Predicting Equity Movements using Structural Models of Debt Pricing and Statistical learning

By

Kiran Ryan Singh

Report presented at the University of Stellenbosch in partial fulfilment of the requirements for the degree of

Masters of Commerce (Financial Risk Management)

Department of Statistics and Actuarial Science
University of Stellenbosch
Private Bag X1, 7602 Matieland, South Africa

Supervisor: Carel van der Merwe

Degree of confidentiality: A

April 2019

PLAGIARISM DECLARATION

1. Plagiarism is the use of ideas, material and other intellectual property of another's work and to present it as my own.
2. I agree that plagiarism is a punishable offence because it constitutes theft.
3. I also understand that direct translations are plagiarism.
4. Accordingly all quotations and contributions from any source whatsoever (including the internet) have been cited fully. I understand that the reproduction of text without quotation marks (even when the source is cited) is plagiarism.
5. I declare that the work contained in this assignment, except otherwise stated, is my original work and that I have not previously (in its entirety or in part) submitted it for grading in this module/assignment or another module/assignment.

Student number	**Signature**
18870295	
Initials and surname	**Date**
K.R. Singh	8 February 2019

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

ABSTRACT

Valuation is not an interesting problem in corporate finance, it is the only problem. Price and value are assumed to be the same number in economic theories of equilibrium and perfect capital markets. The economic theories of equilibrium asset pricing offer very weak practical suggestions for stock price behaviour at the firm level. The fundamental approach to stock price investing operates on the basis that price and value are two separate quantities and the stock price is fully determined by its intrinsic value. In this research the option-theoretic approach to default modelling is amended to provide an alternate view of value.

Structural models apply an option-theoretic approach inspired by Merton (1974) that uses equity market and financial statement data in order to determine default probabilities. Default probabilities obtainable from the reduced form class of models provides the basis for extending the Merton model to estimate the firms value from market observable credit spreads. The probability of default is then a known constant provided from the reduced form model. The Merton model is reformulated with equity or firm value being used as the subject of the formula. The re-appropriated Merton model then provides a unique estimate of the firm's value based on current market information. The expected return on equity is then estimated from market credit spreads using individual capital structure and traded equity information.

In this research it was found that historic estimates of return are poor predictors of future return at the firm level. The structural models provide good forecasts of return in some instances although have many challenges in implementation. The use of statistical learning methods was found to greatly improve predictions of future equity return movements using both debt and equity predictor variables, including unique predictor variables constructed from the structural models of the firm.

OPSOMMING

Waardering is nie 'n interessante probleem in korporatiewe finansies nie, dit is die enigeste problem. Prys en waarde word gesien as dieselfde getal in ekonomiese teorieë rakende ewewig en perfekte kapitale markte. The ekonomiese teorieë rakende die ewewig van bate pryse, verskaf baie swak praktiese voorstelle vir die gedrag van aandeel pryse op besigheidsvlak. Die fundametele uitkyk rondom beleggings in die aandele mark is gebou op die fundament dat prys en waarde twee verskillende bedrae is en dat die aandeel prys ten volle bepaal word deur sy intrensieke waarde. In hierdie navorsing word die opsie-teoretiese benadering tot wanbetaling modellering aangepas om 'n alternatiewe benadering vir waarde te kry.

Gestruktureerde modelle gebruik 'n opsie-teoretiese metode geïnspireer deur Merton (1974) wat gebruik maak van data wat bestaan uit ekwiteit en finansiële state om wanbetaling waarskynlikhede te bereken. Wanbetaling waarskynlikhede verkry van die verminderde klas van modelle, bied 'n basis om die Merton model uit te brei om 'n firma se waarde te voorspel vanaf markverwante krediet premies. Die waarskynlikheid van wanbetaling is dan 'n konstante wat gekry word vanaf die verminderde model. Die Merton model word dan verander sodat die ekwiteit of firma se waarde gebruik word as die inset van die formule. Hierdie model gee dan 'n unieke voorspelling van die firma se waarde gebasseer op huidige mark inligting. Die verwagte opbrengs op ekwiteit word dan bepaal deur die mark se krediet premies, gebasseer op individuele kapitaal strukture en ekwiteit informasie.

In hierdie navorsing was dit gevind dat historiese skattings van opbrengs swak voorspellings van die toekomstige opbrengs op 'n firma vlak is. Die gestruktureerde modelle bied goeie vooruiskattings van opbrengs in sekere gevalle, maar het baie probleme met implimentering. Deur gebruik te maak van statistiese metodes is dit gevind dat vooruitskattings van toekoms opbrengs drasties verbeter wanneer beide skuld en ekwiteit, asook unieke veranderlikes wat opgestel word deur gebruik te maak van die gestruktureerde modelle van die firma, gebruik word.

ACKNOWLEDGEMENTS

The author would like to thank the following people/institutions for their contribution towards this project:

- Mr Carel van der Merwe in his efforts and contributions towards the development of this research. Moreover the guidance, support and opportunities he made available over the course of my academic career have been truly invaluable. Thank you.
- Prof. Willie Conradie for making dreams possible, accepting me into the department as well as making funding opportunities available to me and so many other students – the opportunities you have given me are truly life changing.
- Schroders asset management for making the journey possible and providing the funding over the course of my masters.
- Sarah, the SU international office and the University of Amsterdam for making it possible to study abroad at the University of Amsterdam for 6 months.
- To my fellow masters class mates (the A-team) Jan, Justin, Monique and Nadia, each of whom have made the past two years truly special and absolutely unforgettable. Thank you for making the journey such a memorable one.
- Justine Heald for all her contributions towards the proofreading.
- To Jazz-hands, Douglas, Gusheshe, MJ and all those who endured the long nights.

