W. J. (1986). Sample-Dependent Results Using Accounting and Market Data: Some Evidence. *Journal of Finance*. 41(4):779-793.

Barth, M. & Beaver, W. (2000). Accruals, Cash Flows and Equity Values. *Review of Accounting Studies.* 4:205-229.

Basiewicz, P.G. & Auret, C. J. (2009). Another Look at the Cross-Section of Average Returns on the JSE. *Investment Analysts Journal*. 69(1):23-38.

Basiewicz, P.G. & Auret, C. J. (2010). Feasibility of the Fama and French Three Factor Model in Explaining Returns on the JSE. *Investment Analysts Journal*. 39(71):13-25.

Basu, S. (1977). The Investment of Common Stocks in Relation to their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis. *Journal of Finance*. 32(3):663-682.

Basu, S. (1983). The Relationship Between Earnings Yield, Market Value, and Return for NYSE Common Stocks: Further Evidence. *Journal of Financial Economics*. 12(1):129-156.

Beaver, W.H. & Dukes, R. E. (1972). Interperiod Tax Allocation Earnings Expectations and the Behaviour of Security Prices. *Journal of Finance.* 27:663-682.

Beisland, L. A. (2009). A Review of the Value Relevance Literature. *The Open Business Journal*. 2:7-27.

Bernard, V. & Stober, T. (1989). The Nature and Amount of Information in Cash Flows and Accruals. *The Accounting Review*. 64(4):624–652.

Bhandari, L.C. (1988). Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence. *Journal of Finance*. 43(2):507-528.

Black, F. (1972). Capital Market Equilibrium with Restricted Borrowing. *The Journal of Business*. 45(3):444-455.

Black, F. (1993). Estimating Expected Return. *Financial Analysts Journal*. 49(5):36-38.

Blume, M. E. & Stambaugh, R. F. (1983). Biases in Computed Returns: An Application to the Size Effect. *Journal of Financial Economics*. 12(3):387-404.

Board, J. & Day, J. (1989). The Information Content of Cash Flow Figures. *Accounting and Business Research.* 20(77):3-11.

Bowen, R. M., Burgstahler, D. & Daley, L. A. (1986). Evidence on the Relationships Between Earnings and Various Measures of Cash Flow. *The Accounting Review*. 61(4):713-725.

Bowen, R. M., Burgstahler, D. & Daley, L. A. (1987). The Incremental Information Content of Accrual Versus Cash Flows. *The Accounting Review*. 62(4):723-747.

Bowie, D.C. (1994). Thin Trading, Non-Normality and the Estimation of Systematic Risk on Small Stock Markets. Unpublished PhD Thesis. University of Cape Town.

Breusch, T.S. & Pagan, A.R. (1980). The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics. *The Review of Economic Studies*. 47(1):239-253.

Brown, K. C., Harlow, W. V. & Tinic, S. (1988). Risk Aversion, Uncertain Information, and Market Efficiency. *Journal of Financial Economics*. 22(2):355-385.

Brown, K. C., Harlow, W. V. & Tinic, S. M. (1988). Risk Aversion, Uncertain Information, and Market Efficiency. *Journal of Financial Economics*. 22(2):355-385.

Campbell, J. Y. & R. J. Shiller, (1988). The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. *Review of Financial Studies*. Oxford University Press for Society for Financial Studies. 1(3):195-228.

Campbell, J. Y. & Vuolteenaho, T. (2004). Bad Beta, Good Beta. *American Economic Review*. 94(5):1249-1275.

Campbell, J. Y. (1991). A Variance Decomposition for Stock Returns. *The Economic Journal*. 101(405):157-179.

Campbell, J. Y., Polk, C. & Vuolteenaho, T. (2010). Growth or Glamour? Fundamentals and Systematic Risk in Stock Returns. *Review of Financial Studies*. 23(1):305-344.

Carhart, M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance*. 52(1):57-82.

Chan, K., Chan, L. K.C., Jegadeesh, N. & Lakonishok, J. (2001). Earnings Quality and Stock Returns. *Journal of Business.* 79(3):1041-1082.

Chan, L. K. C., Hamao, Y. & Lakonishok, J. (1991). Fundamentals and Stock Returns in Japan. *The Journal of Finance*. 46(5):1739-1764.

Chan, L. K. C., Lakonishok, J. & Sougiannis, T. (2001). The Stock Market Valuation of Research and Development Expenditures. *Journal of Finance.* 56(6): 2431-2456.

Charitou, A. & Clubb, C. (1999). Earnings, Cash Flows and Security Returns Over Long Return Intervals: Analysis and UK Evidence. *Journal of Business Finance and Accounting.* 26(2):283-312.

Chatfield, M. (1974). A History of Accounting Thought. Hinsdale, Ill., *Dryden Press*. 28.

Christie, A. (1987). On Cross-Sectional Analysis in Accounting Research. *Journal of Accounting and Economics*. 9:231-258.

Clark, T. S. & Linzer, D. A. (2012). Should I use Fixed or Random Effects? Available: https://static1.squarespace.com/static/58d3d264893fc0bdd12db507/t/58d3f756f5e23112262c5dae/1490286422489/ClarkLinzerPSRM2015.pdf

Cohen, J., Cohen, P., West, S.G. & Aiken, L.S. (2013). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. 6th ed. Routledge.

Damodar N. Gujarati, D.N. & Porter, D.C. (2010). *Basic Econometrics*. 5th ed. McGraw-Hill Irwin.

De Jager, L.M. (1997). An Investigation into the Information Content of Traditional and Cash Flow Ratios of Listed Industrial Companies on the Johannesburg Stock Exchange. Master's Thesis. University of Stellenbosch.

De Jager, P. (2008). Panel Data Techniques and Accounting Research. *Meditari Accountancy Research.* 16: 53-68.

De Villiers, P., Lowings, A., Pettit, T. & Affleck-Graves, J. (1986). An Investigation into the Small Firm Effect on the Johannesburg Stock Exchange. *South African Journal of Business Management*. 17(4):191-195.

DeBondt, W. F. M. & Thaler, R. (1985). Does the Stock Market Overreact? *Journal of Finance*. 40:793-805.

Dechow, P. (1994). Accounting Earnings and Cash Flows as Measures of Firm Performance: The Role of Accounting Accruals. *Journal* of *Accounting and Economics.* 18(1):3-42.

Desai, H., Rajgopal, S. & Venkatachalam, M. (2004). Value-Glamour and Accurals Mispricing: One Anomaly or Two? *The Accounting Review*, 79(2):355–385.

Desai, H., Rajgopal, S., Venkatachalam, M., Dichev, I., Hodder, L., Hogan, C., Jim Jiambalvo, H., Magliolo, J. et al. (2004). Value-Glamour and Accruals Mispricing: One Anomaly or Two? *The Accounting Review.* 79(2):355-385.

Dimson, E. (1979). Risk Measurement When Shares are Subject to Infrequent Trading. *Journal of Financial Economics*. 7:197–226.

Drtina, R. & Largay, J. (1985). Pitfalls in Calculating Cash Flow from Operations. *The Accounting Review*. 60(2):314-326.

Easton, P. D. & Sommers, G. A. (2003). Scale and the Scale Effect in Capital Market-Based Accounting Research. *Journal of Business Finance and Accounting.* 30(1-2): 253-288.

Eisfeldt, A.L. & Papanikolaou, D. (2013). Organization Capital and the Cross-Section of Expected Returns. *The Journal of Finance.* 68(4):1365-1406.

Fama, E. & French, K. (2006). The Value Premium and the CAPM. *Journal of Finance.* 61:2163-2185.

Fama, E. & French, K. (2008). Dissecting Anomalies. *The Journal of Finance*. 63(4):653–1678.

Fama, E. & French, K. (2015). A Five-Factor Asset Pricing Model. *Journal of Financial Economics*. 116(1):1-22.

Fama, E. F. & French, K. R. (1988). Dividend Yields and Expected Stock Returns. *Journal of Financial Economics.* 22(1):3-25.

Fama, E. F. & French, K. R. (1992). Common Risk Factors in the Returns on Stocks and Bonds. J*ournal of Financial Economics*. 33:3-56.

Fama, E. F. & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *Journal of Finance.* 47:654-673.

Fama, E. F. & French, K. R. (1996). Multifactor Explanations of Asset Pricing Anomalies. *Journal of Finance*. 51(1):55-84.

Fama, E. F. & MacBeth, J. D. (1973). Risk, Return and Equilibrium: Empirical Tests. J*ournal of Political Economy.* 81(3):607-636.

Fama, E. F. (1965a). The Behaviour of Stock Market Prices. *Journal of Business.* 38(1):34-105.

Fama, E. F. (1965b). 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.

Financial Accounting Standards Board (1978), Statement of Financial Accounting Concepts No. 1. Objectives of Financial Reporting by Business Enterprises.

Financial Accounting Standards Board (1997), Statement of Financial Accounting Standards No. 95. Statement of Cash Flows.

Financial Accounting Standards Board (2000). Statement of Financial Accounting Concepts No. 7. Using Cash Flow Information and Present Value in Accounting Measurements.

Financial Accounting Standards Board (2018). Statement of Financial Accounting Concepts No. 8. Conceptual Framework for Financial Reporting.

Foerster, S., Tsagarelis, J. & Wang, G. (2017) Are Cash Flows Better Stock Return Predictors Than Profits? *Financial Analysts Journal.* 73(1):73-100.

Fourie, J. (1992). An Essay on the Effect of Cash Flows on Firm Value. Hons. Report. University of Cape Town.

Fraser, E. & Page, M. (2000). Value and Momentum Strategies: Evidence from the Johannesburg Stock Exchange. *Investment Analysts Journal.* 51:25-35.

Freeman, R. N. & Tse, S.Y. (1992). A Nonlinear Model of Security Price Responses to Unexpected Earnings. *Journal of Accounting Research.* 30:185-209.

Garrod, N. & Hadi, M. (1998). Investor Response to Cash Flow Information. *Journal of Business Finance and Accounting.* 25(5):613-630

Graham, B. & Dodd, D. (1940). *Security Analysis*. New York: McGraw-Hill.

Graham, C. M. & Knight, J. R. (2000). Cash Flows vs. Earnings in the Valuation of Equity REITs. *Journal of Real Estate Portfolio Management.* 6:17-26.

Williams, M.N., Grajales, C.A.G. & Kurkiewicz, D. (2013). Assumptions of Multiple Regression: Correcting Two Misconceptions. *Practical Assessment, Research & Evaluation*. 18(11):2.

Greene, W. H. (2008). *Econometric Analysis.* 6th ed. New Jersey: Pearson Prentice Hall.

Grewal, R., Cote, J. A. & Baumgartner, H. (2004). Multicollinearity and Measurement Error in Structural Equation Models: Implications for Theory Testing. *Marketing Science*. 23: 519-529.

Grossman, S. J. & Stiglitz, J. (1980). On the Impossibility of Informationally Efficient Markets. *American Economic Review*. 70(3):393-408.

Gumbi, P. (2013). Predictive Ability of Current Earnings and Cash Flows. Master’s Thesis. University of Pretoria.

Habib, A. (2008). The Role of Accruals and Cash Flows in Explaining Security Returns: Evidence from New Zealand. *Journal of International Accounting, Auditing and Taxation*. 17(1):51-66.

Hackel, K. S., Livnat, J. & Rai, A. (1994). The Free Cash Flow/Small-Cap Anomaly. *Financial Analysts Journal*. 50(5):33-42.

Hackel, K. S., Livnat, J. & Rai, A. (2000). A Free Cash Flow Investment Anomaly. *Journal of Accounting, Auditing and Finance.* 15(1): 1-24.

Haugen, R. A. & Baker, N. L. (1996). Commonality in the Determinants of Expected Stock Returns. *Journal of Financial Economics.* 41(3):401-439.

Hirshleifer, D., Hou, K. & Teoh, S. H. (2009). Accruals, Cash Flows, and Aggregate Stock Returns. *Journal of Financial Economics*. 91(3):389-406.

Hirshleifer, D., Hou, K. & Teoh, S. H. (2009). Accruals, Cash Flows, and Aggregate Stock Returns. *Journal of Financial Economics*. 91(3):389-406.

Hodgson, A. & Stevenson-Clarke, P. (2000). Earnings, Cash Flows and Returns: Functional Relations and the Impact of Firm Size. *Accounting and Finance.* 40(1):51-75.

Hoechle, D. (2007). Robust Standard Errors for Panel Regressions with Cross Sectional Dependence. *The Stata Journal*. 7:281-312.

Hoffman, A. J. (2012). Stock Return Anomalies: Evidence from the Johannesburg Stock Exchange. *Investment Analysts Journal*. (75):21-41.

Holthausen, R. W. & Larcker, D. F. (1992). The Prediction of Stock Returns Using Financial Statement Information. *Journal of Accounting and Economics*. 15(2):373-411.

Hou, K., Karolyi, G. A. & Kho, B. (2011). What Factors Drive Global Stock Returns? *The Review of Financial Studies.* 24(8):2527-2574.

Houge, T. & Loughran, T. (2000). Cash Flow is King: Cognitive Errors by Investors. *Journal of Psychology and Financial Markets.* 1:161-175.

Hsiao, C. (2007). Panel Data Analysis - Advantages and Challenges. *Test.* 16(1)-22.

Hunjra, A. I., Ijaz, M. S, Chani, M. I., Hassan, S. and Mustafa, U. (2014). Impact of Dividend Policy, Earning per Share, Return on Equity, Profit after Tax on Stock Prices. *International Journal of Economics and Empirical Research*. 2(3):109-115.

Jooste, L. (2007). An Evaluation of the Usefulness of Cash Flow Ratios to Predict Financial Distress. *Professional Accountant.* 7(1):1-13.

Johannesburg Stock Exchange. (2018). *JSE Limited Listing Requirements*. Available: https://www.jse.co.za/content/JSERulesPoliciesandRegulationItems/JSE%20Listings%20Requirements.pdf

Keim, D. B. & Stambaugh, R. F. (1986). Predicting Returns in the Stock and Bond Markets. *Journal of Financial Economics*. 17(2):357-390.

Keller, G. & Warrack, B. (2003). *Statistics for Management and Economics*. 6th ed. Pacific Grove.

Kim, M. & Kross, W. (2005). The Ability of Earnings to Predict Future Operating Cash Flows Has Been Increasing Not Decreasing. *Journal of Accounting Research*. 43(5):753-780.

Kinnunen, J. & Niskanen, J. (1993). The Information Content of Cash Flows and the Random Walk: Evidence from the Helsinki Stock Exchange. *Accounting and Business Research.* 23(91):263-272.

Kothari, S., Shanken, J. & Sloan, R. (1995). Another Look at the Cross-Section of Expected Returns. *Journal of Finance.* 50:185-224.

Kruger, R. (2011). Evidence of Return Predictability on the Johannesburg Stock Exchange. Ph.D. Thesis. University of Cape Town.

Kruger, R., Toerien, F. & MacDonald, I. (2014). The Consistency of Style Anomalies over Periods of Market Crisis. *The African Finance Journal*. 16(1).

Lakonishok, J., Shleifer, A. & Vishny, R. W. (1994). Contrarian Investment, Extrapolation, and Risk. *Journal of Finance.* 49(5):1541-1578.

Lee, C., Lee, J. D., & Lee, C. C. (2010). Stock Prices and the Efficient Market Hypothesis: Evidence from a Panel Stationary Test with Structural Breaks. *Japan and the World Economy*. 22(1): 49-58.

Lev, B., Thiagarajan, S. R. & Thiagarajant, S. R. (1993). Fundamental Information Analysis. *Journal of Accounting Research.* 31(2):190–215.

Lin, F.J. (2008). Solving Multicollinearity in the Process of Fitting Regression Model using the Nested Estimate Procedure. *Quality and Quantity*. 42(3):417-426.

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*. The MIT Press. 47(1):13-37.

Litzenberger, R. H. & Ramaswamy, K. (1979*).* The Effect of Personal Taxes and Dividends on Capital Asset Prices: Theory and Empirical Evidence. *Journal of Financial Economics*. 7(2):163-195.

Livnat, J. & López-Espinosa, G. (2008). Quarterly Accruals or Cash Flows in Portfolio Construction? *Financial Analysts Journal*. 64(3): 67-79.

Livnat, J. & Zarowin, P. (1990). The Incremental Information Content of Cash Flow Components. *Journal of Accounting and Economics.* 13(1):25-46.

Lou, M. (2008). Unusual Operating Cash Flows and Stock Returns. *Journal of Accounting and Public Policy*. 27(5):420–429.

Maas, C.M. & Hox, J.J. (2004). Robustness Issues in Multilevel Regression Analysis Statistica. *Neerlandica*. 58(2):127–137.