CONTENTS

LIST OF ABBREVIATIONS

- AIC- Akaike Information Criterion
- APT- Arbitrage Pricing Theory
- BIC- Bayes Information Criterion
- BVC- Bidvest Co.
- CAPM- Capital Asset Pricing Model
- CB – Class Balance
- CDS – Credit Default Swap
- CF- Cash Flow
- Cp- Mallow's Cp
- CS- Credit Spreads
- DCF- Discounted Cash Flow
- EM- Equity Multiplier
- EMH- Efficient Market Hypothesis
- FNR – False Negative Rate
- FPR- False Positive Rate
- GARCH- Generalized Auto-Regressive Conditional Heteroscedastic
- GRF- Group Five Construction
- INL- Investec Limited
- JSE- Johannesburg Stock Exchange
- KNN – K Nearest Neighbours
- MC- Market Capitalized
- MM- Miller and Modigliani
- MSE- Mean Squared Error
- PD- Probability of Default
- PV- Present Value
- ROA- Return on Assets
- ROE- Return on Equity
- TN – True Negative
- TP- True Positive
- TSS – Test Sample Size
- WACC- Weighted Average Cost of Capital

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

"As recently as a generation ago, finance theory was still little more than a collection of anecdotes, rules of thumb, and manipulations of accounting data. The most sophisticated tool of analysis was discounted value and the central intellectual controversy centred on whether to use present value or internal rate of return to rank corporate investments. The subsequent evolution from this conceptual potpourri to a rigorous economic theory subjected to scientific empirical examination was, of course, the work of many, but most observers would agree that Arrow, Debreu, Lintner, Markowitz, Miller, Modigliani, Samuelson, Sharpe, and Tobin were the early pioneers in this transformation" (Robert Merton, 1990).

1.1 PRELUDE

For the modern connoisseurs of uncertainty and quantitative methods, it may be challenging to understand what finance was like before modern portfolio theory. Risk and return are such fundamental concepts of finance courses that it is hard to imagine a time where these concepts were once a theoretical novelty (Varian, 1993). A brief chronological review of the development of models around asset pricing in capital markets reveals some of the great insights provided by brilliant theories and theorists in the last century.

Modigliani and Miller (1958) suggest that under perfect capital market conditions the valuation of a company should be independent of capital structure, such as debt to equity ratios. Sharpe, Lintner, and Treynor (1964), in their capital asset pricing model (CAPM), propose that expected return is singularly related to non-diversifiable risk associated with the market portfolio. Ross and Roll (1976) introduce arbitrage pricing theory, opining that arbitrage trading by smart money would eliminate price deviation from fundamentals caused by irrational investors. Kahneman and Tversky (1979) develop behavioural finance establishing a behaviour basis for market inefficiencies. More recently Fama and French (1993) observe the importance of more than one priced risk factor.

Existing theories about the behaviour of capital markets share a common denominator, they all provide a theoretical construct of how the world should work. The purpose of formulating a theoretical construct is not complete realism, rather a framework from which meaningful inference or prediction can be made (James, Witten, Hastie and Tibshirani, 2015). Many of the original engineers of these brilliant theories willingly acknowledge that the theories are based on an array of implausible assumptions. Iconoclastically, the oversimplification achieved in many of these theories are touted as their brilliance as opposed to their demise.

Arnott (2004) theorizes that the sheer brilliance of these theories blinds us to their limitations, adding that to accept theories as facts we often accommodate the assumptions as facts. The limitations of existing theories in finance has not gone unnoticed, with some expressing their dissatisfaction more assertively than others. George Soros, one of the most powerful and profitable modern investors, goes as far as saying *"existing theories about the behaviour of stock prices are remarkably inadequate. They are of so little value to the practitioner that I am not even fully familiar with them. The fact that I could get by without them speaks for itself"* (Soros, 1987).

The *Alchemy of Finance* written by George Soros in 1987 has been described as somewhat of a revolutionary book. Mr Soros puts forth his theory of reflexivity in the stock market as well as highlights the severe limitations of preceding theory. Mr Soros's theory of reflexivity in the book has been described as the first modern non-technical effort to describe the dynamics of the path between points of extreme valuation and equilibrium in the market place.

1.2 REFLEXIVITY IN THE STOCK MARKET

There is a beautiful synchronicity present in the taxonomy of investment styles, theories of stock price behaviour, and beliefs regarding the degree of efficiency in the market. Theories of stock price behaviour are characterised by the three broad classes of investment management styles, namely: passive, technical and fundamental. Technical analysis operates under the weak form of market efficiency and suggests that the past experience is relevant in predicting the future. The random walk hypothesis operates under the assumption of an efficient market, and that prices quickly incorporate all information leaving no economic profit opportunities in the market (Elton *et al.*, 2011). The random walk or efficient market hypothesis (EMH) is often the reason spouted for investing in many passive funds.

The fundamentalist approach along with the EMH bares the bulk of criticism from Soros (1987). The performance of well renowned successful investors such as Warren Buffet and Soros is often cited as sufficient anecdotal evidence to refute the random walk hypothesis. The fundamentalist view of stock price behaviour is an out of equilibrium model, where the price and intrinsic value of a stock are two distinctly separate quantities. The price of a stock is assumed to revert towards the intrinsic value in line with equilibrium fair market price of the firm. The classic or fundamentalist asset pricing theories stress that asset prices are determined by the intrinsic value only. In other words asset prices are completely determined by expected future cash flows and the risk premium for bearing the risk (Mpofu *et al.*, 2013).