MacKinlay, C. (1995). Multifactor Models Do Not Explain Deviations from the CAPM. *Journal of Financial Economics.* 38(1):3-28.

Malkiel, B. G. (2003). The Efficient Market Hypothesis and Its Critics. *Journal of Economic Perspectives*. American Economic Association. 17(1):59-82.

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*. 7(1):77–91.

Minton, B. A., Schrand, C. M., & Walther, B. R. (2002). The Role of Volatility in Forecasting. *Review of Accounting Studies*. 7(2-3):195-215.

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*. 34(4):768–783.

Muller, C., & Ward, M. (2013). Style-Based Effects on the Johannesburg Stock Exchange: A Graphical Time-Series Approach. *Investment Analyst Journal*. 42(77):1-16.

Narayan, P. K. & Westerlund, J. (2014). Does Cash Flow Predict Returns? *International Review of Financial Analysis*. 35:230–236.

Novy-Marx, R. (2013). The Other Side of Value: The Gross Profitability Premium. *Journal of Financial Economics*. 108(1):1-28.

Nyamgero, F. (2015). Are South African Firms Holding More Cash Than Before? Master's Thesis. University of Cape Town.

Ohlson, J. (1991). The Theory of Value and Earnings, and an Introduction to the Ball-Brown Analysis. *Contemporary Accounting Research.* 8:1-19.

Ou, J. A. & Penman, S. H. (1989). Financial Statement Analysis and the Prediction of Stock Returns. *Journal of Accounting and Economics.* 11(4):295-329.

Page, D. & Auret, C. (2014). A Time-Series Approach to Testing Cash-Flow Beta on the JSE. *Investment Analysts Journal.* 80:59-69.

Page, D., Britten, J. & Auret, C. (2013). Momentum and Liquidity on the Johannesburg Stock Exchange. *International Journal of Economics and Finance Studies.* 5(1):56-73.

Page, M. J. & Palmer, F. (1993). The Relationship Between Excess Returns, Firm Size and Earnings on the Johannesburg Stock Exchange. *South African Journal of Business Management.* 22:63-73.

Page, M. J. (1996). Further Evidence of Firm Size and Earnings Anomalies on the Johannesburg Stock Exchange. *De Ratione.* 10(1):22-44.

Page, M. J. (1986). Empirical Testing of the Arbitrage Pricing Theory Using Data from the Johannesburg Share Exchange. *South African Journal of Business Management*. 17(1):38-42.

Park, H. M. (2011). *Practical Guides to Panel Data Modeling: A Step by Step Analysis Using Stata.* International University of Japan. Available: www.iuj.ac.jp/faculty/kucc625/documents/panel_iuj.pdf

Pincus, M., Rajgopal, S. & Venkatachalam, M. (2007). The Accrual Anomaly: International Evidence. *The Accounting Review*. 82(1):169–203.

Quirin, J., O'Bryan, D. & Wilcox, W. (1999). The Corroborative Relation Between Earnings and Cash Flows. *Quarterly Journal Business and Economics.* 38(2):3-16.

Rayburn, J. (1986). The Association of Operating Cash Flow and Accruals with Security Returns. *Journal of Accounting Research.* 24:112-133.

Razali, N. M. & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests. *Journal of Statistical Modelling and Analysis*. 2:21-33.

Reinganum, M. R. (1982) A Direct Test of Roll's Conjecture on the Firm Size Effect. *Journal of Finance*. 37:27-35.

Reinganum, M.R. (1981). Misspecification of Capital Asset Pricing, Empirical Anomalies Based on Earnings' Yields and Market Values. *Journal of Financial Economics*. 9:19−46.

Roll, R. & Ross, S. A. (1980). An Empirical Investigation of the Arbitrage Pricing Theory. *Journal of Finance*. 35:1073-1103.

Rosenberg, B., Reid, K. & Lanstein, R. (1985). Persuasive Evidence of Market Inefficiency. *Journal of Portfolio Management*. 11:9-17.

Ross, S. (1976). The Arbitrage Theory of Capital Asset Pricing. *Journal of Economic Theory*. 13(2):341-360.

Samuelson, P. A. (1969). Lifetime Portfolio Selection by Dynamic Stochastic Programming. *The Review of Economics and Statistics*. The MIT Press. 51(3):239–246.

Scholes, M. & Williams, J. (1977). Estimating Betas from Nonsynchronous Data. *Journal of Financial Economics*. 5:309-327.

Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk. *The Journal of Finance*. 19(3):425–442.

Sloan, R. G. (1996). Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings? *The Accounting Review*. 71(3):289–315.

Stober, T. L. (1992). Summary Financial Statement Measures and Analysts' Forecasts of Earnings. *Journal of Accounting and Economics.* 15(2):347-372.

Strugnell, D., Gilbert, E. & Kruger, R. (2011). Beta, Size and Value Effects on the JSE. *Investment Analysts Journal.* 40(74):1-17.

Timmerman, A. (2008). Elusive Return Predictability. *International Journal of Forecasting.* 24(1):1-18.

Tobin, J. (1958). Liquidity Preference as Behavior Towards Risk. *The Review of Economic Studies*. 25(2):65-86.

Torres-Reyna, O. (2007). *Panel Data Analysis Fixed and Random Effects Using Stata* (V.4.2). Princeton University. Available:
https://www.princeton.edu/~otorres/Panel101.pdf

Tukey, J. W. (1977). *Exploratory Data Analytics.* Reading, Massachusetts: Addison-Wesley.

Van Niekerk, M. C. (1992). Die Gebruikswaarde van Kontantvloei-inligting: N' Empiriese Studie. Master’s Thesis, University of Stellenbosch.

Van Rensburg, B. (1999). Macroeconomic Identification of Candidate APT Factors on the Johannesburg Stock Exchange. *Journal of Studies in Economics and Econometrics.* 23(2):27-53.

Van Rensburg, P. & Robertson, M. (2003a). Style Characteristics and the Cross-Section of JSE Returns. *Investments Analysts Journal*. 57:1-10.

Van Rensburg, P. & Robertson, M. (2003b). Size, Price-to-Earnings and Beta on the Johannesburg Securities Exchange. *Investments Analysts Journal.* 58:7-16.

Van Rensburg, P. & Slaney, K. B. E. (1997). Market Segmentation on the Johannesburg Share Exchange. *Journal for Studies in Economics and Econometrics*. 23(3):1-23.

Van Rensburg, P. (2001). A Decomposition of Style-Based Risk on the JSE. *Investment Analysts Journal*. 54:45-60.

Van Rensburg, P. (2002). Market Segmentation on the Johannesburg Share Exchange II. *Journal for Studies in Economics and Econometrics.* 26(1):83-100.

Vuolteenaho, T. (2002). What Drives Firm-Level Stock Returns? *Journal of Finance.* 57(1):233-264.

Wapenaar, J. (1996). The Predictive Value of Cash Flow Information. Hons. Research Report. University of the Witwatersrand.

Ward, M. & Muller, C. (2012). Empirical Testing of the CAPM on the JSE. *Investments Analysts Journal.* 41(76):1-12.

Weisberg, S. (2005). *Applied Linear Regression*. 3rd ed. John Wiley & Sons.

Wessels, J. A., Smith, J. & Gevers, W.R. (1993). The Association Between Cash Flow Variables and Market Risk on the Johannesburg Stock Exchange: An Empirical Analysis. *South African Journal of Business Management.* 24:101-109.

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica.* 48(4):817-838.

Wilson, G. P. (1987). The Incremental Information Content of the Accrual and Funds Components of Earnings After Controlling for Earnings. *The Accounting Review*. 62: 293-322.

Wilson, G.P. (1986). The Relative Information Content of Accruals and Cash Flows: Combined Evidence at the Earnings Announcement Date and Annual Report Release Date. *Journal of Accounting Research.* 24:165-200.

Winjum, J.O. (1972). The Role of Accounting in the Economic Development of England. *University of Illinois Center for International Education and Research in Accounting*. 1500-1750.

Wooldridge, J. M. (2013). *Introductory Econometrics: A Modern Approach*. 5th ed. Mason, USA: South Western Cengage Learning.

Appendix

Appendix A: Definitions of Datastream Datatypes

This study uses fundamental and market data from *Datastream*, either directly or as an input into a calculation. The definition of each variable, as provided by the *Datastream* service is given below.

Mnemonic	Name	Definition
MNEM	Mnemonic	"This is a unique identification code, assigned by Datastream. It can be used to access data for a particular issue on all Research programs (that is, it is interchangeable with the Datastream code number). It consists of up to 6 characters, for example, RLRC for Rolls-Royce."
NOSH	Number of shares in issue	"This is the total number of ordinary shares that represent the capital of the company. The datatype is expressed in thousands. For shares with more than one class of equity issue, (NOSH) is held separately for each issue. The amount is updated whenever new tranches of stock are issued or after capital changes."
P X(P) ~ ZAR	Price (Adjusted – Default)	"Datatype (P) represents the official closing price. This is the default datatype for all equities and ETF's. The 'current' price on Datastream's equity programs is the latest price available to us from the appropriate market in primary units of currency (except in the case of the UK where price is given in pence). It is the previous day's closing price from the default exchange except where more recent or real-time prices are available, as listed in the Global Data Coverage section of this help system. The 'current' prices taken at the close of market are stored each day. These stored prices are adjusted for subsequent capital actions, and this adjusted figure then becomes the default price offered on all Research programs. The actual historical prices can be

		accessed using the unadjusted price datatype (UP). Prices are generally based on 'last trade' or an official price fixing. For stocks which are listed on more than one exchange within a country, default prices are taken from the primary exchange of that country (note that this is not necessarily the 'home' exchange of the stock). For Japan and Germany, prices from the secondary markets can be obtained by qualifying the price datatype with an exchange code."
RI X(RI) ~ ZAR	Total Returns Index	"A return index (RI) is available for individual equities and unit trusts. This shows a theoretical growth in value of a share holding over a specified period, assuming that dividends are re-invested to purchase additional units of an equity or unit trust at the closing price applicable on the ex-dividend date. Method: the discrete quantity of dividend paid is added to the price on the ex-date of the payment. Then: $RI_t = RI_{t-1} * \frac{P_t}{P_{t-1}}$ except when t = ex-date of the dividend payment D_t then: $RI_t = RI_{t-1} * \frac{P_t + D_t}{P_{t-1}}$ Where: P_t = price on ex-date P_{t-1} = price on previous day D_t = dividend payment associated with ex-date t Gross dividends are used where available and the calculation ignores tax and re-investment charges. Adjusted closing prices are used throughout to determine price index and hence return index."

MV X(MV) ~ ZAR	Market Value (Capitalisation)	“Market value on Datastream is the share price multiplied by the number of ordinary shares in issue. The amount in issue is updated whenever new tranches of stock are issued or after a capital change. • For companies with more than one class of equity capital, the market value is expressed according to the individual issue. • Market value is displayed in millions of units of local currency.”
DWTA X(DWTA) ~ ZAR	Total Assets	“Total Assets represent the sum of total current assets, long term receivables, investment in unconsolidated subsidiaries, other investments, net property plant and equipment and other assets.”
QTLE X(QTLE) ~ ZAR	Total Equity	“Total Equity consists of the equity value of preferred shareholders, general and limited partners, and common shareholders, but does not include minority shareholders' interest.”
OTLO X(OTLO) ~ ZAR	Net Cash Flow - Operating Activities	“Net Cash Flow – Operating Activities represent the net cash receipts and disbursements resulting from the operations of the company. It is the sum of Funds from Operations, Funds From/Used for Other Operating Activities and Extraordinary Items.”
FTLF X(FTLF) ~ ZAR	Net Cash Flow – Financing Activities	“Net Cash Flow – Financing Activities represents the net cash receipts and disbursements resulting from reduction and/or increase in long or short term debt, proceeds from sale of stock, stock repurchased/ redeemed/ retired, dividends paid and other financing activities.”
ITLI X(ITLI) ~ ZAR	Net Cash Flow – Investing Activities	“Net Cash Flow – Investing Activities represents the net cash receipts and disbursements resulting from capital expenditures, decrease/increase from investments, disposal of fixed assets, increase in other assets and other investing activities. A positive value in this field represents an outflow (use) of funds. A negative value in

		this field represents an inflow (source) of funds.”
SNCC X(SNCC) ~ ZAR	Net Change in Cash	“Net Change in Cash represents the sum of: Cash From Operating Activities [OTLO] Cash From Investing Activities [ITLI] Cash From Financing Activities [FTLF] Foreign Exchange Effects [SFEE]”
FCF X(FCF) ~ ZAR	Free Cash Flow	“Free Cash Flow represents Cash From Operating Activities for the time period minus Capital Expenditures for the same period.”
SOPI X(SOPI) ~ ZAR	Operating Income	“Operating Income [SOPI] represents total revenues from all of a company’s operating activities, after deducting any sales adjustments and excise taxes, reduced by total expenses that are operating in nature, such as variable costs directly related to the volume of sales, indirect operating costs, depreciation or amortization, operating provisions and other expenses incurred from operating activities. Operating income is commonly referred to as earnings before interest and taxes (EBIT).”
NIBX X(NIBX) ~ ZAR	Net Income Before Extraordinary Items/Preferred Dividends	“Net Income Before Extraordinary Items /Preferred Dividends represents income before extraordinary items and preferred and common dividends, but after operating and non-operating income and expense, reserves, income taxes, minority interest and equity in earnings.”
NINC X(NINC) ~ ZAR	Net Income	“Net Income represents net income after taxes, adjusted by minority interest, equity in affiliates, the GAAP adjustment and extraordinary items, before preferred distributions and other adjustments to net income.
INDM X(INDM) ~ ZAR	Industrial Grouping	“This datatype returns the Datastream level 6 industrial classification name, for example, ‘Breweries’.”

Appendix B: Equity Sample and Industry Groupings

There was a total of 38 industry groupings of JSE constituents used in the study, pictured below.

Company	Industry
Mr Price Group	Apparel Retailers
The Foschini Group	Apparel Retailers
Truworths International	Apparel Retailers
Metair Investments	Auto Parts
Massmart	Broadline Retailers
Pepkor Holdings	Broadline Retailers
Woolworths Holdings	Broadline Retailers
PPC	Building Mat.& Fix.
Compagnie Financiere Richemont	Clothing & Accessory
Exxaro Resources	Coal
Allied Electronics Corporation	Computer Services
Datatec	Computer Services
EOH Holdings	Computer Services
Nampak	Containers & Package
Distell Group Holdings	Distillers & Vintners
Barloworld	Divers. Industrials
Bidvest Group	Divers. Industrials
KAP Industrial	Divers. Industrials
Murray & Roberts Holdings	Divers. Industrials
Remgro	Divers. Industrials
Clicks Group	Drug Retailers
Dis-Chem Pharmacies	Drug Retailers

Reunert	Electrical Equipment
Astral Foods	Farm Fish Plantation
Oceana Group	Farm Fish Plantation
RCL Foods	Farm Fish Plantation
Telkom SA	Fixed Line Telecom.
AVI	Food Products
Pioneer Food Group	Food Products
Tiger Brands	Food Products
Tongaat Hulett	Food Products
Bid Corporation	Food Retail,Wholesale
Pick n Pay Stores	Food Retail,Wholesale
Shoprite	Food Retail,Wholesale
Spar Group	Food Retail,Wholesale
Steinhoff International Holdings	Furnishings
Sun International	Gambling
Tsogo Sun	Gambling
Anglo American	General Mining
African Rainbow Minerals	General Mining
Assore	General Mining
Bhp Billiton	General Mining
Glencore	General Mining
South32	General Mining
AngloGold Ashanti	Gold Mining
Gold Fields	Gold Mining
Harmony Gold Mining	Gold Mining
Sibanye Gold	Gold Mining
	Healthcare Providers

Life Healthcare Group Holdings	
Mediclinic Intl.	Healthcare Providers
Netcare	Healthcare Providers
Aveng	Heavy Construction
Raubex Group	Heavy Construction
Wilson Bayly Holmes – Ovcon	Heavy Construction
Cashbuild	Home Improvement Ret.
Italtile	Home Improvement Ret.
Lewis Group	Home Improvement Ret.
City lodge Hotels	Hotels
Invicta Holdings	Industrial Machinery
Hudaco Hndustries	Industrial Suppliers
Montauk Holdings	Integrated Oil & Gas
Kumba Iron Ore	Iron & Steel
Blue Label Telecoms	Mobile Telecom.
MTN Group	Mobile Telecom.
Vodacom Group	Mobile Telecom.
Naspers	Media
Mondi	Paper
Sappi	Paper
Adcock Ingram Holdings	Pharmaceuticals
Aspen Pharmacare	Pharmaceuticals
Anglo American Platinum	Plat.& Precious Metal
Impala Platinum	Plat.& Precious Metal
Lonmin PLC	Plat.& Precious Metal
Northam Platinum	Plat.& Precious Metal

Royal Bafokeng Platinum	Plat.& Precious Metal
Famous Brands	Restaurants & Bars
Spur Corporation	Restaurants & Bars
Advtech	Spec.Consumer Service
Curro Holdings	Spec.Consumer Service
AECI	Specialty Chemicals
African Oxygen	Specialty Chemicals
Omnia Holdings	Specialty Chemicals
Sasol	Specialty Chemicals
British American Tobacco	Tobacco
Grindrod	Transport Services
Imperial	Transport Services
Super Group	Transport Services
Trencor	Transport Services

Appendix C: Year-End Groupings

For purposes of comparability, all shares in the sample were grouped by their respective year-end dates. Six-month ahead returns, deflators and lagged values were then calculated for each group. Any changes in company year-end dates were considered, and adjusted calculations were performed accordingly.