The key insight or scathing criticism of the fundamentalist approach by Soros lies around the assumption that the market cannot influence the price. The classic economic theory of pricing under the perfect competition paradigm shares the analogue of a unilateral relationship between price and value. The shared axiomatic beliefs in the market pricing mechanism is no accident here since the fundamentalist view of stock price behaviour is derived from economic theories of pricing in perfect competition. Soros (1987) so aptly points out that the omission of the reflexive relationship between price and value is much more glaring in stock markets than in others.

Stock market valuations of the firm's equity have a direct way of influencing the underlying values. The issue and repurchase of shares by a company or corporate transactions such as mergers and acquisitions directly translate to influences on the underlying value. There are other more subtle ways in which share prices may influence the underlying value, such as credit rating, consumer acceptance and management credibility to name a few. Granting the manner in which equity prices impact these factors is subtle. There is nothing subtle about the magnitude to which these factors impact equity prices. The influence of these factors on stock prices is of course well recognized, it is the influence of stock prices on these factors that is so strangely ignored by the fundamentalist approach (Soros, 1987).

1.3 CAPITAL STRUCTURE AND RISK & RETURN

The influence of the capital structure of the firm on underlying value or stock price returns have been assumed away in the macroeconomic theorists endeavour to provide a generalized theory of market pricing mechanisms in the utopic setting. The influence and importance of the firm's capital structure on risk and return has not been forgotten in other fields of academic financial theory and practice.

The DuPont model expression for the measure of a firms return on equity (ROE) suggest that a firms ROE depends on operating efficiency; asset use efficiency and financial leverage. The firms ROE can then be expressed as the return on assets (ROA) times the equity multiplier (Mpofu *et al.*, 2013). To simplify further the DuPont model provides that $ROE = ROA \times EM$ in notation terms. The equity multiplier is simply the portion of the firm's assets financed by equity, capturing the use of financial leverage in the ROE. The DuPont model decomposition illustrates the reflexive nature between capital structure, ROE and the market price of a firm's equity.

The relationship between leverage and equity risk is well documented within market risk literature and behavioural finance. The asymmetric generalized auto regressive conditional heteroscedastic (A-GARCH) and GJR GARCH are models for conditional volatility designed to incorporate the 'leverage effect' within equity returns (Sui *et al.*, 2011). The leverage effect describes the asymmetric response of investors to increases and decreases in equity prices. For a decrease in equity stock prices the volatility of the equity price escalates more than for an equivalent increase in equity prices, since investors are loss averse *ceteris paribus*. The increased financial leverage resulting from decreased equity prices, in conjunction with more volatile equity prices ultimately increases the risk on equity (Alexander, 2008). The leverage effect further substantiate that capital structure plays a significant role in the risk and return on equity.

The banking or lending institutions devote a considerable amount of time and resources to the assessment and quantification of the firm's capital structure and risk and return. According to Zaik *et al.* (1996), Bankers Trust developed the risk-adjusted return on capital (RAROC) methodology in the late 1970s with the intent to measure the risk of a bank's credit portfolio and the amount of equity capital required to limit the bank to a specified probability of loss. The risk measure within the RAROC framework moves away from a market-driven definition of risk to a measure of risk that is firm specific. Crouhy *et al.* (1999) contends that the underlying premise of the risk-adjusted return on capital (RAROC) approach is that it is possible to construct a risk-adjusted rate of return measure such that it can be compared with a firm's cost of equity capital. The implicit assumption is that the RAROC measure adjusts the risk of a business relative to that of a firm's equity. The RAROC framework by Bankers Trust has long acknowledged the impact of capital structure on risk and return at the firm level.

1.4 RESEARCH PROPOSITION

The property of reflexivity in stock prices, DuPont's partition of the firms ROE, the '*leverage effect*' and RAROC framework all corroborate that at the security level, the capital structure (price of debt and equity) have large influences on firm specific risk and return. If both of these are traded in the market, can the values of debt and equity be used to predict stock price behaviour? The more pertinent question undoubtedly is how to make use of market variables of debt and equity to capture forward looking expectations around the value of individual firms? The pricing of credit derivatives for individual firms may yield some insight in this regard.

In a recent study by Bai and Wu (2016), the researchers observed that firm fundamentals are able to adequately explain cross-sectional variation in credit default swap (CDS) spreads. It is then a tenable assumption that discrepancy in CDS or credit spreads may adequately describe variations in firm fundamental values. Defining credit spreads from the premiums of single-name Credit Default Swaps (CDSs) instead of bond yields compared to some benchmark would give a more accurate measure of counterparty credit risk (CCR), but CDS data is complex and not readily available (Gregory, 2012). CDS spreads are arguably the purest market instrument from which to define the markets view of riskiness of a firm's debt, although sadly in the South African context these instruments are virtually non-existent.

Vassalou and Xing (2004) is the first study that uses Merton's (1974) option pricing model to compute default measures for individual firms and assess the effect of default risk on equity returns. The undertaking by Vassalou and Xing (2004) provides the inspiration for exploring the use of default risk in linking the firm's capital structure and equity returns. There are essentially three broad classifications for default modelling approaches as summarized by Trujillo and Martin (2005): the first is the historical approach where probabilities of default are estimated from statistical models applied to series of historic data and/or credit ratings. The second class consists of the reduced form models where probabilities of default are derived from a market observable credit spreads. Last being the structural model paradigm under which default is modelled using an option theoretic approach.

The last alternative is the basis for so-called structural models, which will constitute a major area of focus within this research. The theoretical inspiration for the series of structural models is that of Merton (1974). The basis of the structural approach is that the debt and equity of a firm can be regarded as contingent claims on the firm's assets. The value of the debt and equity of a firm thus depends on the value of its assets as well as the forward-looking expectation surrounding the value of those assets. While scrutinizing the assumption of the RAROC framework, Crouhy *et al.* (1999) demonstrates that the Merton (1974) contingent claims framework can be used to describe the relationship between capital structure, expected return and the probability of default.