Year-End	Tickers
February	AEL DCP PSG DTC FBR PIK RBX
March	NPN BAT CFR IVT LEW MEI MNK MRP OMN PGR PPC TFG TKG TON TSH RCL* VOD
May	BLU
June	HAR AIP WBO S32 CLH ITE SPG KAP IPL BIL ASR BVT

	SUR APN ARI AVI BID DGH IMP NHM RCL MUR CSB AEG WHL TRU SOL SHP GFI* SUI* SNH*
July	EOH
August	CLS
September	SAP BAW LON NPK NTC OCE PFG PPH RLO SNH SPP TBS ARL CML LHC AIP* PPC*
November	HDC
December	ANG GFI TRE SUI SGL

	RBP MTN MTA MSM MND KIO JSE GND GLN EXX COH BTI BRT AMS AGL AFX AFE ADH TSH*

Companies marked with a * had a change in year-end during the sample period. Adjustments were made accordingly.

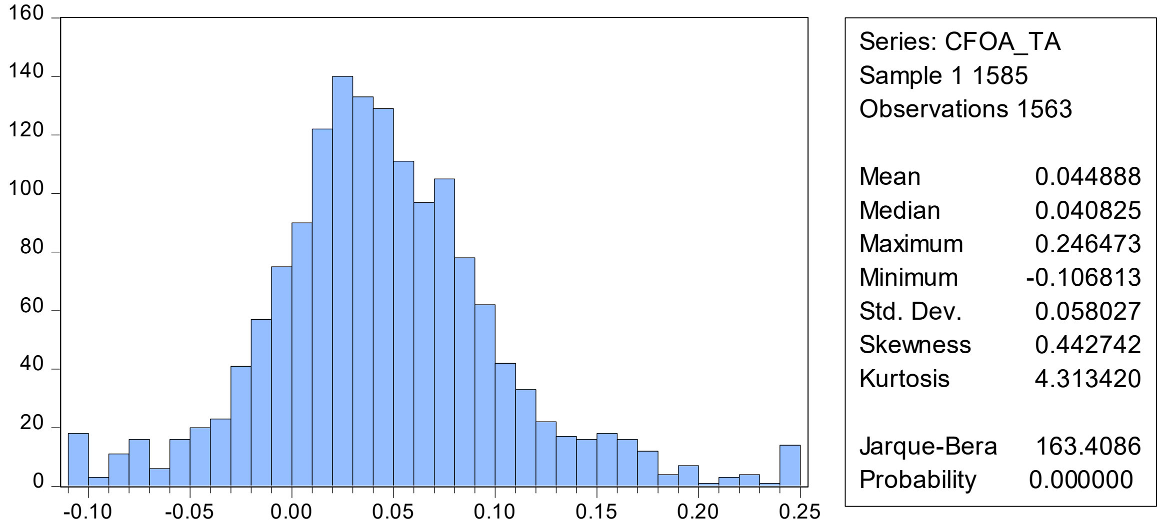
Appendix D: Correlations before Winsorisation

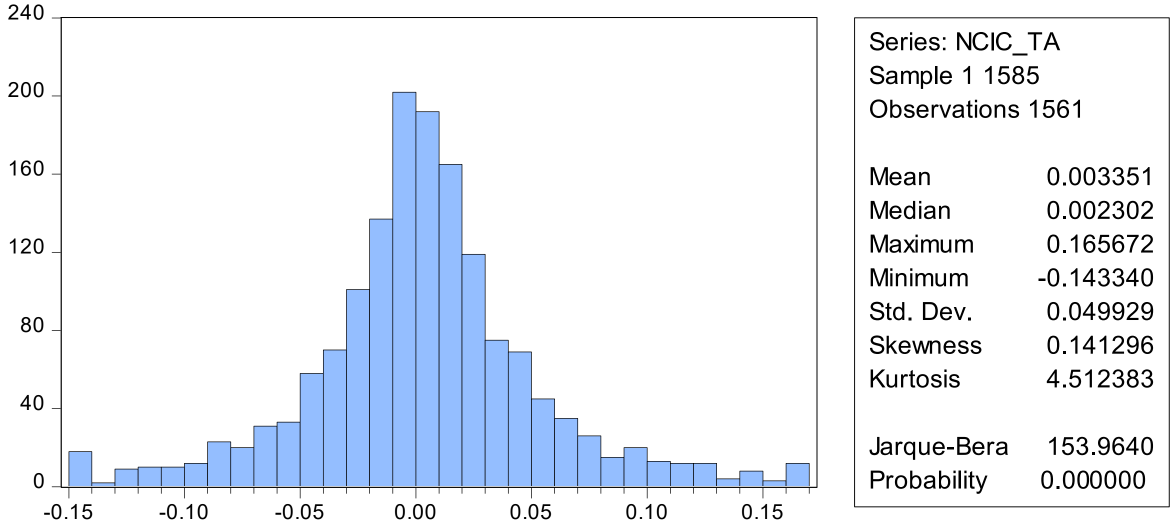
The initial data was deflated by Total Assets (TA) and Market Value of Equity (MVE), and a correlation matrix was performed before and after winsorisation (at 1% and 99%). Correlations between cash flow and accrual variables before any winsorisation procedure are pictured below.

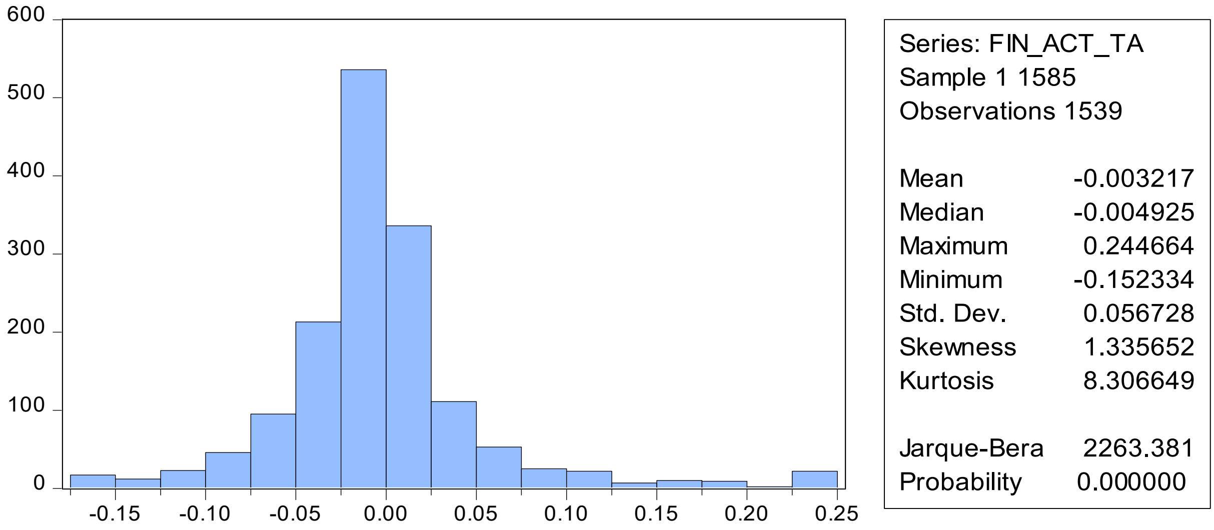
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	CFOA/TA	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	FCF/TA	0.90	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	NCIC/TA	0.51	0.55	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-
4	Fin Act/TA	-0.38	-0.37	0.21	1.00	-	-	-	-	-	-	-	-	-	-	-	-
5	Inv Act/TA	-0.18	-0.05	0.16	-0.55	1.00	-	-	-	-	-	-	-	-	-	-	-
6	OP/TA	0.40	0.32	0.03	-0.25	-0.12	1.00	-	-	-	-	-	-	-	-	-	-
7	NIBX/TA	0.29	0.25	0.04	-0.23	-0.02	0.89	1.00	-	-	-	-	-	-	-	-	-
8	NI/TA	0.22	0.18	0.02	-0.18	-0.02	0.71	0.79	1.00	-	-	-	-	-	-	-	-
9	CFOA/MVE	0.49	0.45	0.30	-0.14	-0.08	-0.06	-0.09	-0.09	1.00	-	-	-	-	-	-	-
10	FCF/MVE	0.46	0.53	0.30	-0.15	-0.06	0.06	0.05	0.06	0.80	1.00	-	-	-	-	-	-
11	NCIC/MVE	0.29	0.32	0.58	0.11	0.11	-0.03	-0.02	-0.01	0.43	0.49	1.00	-	-	-	-	-
12	Fin Act/MVE	-0.15	-0.14	0.10	0.51	-0.35	-0.07	-0.08	-0.07	-0.19	-0.14	0.25	1.00	-	-	-	-
13	Inv Act/MVE	-0.05	-0.01	0.07	-0.31	0.49	0.10	0.15	0.13	-0.32	-0.15	0.12	-0.66	1.00	-	-	-
14	OP/MVE	0.02	0.04	0.00	-0.03	0.03	0.55	0.55	0.43	0.12	0.14	0.08	-0.07	0.03	1.00	-	-
15	NIBX/MVE	0.02	0.04	0.00	-0.06	0.05	0.49	0.65	0.52	0.09	0.22	0.09	-0.10	0.09	0.83	1.00	-
16	NI/MVE	0.03	0.05	0.01	-0.04	0.02	0.33	0.43	0.44	0.09	0.39	0.17	-0.09	0.14	0.51	0.74	1.00

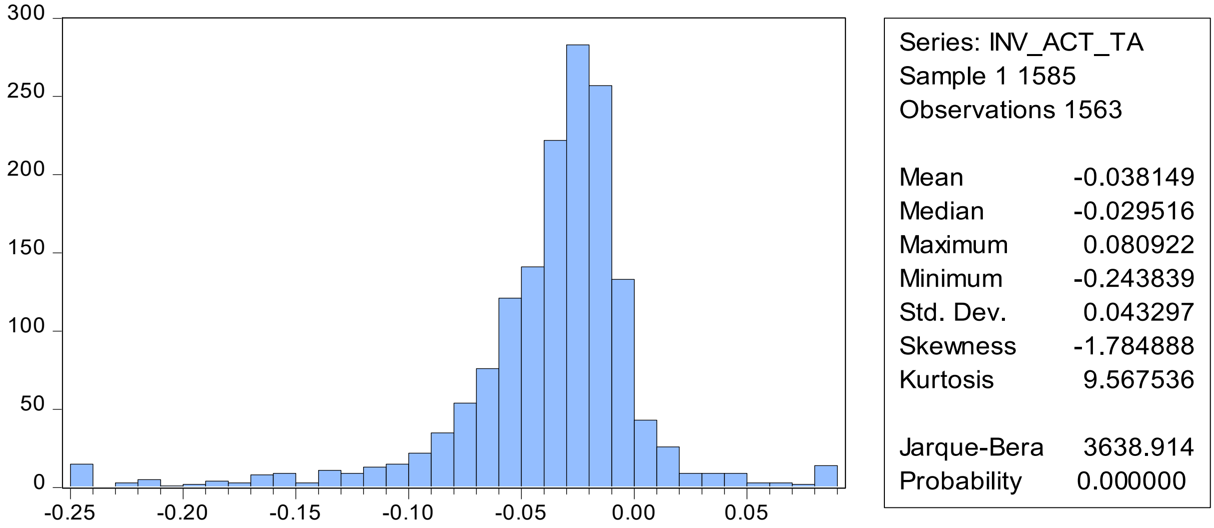
Appendix E: Histograms

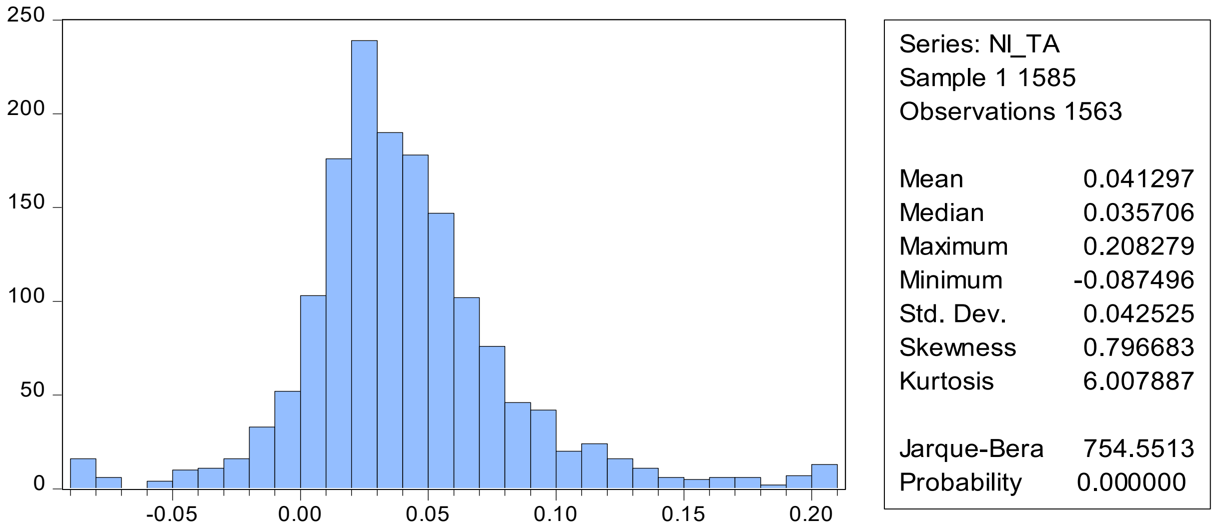
Pictured below are the histograms of the initial dataset (of the final variables used in panel regressions) before any adjustments or winsorisation procedures, obtained using *EViews*. The data required additional investigation in terms of descriptive statistics and visual representations of distributions.