The above points culminate in giving rise to the first proposition explored within this research-is it possible to use structural models of default to capture forward looking expectations of return for individual firms? More concretely, the proposition is to explore the use of structural models to link market observable credit spreads and forward looking expectations of equity returns. Tackling the broader topic of risk and return under the structural model approach is more appropriately left for the possibility of PhD research.

1.5 Research design / Chapter overview

The research paper has both quantitative and qualitative aspects. The qualitative aspect is the review of the various models and methods for predicting stock price returns. The research focus of the paper is more specifically on the use of structural models in the prediction of stock price returns. Chapter 1 served as an introduction discussing the background/rationale as well as the context and need for the research.

Chapter 2 of the research paper contains an outline of the literature that is relevant to the theories of stock price behaviour and theories of firm valuation. The literary review presented briefly covers the theoretical development of the various methods available for estimating the probability of default. Furthermore, the proposed use of the Merton (1974) model to link market default probabilities and unique firm valuations and expected return will be discussed comprehensively.

Chapter 3 describes the methodology followed within the research to test whether structural models of default can be used to provide estimates of the firm value and expected return on the stock price. The methodology narrates fully how predictors of firm returns are created under different theories as well as how the usefulness of these predictors is evaluated. The methodology elucidates the process under which the validation of stock return predictions is performed in an analogous manner to credit risk model validation.

Chapter 4 follows by reviewing the results obtained by following the methodology and theory set out in the previous chapters. The Merton model predictors of expected return are evaluated against those from the CAPM model for five firms traded on the Johannesburg stock exchange (JSE). The predictors of firm return are evaluated on how well they predict the class (positive or negative) of future excess returns. The forward excess returns are also defined for a variety of time horizons under which the return is earned and further evaluated in terms of class predictive capability through the passage of time. The comparison and return class predictions from different models and theories provide an indication of whether structural models of default yield decent predictors of stock price returns.

A summary of the overall results and outcomes of the research, along with the overall conclusions drawn from this research are presented lastly in Chapter 5. This also includes the scope and limitations of the investigation along with recommendations for further research.

2 ACADEMIC LITERATURE REVIEW

"If I have seen further it is by standing on the shoulders of Giants" Sir Isaac Newton (1676)

The academic review of the relevant literature concerning the theories of stock price behaviour and firm valuation is recapitulated at the necessary level of granularity. The review of modern portfolio theory reveals the model pedagogies for stock price behaviour in equilibrium models. Within the review of these models special effort is made to highlight the distinction between the *brilliance of theory* and the limitations encountered in practical implementation. Thereafter the fundamental approach to firm valuation and stock price behaviour is examined in detail, disbursing special consideration to highlight the distinction between the price and intrinsic value of the company's equity.

The literary review presented briefly covers the theoretical development of the various methods available for estimating the probability of default. A discussion of the theoretical and conceptual basis behind the structural models as well as the implementation of structural models is included. Moreover, the proposed use of the Merton (1974) model to link market default probabilities with unique firm valuations and expected return is discoursed in prodigious detail. Further theoretical substantiations for the contingent claims approach to firm valuation are weaved into to the arguments covering the mathematical specification of the suggested framework.

2.1 MODERN PORTFOLIO THEORY

The portfolio selection problem stated in classical economic terms is the problem of selecting the portfolio that maximizes the expected utility of an individual's end-of-period wealth (Ross, 2009). Since future asset returns are unknown it is the expected asset returns that should be used in the portfolio selection problem. However to maximize the expected return for a portfolio of stocks then, an investor should purchase the single stock with the highest expected return. Markowitz (1952) formulation of portfolio optimization leads quickly to the fundamental point that riskiness of a stock should not be measured by the variance of the stock in isolation, but also by covariance.

2.1.1 Markowitz portfolio selection

The Markowitz efficient frontier, developed in 1952, laid the foundations for modern portfolio theory for portfolio selection. The efficient portfolios (combination of securities) are defined as the set of portfolios with returns that are maximized for a given level of risk based on mean-variance construction (Elton, Gruber, Brown & Goetzmann, 2011). While Markowitz provided the insight for diversification in portfolio selection, it is the measures of risk and return that are the most contentious parts of the framework.

The exact definition of risk is quite a contentious issue as no single definition of risk will be sufficient under all scenarios. Broad definitions of risk allow for interpretations of risk as the uncertainty surrounding achievement of the expected outcome. The definition of risk according to the oxford dictionary defines risk as a situation involving exposure to danger or to expose someone or something to danger, harm, or loss. It is necessary to distinguish between financial and non-financial risks in the risk and rewards conundrum. This is since no rational person can expected to yield benefit from additional exposure to non-financial or pure risk.

The definition of risk in the investment setting should thus analogously imply the possibility of loss. Risk Metrics confirms the intuitive rationale, defining risk as the explicit possibility of loss. A more comprehensive definition of financial risk is provided by Mpofu, De Beer, Myhnardt and Nortje (2013), financial risk can be described as the probability of experiencing an event that has a negative financial implication, thus a loss. The semantics of risk provide that rewards for taking on risk in the investment context, should be seen as the reward gained for exposure to possible financial losses.

The problem with interchanging volatility and risk is that volatility is a measure of deviation from the desired outcome, in this case the expected portfolio or securities return. Volatility is an appropriate measure of risk where risk may be viewed as the uncertainty surrounding achieving the expected outcome. Moreover, volatility is only an adequate description of the possibility of loss in the case of normally distributed portfolio returns (Dowd, 2005).