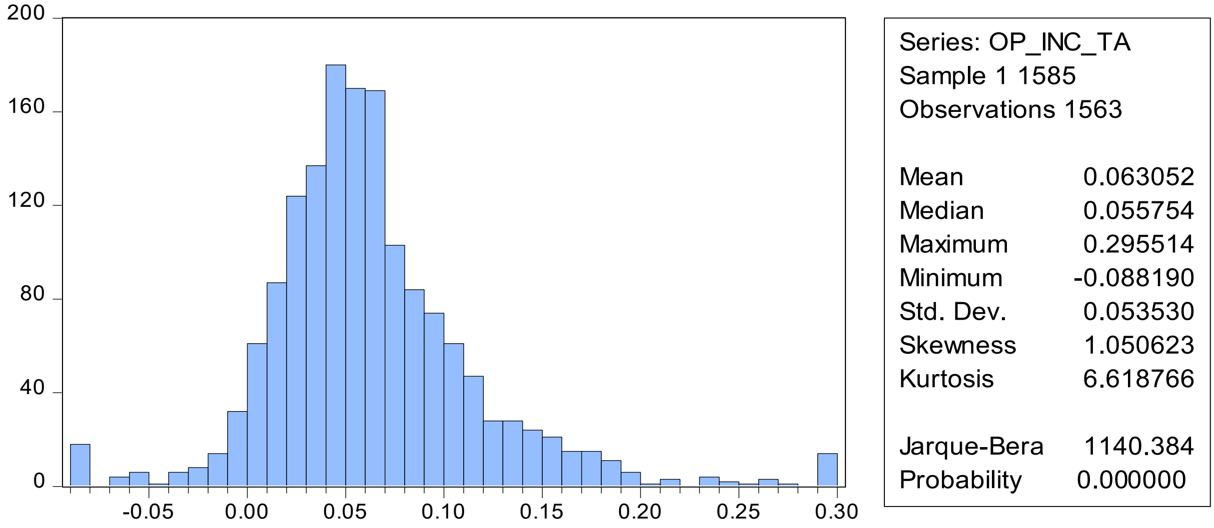


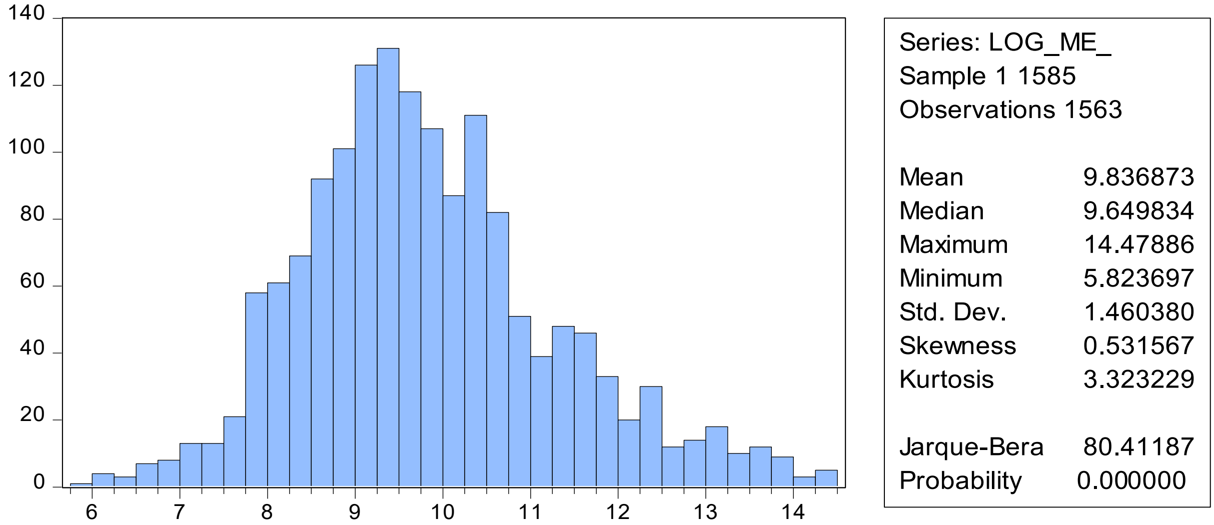


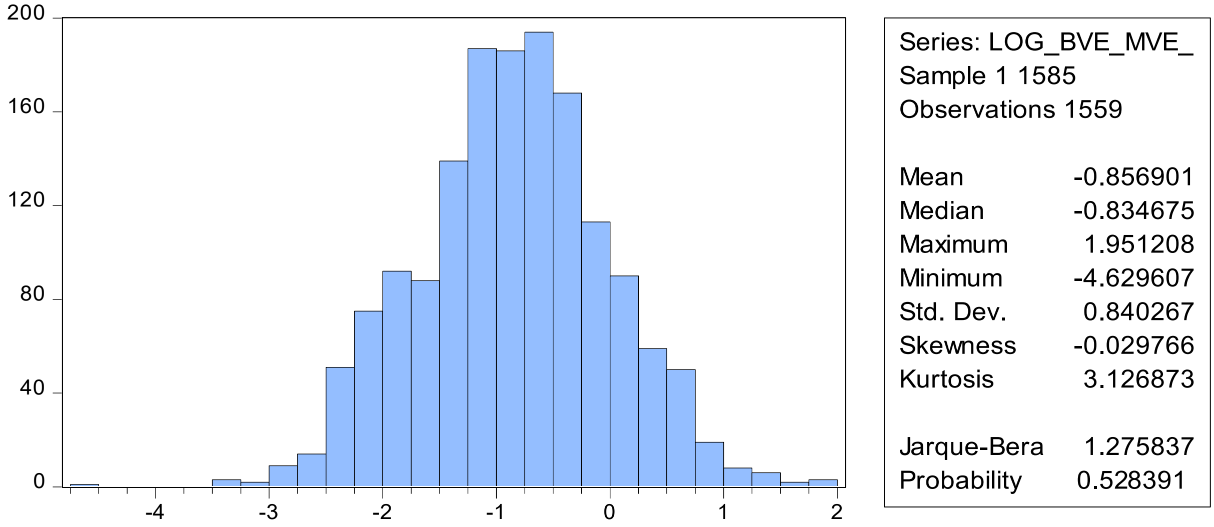

Series: FIN_ACT_TA
Sample 1 1585
Observations 1539
Mean -0.003217
Median -0.004925
Maximum 0.244664
Minimum -0.152334
Std. Dev. 0.056728
Skewness 1.335652
Kurtosis 8.306649
Jarque-Bera 2263.381
Probability 0.000000

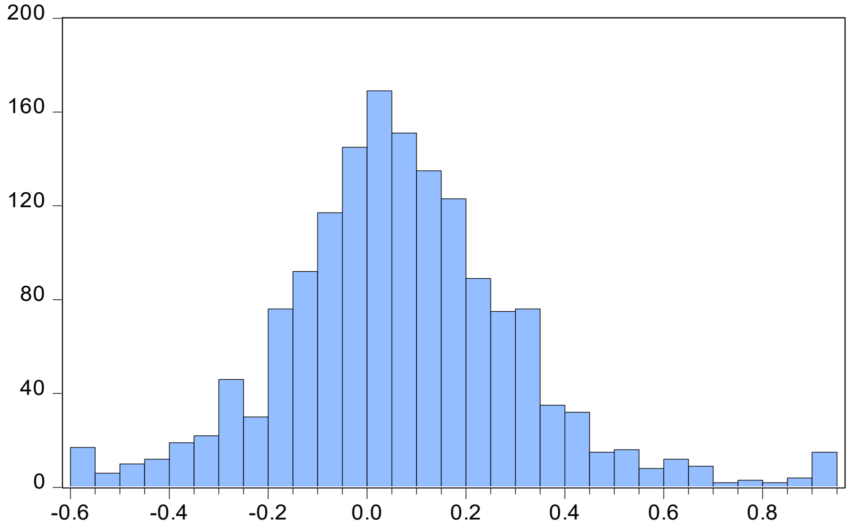

Series: INV_ACT_TA
Sample 1 1585
Observations 1563
Mean -0.038149
Median -0.029516
Maximum 0.080922
Minimum -0.243839
Std. Dev. 0.043297
Skewness -1.784888
Kurtosis 9.567536
Jarque-Bera 3638.914
Probability 0.000000


Series: NI_TA
Sample 1 1585
Observations 1563
Mean 0.041297
Median 0.035706
Maximum 0.208279
Minimum -0.087496
Std. Dev. 0.042525
Skewness 0.796683
Kurtosis 6.007887
Jarque-Bera 754.5513
Probability 0.000000

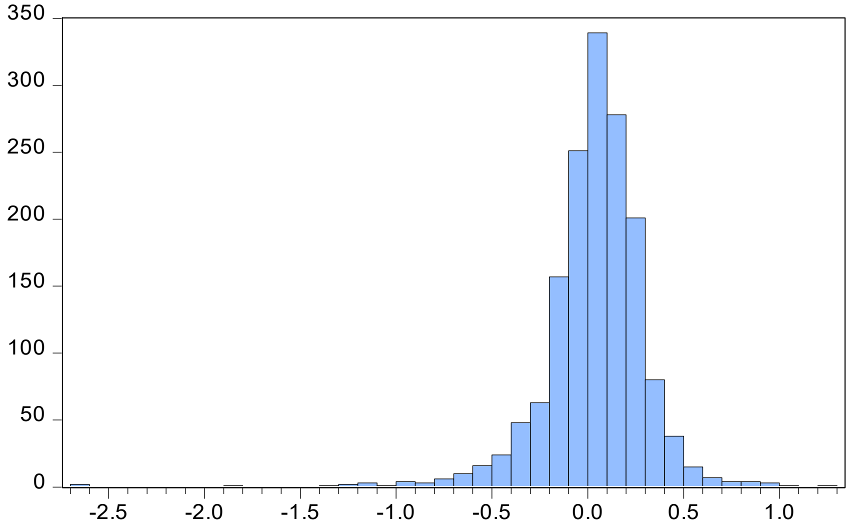

Series: OP_INC_TA
Sample 1 1585
Observations 1563
Mean 0.063052
Median 0.055754
Maximum 0.295514
Minimum -0.088190
Std. Dev. 0.053530
Skewness 1.050623
Kurtosis 6.618766
Jarque-Bera 1140.384
Probability 0.000000


Series: LOG_ME_
Sample 1 1585
Observations 1563
Mean 9.836873
Median 9.649834
Maximum 14.47886
Minimum 5.823697
Std. Dev. 1.460380
Skewness 0.531567
Kurtosis 3.323229
Jarque-Bera 80.41187
Probability 0.000000


Series: LOG_BVE_MVE_
Sample 1 1585
Observations 1559
Mean -0.856901
Median -0.834675
Maximum 1.951208
Minimum -4.629607
Std. Dev. 0.840267
Skewness -0.029766
Kurtosis 3.126873
Jarque-Bera 1.275837
Probability 0.528391



Series: RETURN	
Sample 1 1585	
Observations 1563	
Mean	0.069305
Median	0.055127
Maximum	0.904107
Minimum	-0.588210
Std. Dev.	0.244456
Skewness	0.383501
Kurtosis	4.326443
Jarque-Bera	152.8968
Probability	0.000000



Series: LOG_RETURNS_	
Sample 1 1563	
Observations 1563	
Mean	0.037008
Median	0.053661
Maximum	1.212971
Minimum	-2.685655
Std. Dev.	0.273343
Skewness	-1.985174
Kurtosis	18.81499
Jarque-Bera	17315.28
Probability	0.000000

Appendix F: Panel Data Selection Tests

The decision to use a fixed effects model in the study was informed by statistical tests carried out, comparing the various panel data approaches against each other. These tests were performed using *EViews*, shown below.