Additionally, the measure of volatility is based on historical information, arguably providing a limited indication of future risk to the return achievable by security or portfolio. The standard measure of volatility does not distinguish between calculating historical volatility and estimating future volatility (Alexander, 2008). Historical mean-variance optimization similarly forecasts expected return as the historical return.

Herein lies the fundamental limitation of mean-variance portfolio construction in the portfolio selection problem. The mean-variance constructed efficient frontier does not distinguish between estimating past mean-variance structures and forecasting future mean covariance structures. This leads to the necessary distinction around the use of the mean-variance constructed efficient frontier in portfolio selection problems. Mean-variance efficient frontiers are more appropriately used for evaluating past portfolio performance as opposed to selecting a portfolio that will be most efficient at the end of period.

Mean-risk models are the ubiquitously used approach in portfolio selection in practice (AlHalaseh, Islam and Bakar, 2016). In the portfolio selection problem it is the **future** asset returns that are unknown, utility maximization thus requires estimates of expected asset returns for the optimal solution. The mean variance construction only provides a wholesome solution to portfolio selection problems where markets are perfectly efficient and future returns are not predictable.

Correctly estimating or forecasting asset returns and risks is self-evidently imperatively reliant on the specification for the statistical model which generates the portfolio or asset returns. Forecasting equity returns provides the basis for correctly solving the portfolio selection problem in markets of varying degrees of efficiency.

2.1.2 Capital asset pricing model

Asset pricing theories and models such as the Capital Asset Pricing Model (CAPM) are extensions of mean-variance portfolio optimization problems. These set of models make use of historical mean-covariance structures to estimate the expected return and risk of securities portfolios or individual assets. The CAPM extends the efficient portfolio idea by relating the expected or required return of an asset to its relative exposure to systemic risk. In this sense additional financial rewards are only received for taking on additional exposure to systemic market risk. The exposure to systemic market risk is captured through historical mean covariance structures in the following way.

$$R_i = R_f + \beta(R_m - R_f) \tag{2.1.1}$$

Hence, the required rate of return on an asset, R_i is estimated from its riskiness relative to the market, determined by historical covariance structures (Elton *et al.*, 2011). The CAPM is appealing since it captures both risk and return through a single parameter β. The key insight of the CAPM is that the equilibrium value of an asset depends on how it co-varies with other assets, not on its risk as a stand-alone investment (Varian, 1993).

For stock pricing the CAPM estimates the required rate of return of an asset as part of a well-diversified portfolio in a well-functioning securities market. The CAPM and the performance measures that stem therefrom are generally used to analyse past performance. Any insights investors hope to gleam into future performance is largely contingent on beta and the expected return on the market. Beta is often thought of in a forward-looking sense, yet it is based on historical price movements and predictability is limited. An important concept to remember is that beta quantifies the degree to which a portfolios returns are influenced by the same factors that influence the market return; the portfolio returns are not actually *caused* by the market (Kidd, 2011).

In an explanation of the 'High Risk, Low Return Puzzle' by Frazzini and Pedersen (2014), zero-beta long short portfolios are constructed to empirically evaluate the risk return relationship suggested by the CAPM. The beta-neutral portfolios consistently provided excess returns well above the zero which is what the standard CAPM suggests the strategy should earn. The behavioural economists' school of thought explains the high risk low return puzzle by introducing heterogeneous beliefs, short-sales and leverage constraints. The behavioural finance model then argues that high beta stocks are more sensitive to disagreement and are more likely to have binding short-sales constraints, ultimately yielding over inflated prices for high beta stocks.

The work of Ang *et al.* (2006) along with Baker, Bradley and Wurgler (2011) also provide that risk measured using return volatility yields the same outcome as evidenced in the infamous 'High Risk, Low Return Puzzle'. Empirically both volatility and beta are shown to be limited descriptions of the risk and return relationship in the investment context. The failure of the CAPM model to empirically generalize suggests that the model provides useful insight into the evaluation of past performance portfolio performance but rather fragile practical suggestions for explaining stock price behaviour at the security level.

2.1.3 Arbitrage pricing theory

Ross (1976) and Roll (1977) criticize the CAPM and suggest a multifactor approach called the Arbitrage Pricing Model. Fama and French (1993) suggested the three-factor model that considers beta, size and book-to-market as risk factors for describing the risk premium. Arbitrage Pricing Theory (APT) provides extensions of the CAPM by allowing for more than one factor to describe the expected return. APT models estimate the expected return of an asset or portfolio by multiple regression. Advantageously, this allows for the incorporation of exogenous factors that may describe asset returns. Most notably, traditional asset pricing theories and models have an additional weakness in the sense that they are only accurate descriptions of risk and return where ordinary regression assumptions hold.

It is widely observed that asset returns are not normally distributed, a key assumption of ordinary linear regression Alexander (2008). Additionally, volatility is not a comprehensive measure of dispersion for non-normally distributed random variables, let alone the best measure of risk. Curiously, Elton *et al.* (2011) contends that, empirically, APT and multi-factor models produce more reasonable explanations of variations in portfolio returns relative to the single index factor model of the CAPM. Once again there is a critical distinction necessary in the real world application of these models. The models as formulated offer an explanation of past asset returns and do not forecast the future asset return in its generic formulation except under perfect market conditions.

The incorporation of a distinction between historic estimates and forecasts for asset returns would require a recognition of non-stationarity in price series. Mandelbrot (1966) clarifies that so little is known of non-stationary time series that accepting non-stationarity amounts to giving up any hope of performing a worthwhile statistical analysis. The lack of inclusion of a distinction between past and future estimates in the equilibrium models are conceivably more for the ease of model formulation as opposed to accurately describing the complex nature of reality. The second proposition underscored in this research is that measures of historic mean-variance are not good predictors of future return at individual firm level.

2.2 Valuation models

''If we do not recognize the fundamental difference that exists between price and value, then we are doomed. Historically this distinction did not really matter in corporate finance because the two, price and value, were supposed to be the same, to remain equal forever. Who has been telling us that? These people do not live in New York; they live in Chicago. The Chicago School of Economics has been telling us for a century that price and value are identical, i.e. that they are the same number'' – Sylvain Raynes – The subprime Crisis & Ratings: PRMIA Meeting Notes 2007.

Sylvain Raynes highlights the essential difference between the price of equity and the value of equity, the distinction between the two so often blurred with economic theories about market efficiency. If price and value are the same number it implies there is no such thing as a good deal or a bad deal, there are only fair deals. The deliberation provided in this next segment paints a clearer picture of the fuzzy link between equilibrium models of asset pricing, concepts of price and value, and the translation into theories of stock price behaviour.

2.2.1 Miller Modigliani theory

The insight of Modigliani and Miller (1958), referred to hereafter as the MM theory, showed that under the assumption of frictionless markets and perfect completion, that debt and equity are perfect substitutes in the absence of taxes. Thus, under perfect competition conditions a firm cannot increase its value by changing its capital structure since this would create arbitrage opportunities for investors of the firm's debt and equity.

The MM theorem is a consequence of value additivity; a portfolio of assets must be worth the sum of the values of the assets that make it up. Initially the proposition of value additivity appears to be at odds with the insights about diversification. An asset should be worth more combined in a portfolio with other assets than it is standing alone due to diversification benefits. Most critically the point is that asset values in a well-function securities market already reflect the value achievable by portfolio optimization. The principal of value additivity is even more fundamental than the CAPM since it relies solely on no arbitrage considerations (Varian, 1993).

2.2.2 Discounted future cash flow

The discounted future cash flow (DCF) approach to value was introduced by John Burr Williams (1938) in his book *The Theory of Investment Value*. Williams argued that the value of a stock should be the present value of its dividends, a novel theory for its time. By extension, the intrinsic value of a firm is determined as the expected present value of future cash flows streams discounted at the weighted average cost of capital (WACC), represented in Equation 2.2 The WACC is comprised of the cost of debt and the cost of equity, where the cost of equity is usually determined by the CAPM. The mathematical expression for intrinsic value or price helps discern the critical insights from different theories of stock price behaviour and value.

$$PV = \sum_{i=1}^{n} \frac{CF_i}{(1+WACC)^i} \tag{2.2}$$

In the fundamentalist view, if markets are perfectly efficient then the intrinsic value will equal the price since all future cash flows and developments are correctly discounted by the market. The MM theory states that under no-arbitrage considerations the mix of debt to equity should not influence the discount rate in the valuation of the firm in Equation 2.2. The nomenclature of required rate of return or expected return in the CAPM framework is easily understood in lieu of the above two points. In equilibrium, the price is equal to value and the discount rate is independent of capital structure. Thus much like a bond is traded at par value when the yield to maturity (YTM) is equal to the coupon rate, the expected return on equity or required rate of return on equity is the discount rate that correctly equivocates price and value in equilibrium.

The theories of asset pricing models in equilibrium offer little with regards to explanations of stock price behaviour at the security level. In the fundamentalist approach to stock price behaviour as Soros calls it, the price and intrinsic value are distinctly separate quantities and the market is not always in equilibrium rather continuously moving towards equilibrium. In the fundamentalist approach to stock price behaviour, stock prices are assumed to be fully determined by the firm's intrinsic value. Moreover, the market will always tend towards equilibrium and stock prices should tend towards their intrinsic values. The rate at which stock prices are assumed to tend towards its intrinsic value, reveals the belief held regarding the degree to which markets are thought to efficient.

The use of WACC as the appropriate discount rate in the fundamentalist approach is a movement away from equilibrium conditions where valuation is independent of capital structure as found within the MM theorem. Interestingly then the disposition from equilibrium price and value allows the capital structure to enter into determinations of value through the WACC, however the disposition from equilibrium has not allowed for the influence of price on value. The lack of acknowledgement of a reflexive relationship between price and value seems a bit absurd, if a firm's capital structure contains part of its own traded equity then equity prices are a natural influence on WACC.

The maintenance of a one-way relationship between price and value - even where the critical conditions for this relationship are assumed not to be present - is an ideological inconsistency within the fundamentalist approach to stock market investing. The DCF approach is a central tool for the pricing of financial contracts and instruments throughout quantitative finance. However this fundamentalist approach of value, fitted to modeling stock price behaviour is still sorely dissatisfactory. The estimation of future cash flows to a firm are highly subjective along with the appropriate discount rate and cash flow timings. There is also very little empirical evidence to support the hypothesis that share price moves towards intrinsic value (Soros, 1987).

The DCF approach to asset pricing does not lend itself well to the portfolio selection problem. The major disadvantage of the DCF valuation models is that risk is not an explicit parameter of the model. Valuation models may provide good means to estimate the expected return on a security, however these models fail to provide any intuitive means to measure risk associated with the expected outcome (Mpofu et al., 2013).

2.2.3 The subjective theory of value

The notion of *subjective* value and the theory of greater fools provides an alternate view to value. The subject theory of value is the idea that the value of the firm is not inherent and instead worth the amount market participants are willing to pay. The subjective view of value is a movement away from the notion of an *intrinsic value*. The subjectivity contained within value offers a suggestion for the lack of empirical evidence to support the claim that asset prices move towards their intrinsic values over time.