Panel Selection	Fixed vs Pooled	Random vs Pooled	Fixed vs Random
Regression Model	F-stat	Breusch-Pagan LM	X^2
1	2.09***	0.56	133.93***
2	2.06***	0.20	137.91***
3	2.34***	0.77	139.27***
4	2.48***	0.07	128.75***
5	2.36***	1.12	138.57***
6	2.50***	0.28	128.40***
7	2.07***	0.02	133.25***
8	2.49***	0.01	125.99***
9	2.17***	0.70	141.19***
10	2.13***	0.12	140.54***
11	2.41***	1.06	147.31***
12	2.55***	0.29	136.53***
13	2.57***	0.30	136.72***
14	2.61***	0.04	135.31***

* Statistical significance <0.10, ** Statistical significance <0.05, *** Statistical significance <0.01

Appendix G: Regression Model Diagnostic Tests

Model diagnostics were performed on the data, given the underlying assumptions of OLS regressions - homoscedasticity, autocorrelation, normality and multicollinearity. The tests were performed using *Eviews*, with outputs and significance levels pictured below.

Table G1: Homoscedasticity and Autocorrelation Test Results

OLS Assumption	Homoscedasticity	Autocorrelation
Regression Model	BPLM	Durbin-Watson
1	0.56	1.9989
2	0.20	1.9965
3	0.77	1.9981
4	0.07	1.9974
5	1.12	2.0016
6	0.28	2.0017
7	0.02	2.0004
8	0.01	2.0049
9	0.70	2.0173
10	0.12	2.0158
11	1.06	2.0154
12	0.29	2.0114
13	0.30	2.0199
14	0.04	2.0206

* Statistical significance <0.10, ** Statistical significance <0.05, *** Statistical significance <0.01

Table G2: Test for Normality

	CFOA	NCIC	OP	NI	Fin Act	Inv Act	Log(BVE /MVE)	Log(MVE)	$R_{6,6}$
Shapiro -Wilk	0.9746 ***	0.9660 ***	0.9221 ***	0.9306 ***	0.8635 ***	0.8338 ***	0.9977 **	0.9795 ***	0.9796 ***

* Statistical significance <0.10, ** Statistical significance <0.05, *** Statistical significance <0.01

Table G3: Multicollinearity Test Results

Regression Model 1

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.001757	46.69265	NA
CFOA	0.011830	1.680853	1.055915
LOG_BVE_MVE_	6.02E-05	2.299661	1.125747
LOG_ME_	1.86E-05	48.89374	1.050659
R6_6	0.000644	1.123485	1.041613

Regression Model 2

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.001767	46.66248	NA
NCIC	0.015321	1.005236	1.001122
LOG_BVE_MVE_	5.82E-05	2.210280	1.082651
LOG_ME_	1.86E-05	48.57094	1.045064
R6_6	0.000648	1.123816	1.041692

Regression Model 3

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.001852	49.69936	NA
OP_INC	0.018879	3.472116	1.452404
LOG_BVE_MVE_	8.08E-05	3.119232	1.526949
LOG_ME_	1.87E-05	49.68264	1.067612
R6_6	0.000642	1.130491	1.048108

Regression Model 4

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.001734	47.19136	NA
NI	0.028629	2.749403	1.406563
LOG_BVE_MVE_	7.58E-05	2.964763	1.451332
LOG_ME_	1.81E-05	48.66722	1.045792
R6_6	0.000638	1.139700	1.056646

Regression Model 5			
Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.001862	50.07899	NA
CFOA	0.013508	1.942593	1.220340
OP_INC	0.021769	4.012788	1.678570
LOG_BVE_MVE_	8.07E-05	3.122270	1.528436
LOG_ME_	1.90E-05	50.61213	1.087585
R6_6	0.000641	1.131632	1.049166

Regression Model 6			
Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.001730	47.20829	NA
CFOA	0.012408	1.809665	1.136834
NI	0.030747	2.960104	1.514355
LOG_BVE_MVE_	7.57E-05	2.969587	1.453693
LOG_ME_	1.82E-05	49.02256	1.053428
R6_6	0.000637	1.140994	1.057846

Regression Model 7			
Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.001760	46.66464	NA
CFOA	0.017132	2.431757	1.527463
NCIC	0.022067	1.454145	1.448193
LOG_BVE_MVE_	6.18E-05	2.356482	1.154264
LOG_ME_	1.87E-05	49.04347	1.055232
R6_6	0.000645	1.123845	1.041719

Regression Model 8			
Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.001729	47.15214	NA
NCIC	0.014878	1.007985	1.003859
NI	0.028618	2.756418	1.410438
LOG_BVE_MVE_	7.57E-05	2.969437	1.454505
LOG_ME_	1.80E-05	48.61177	1.045943
R6_6	0.000637	1.140198	1.056877

Regression Model 9			
Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.001776	47.04498	NA
CFOA	0.017048	2.418135	1.519215
FIN_ACT	0.022682	1.882872	1.875254
INV_ACT	0.034194	2.952365	1.656234
LOG_BVE_MVE_	6.30E-05	2.418732	1.177825
LOG_ME_	1.87E-05	49.00560	1.056665
R6_6	0.000652	1.127673	1.045064

Regression Model 10			
Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.001780	46.98419	NA
NCIC	0.017507	1.138195	1.133529
FIN_ACT	0.017845	1.477579	1.471461
INV_ACT	0.030765	2.648084	1.486592
LOG_BVE_MVE_	6.31E-05	2.416165	1.177297
LOG_ME_	1.87E-05	48.93837	1.056594
R6_6	0.000654	1.127958	1.045095

Regression Model 11			
Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.001896	50.71602	NA
OP_INC	0.022145	4.043443	1.689184
FIN_ACT	0.018261	1.530848	1.524655
INV_ACT	0.029542	2.575841	1.445009
LOG_BVE_MVE_	8.35E-05	3.237269	1.576420
LOG_ME_	1.92E-05	50.81488	1.095677
R6_6	0.000650	1.134015	1.050942

Regression Model 12			
Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.001759	47.70126	NA
NI	0.031602	3.015322	1.544103
FIN_ACT	0.017102	1.453981	1.448099
INV_ACT	0.027845	2.462244	1.381283
LOG_BVE_MVE_	7.94E-05	3.120241	1.519432
LOG_ME_	1.83E-05	49.12786	1.059301
R6_6	0.000645	1.142474	1.058781

Regression Model 13			
Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.001912	51.96088	NA
CFOA	0.153375	22.35294	14.04184
NCIC	0.154733	10.36013	10.31766
OP_INC	0.057837	10.74013	4.489113
NI	0.081142	7.766062	3.977720
INV_ACT	0.166245	14.73653	8.272851
FIN_ACT	0.144655	12.33512	12.28405
LOG_BVE_MVE_	8.35E-05	3.292685	1.604389
LOG_ME_	1.94E-05	52.21427	1.127322
R6_6	0.000644	1.144048	1.060003

Regression Model 14			
Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.001749	47.70268	NA
CFOA	0.016990	2.481501	1.559025
NI	0.032256	3.094337	1.584566
FIN_ACT	0.022634	1.934739	1.926912
INV_ACT	0.033247	2.955833	1.658179
LOG_BVE_MVE_	7.90E-05	3.121049	1.519825
LOG_ME_	1.82E-05	49.18851	1.060609
R6_6	0.000642	1.143029	1.059295