The theory of greater fools simply states that there will always be a "greater fool" in the market who will be ready to pay a price based on a higher valuation for an already overvalued security. The greater fool theory approach to investing focuses on determining the likelihood that the investment can be resold for a higher price instead of trying to accurately discern the *intrinsic* value of the investment in the firm. The greater fool theory is not really designed to provide investors with a trading strategy based on finding tools. The greater fool theory is articulated in a manner that aids explanations surrounding the formations of speculative bubbles in markets.

2.3 Structural models of default probability

The next section introduces the structural models of default of the firm. The cross-over from modern portfolio theory and stock price behaviour to credit risk modeling may appear very tangential at first glance. The reader is encouraged to bear in mind that the default model pedagogies may reveal a different paradigm for estimating value from increased market observable information. The exact suggestions for the re-appropriation of the structural model framework is deferred to the penultimate section of the literature review.

2.3.1 Merton (1974) model

In 1974, Robert Merton introduced a new option-theoretic approach to credit risk modelling and measurement based on ideas and formulations that were implicit in the Black and Scholes (1973) option-theoretic framework. The firm value in the context of the Merton (1974) model is the economic value of the total assets of the firm. As with all structural models, the Merton model begins with a specification of a stochastic process for the firm value. The Merton model assumes that the firm value follows a geometric Brownian motion:

$$\mathrm{d}V_t = \mu V_t \mathrm{d}t + \sigma V_t \mathrm{d}W_t \tag{2.3.1}$$

Where, V_t is the firm value at time t, μ is the drift of the firm value, and σ is the volatility of the firm value. The second assumption of the Merton model is that the capital structure of the firm consists solely of equity and debt. Furthermore, the debt is assumed to be a single issue of zero-coupon form where the face value of the debt is denoted by F and the maturity date is T.

To complete the model, further assumptions regarding the conditions that trigger default and the costs incurred in the event of default are required. The Merton (1974) model assumes that default can only occur at time T when the debt becomes due and no covenants can trigger default before time T. Furthermore, debt holders are assumed to have absolute priority over equity holders in the event of default and there are no frictional market costs associated with liquidation in the event of default.

Under these assumptions Merton (1974) shows that holding the risky debt of the firm is equivalent to holding a portfolio consisting of a long position in default risk free bond paying F at time T and short a put on the firm's assets with strike F and maturity T. The following decomposition follows naturally:

$$D_t = B - P \tag{2.3.2}$$

Where D_t represents the value of risky debt, B is the value of riskless debt and P is the value of the put on the firm's assets. This decomposition importantly shows that the spread on the risky debt is completely determined by the value of the put, P (Smit, Swart and Van Niekerk, 2003). The value of the put can be determined using the Black-Scholes formula since all the conditions of Black-Scholes have been met in the assumptions. Merton (1974) expresses the value of the put slightly differently to the standard Black-Scholes formula:

$$B = Fe^{-r(T-t)} \qquad \text{Value of riskless debt.} \qquad (2.3.3)$$

$$P = e^{-r(T-t)}\, F.N\left(-d + \sigma\sqrt{T-t}\right) - V_t N(-d) \quad \text{Merton put formula.} \qquad (2.3.4)$$

Where:

$$d = \frac{1}{\sigma\sqrt{T-t}}\left[\ln\left({}^{1}/{}_{L}\right) + \frac{1}{2}\sigma^2\,(T-t)\right] \qquad (2.3.5)$$

$$L = \frac{e^{-r(T-t)}\,F}{V_t} \qquad \text{Leverage ratio of the firm} \qquad (2.3.6)$$

Simplified expressions for the risky value of debt then include:

$$D_t = \ e^{-r(T-t)}\, F.N\left(d - \sigma\sqrt{T-t}\right) + V_t N(-d) \qquad (2.3.7)$$

$$D_t = V_t - E_t \quad \text{Accounting value} \qquad (2.3.8)$$

The firm's equity price is represented by E_t and In addition $N(\,.\,)$ is the standard cumulative normal distribution function. The risk-neutral probability of default is easily extracted as the probability that $V_t < F$. From Black-Scholes formula, this is simply the probability that the put P finishes 'in the money'. Smit *et al.* (2003) show that the risk-neutral probability of default is given by:

$$N(\,-d + \ \sigma\sqrt{T-t}\,) \qquad (2.3.9)$$

The actual or real-world probability is given similarly as the probability that $V_t < F$. However, the process for the firm value has drift μ as opposed to drift in the risk-neutral world where the risk-free rate, r is the drift of the firm value process. The actual probability of default will typically be less than risk-neutral probabilities since $\mu > r$ usually. The higher risk-neutral default probability can be interpreted as comprising of actual default probability and a premium for uncertainty of timing and magnitude of the default (Sundaran and Das, 2010).

Another useful feature of the Merton (1974) model is that is allows for estimation of expected recovery rates in the risk-neutral setting. Under the Merton framework, Smit *et al.* (2003) provide the following closed form expression for the expected recovery rate:

$$\frac{1}{D}E_T[V_T|\,V_T < D] = \ e^{-r(T-t)}\left(\frac{V_T}{D}\right)\left(\frac{N(-d)}{N(\,-d+\sigma\sqrt{T-t}\,)}\right) \qquad (2.3.10)$$

This feature is extremely useful since both CDS spreads and bonds prices require an estimate of the recovery rate to estimate default probabilities. Although the model is theoretically very appealing, since it provides a simplistic model for the credit spread along with default probability and recovery rates, the model encounters a number of major challenges in practical implementation. The first of these challenges is that both the firm value V_t and its volatility σ are unobservable in the market. Wang and Suo (2006) argue that in the Merton model, the firm's equity is treated as a European call option on the firm's assets and hence the firm value and volatility should satisfy the following set of simultaneous equations:

$$E_T[V_T, \sigma] = V_T\, N(d) - e^{-r(T-t)}\, D\, N(d - \sigma\sqrt{T-t}\,) \tag{2.3.11}$$

$$\sigma_E = \sigma V_T \frac{N(d)}{E_T} \tag{2.3.12}$$

The relationship between the equity and asset volatility only holds instantaneously and the algorithm forces stochastic variables to be constant, where in practice the hedge ratio and leverage ratio are not stable enough to provide meaningful estimates (Holman *et.al,* 2011). Crosbie and Bohn (2003) illustrate that the procedure biases the probability of default in exactly the wrong direction as increased leveraging will drive down asset volatility and under predict default.

Vassalou and Xing (2004) describe a more complex iterative procedure to solve for the asset volatility. Alternatively, Duan (1994) describes an intricate maximum likelihood approach based on observed market equity or bond prices in order to solve the unknown parameters relating to the firms value and volatility. A distinct advantage of the maximum likelihood approach is that it directly provides an estimate of the real-world drift parameter $\{\mu\}$ of the unobserved asset value process under the physical probability measure (Wang and Suo, 2006).

The second major issue encountered with implementing the Merton (1974) model is that the capital structure assumption is too simplistic. In practice, capital structures consist of many issues of debt outstanding, with varied coupons, maturities and subordination structures (Sundaran and Das, 2010). In order to simplify reality, Delianedis and Geske (1998) suggest a zero-coupon bond that has an equivalent duration of the existing structure replacing the capital structure. An alternative in the popular Moody's KMV vendor model is to use the aggregate of short term and long-term liabilities to estimate the face value of the zero-coupon debt F.

2.3.2 KMV proprietary model

One on the most notable implementation of a structural credit risk measurement model is the Moody's KMV (MKMV), Trujillo & Martin (2005) summarizes the MKMV approach in four stages:

(i) Calculate a default boundary.

(ii) Estimate asset value and volatility.

(iii) Calculate the Distance to Default (DD).

(iv) Map DD into Expected Default Frequency (EDF).

2.3.2.1 Parameter estimation

In the first stage, the capital structure is collapsed into a single debt issue or default boundary calculated as the sum of the short-term liabilities and a fraction of the longer-term liabilities. In the second stage, the asset value and volatility are backed out from observed equity value, volatility and capital structures. The KMV estimation process that is implemented can be outlined as:

1. Let, $\{E_t\}_{t=0}^{n}$ be a time series of equally spaced observed equity prices.
   - $E_t = BS(V_t, \sigma, 'Call')$.
2. Express $V_t = BS^{-1}(E_t, \sigma,\ 'Call')$.
   - Begin with a guess for σ_v.
   - Obtain $\{V_t(\sigma_v)\}_{t=0}^{n}$, time series of asset values.
3. Determine Series of continuously compounded asset returns.

   $$r_{t+1} = \log\left(\frac{\widehat{V_{t+1}}}{\widehat{V}_t}\right)$$
4. σ_v^2: Sample variance of implied asset returns.
5. Update initial guess of σ_v and reiterate several times until convergence is achieved.
   - Asset value is then most recent asset value in $\{V_t(\sigma_v)\}_{t=0}^{n}$ in the final iteration.
   - Volatility of assets is then taken as σ_v for which convergence is achieved.

2.3.2.2 Distance-to-Default (DD)

In the third stage the MKMV approach moves away from the Merton approach and defines the 'distance to default' (DD) as the number of standard deviations the firm value has to move make before the firm is in default (Hayne, 2004). The MKMV approach defines the distance to default δ in a simplified manner as shown by Crosbie and Bohn (2003):

$$\delta = \frac{V_t - D}{\sigma V_t}$$

Sundaran and Das (2010) illustrate that normalizing the distance in this fashion allows for comparability between firms of how far the firm is from default even though the firms may differ substantially in other ways. The final stage uses the estimated 'DD' to determine an 'expected default frequency' (EDF) from a proprietary default database, which represents the likelihood of the given firm defaulting over the specified horizon (Hayne, 2004). The MKMV practioner model thus uses a blend of market and historical data in a structural framework to estimate the probability of default for a given firm.

Another priopertery model of great use is the Bloomberg issuer risk model. According to Bloomberg (2012) the issuer risk model provides an independent assesment of credit health, using market and fundamental data with innovative quantitative models. The bloomberg issuer risk model provides one and five year default probabilities along with implied CDS spreads. In this research paper, the Bloomberg issuer risk model is assumed to provide reasonable and consistent estimates for default probabilities and can thus be used as stable benchmark for market view of the firms default risk.

2.3.3 Delianedis & Geske (1998) model

The challenge of applying the contingent claims model of the firm in the case of increased capital structure complexity can be addressed in one of two ways. Either simplifying the firm's capital structure to fit within the existing model framework, or extending the theoretical framework. Sundaran and Das (2010) recommend that extending the theoretical structure of the model to incorporate more complex debt structures is the more academically appreciated approach to the undertaking.

Delianedis and Geske (1998) (DG) provide extensions of the Merton model, which allows for more complex capital structures. These models allow for multiple debt issues of varying coupons, maturities and seniority or subordination (Chen, 2013). In the simplest extension of the Merton Model, DG (1998) allow for two tranches of zero-coupon debt in the firm's debt structure with face values F_1 and F_2 and maturities T_1 and T_2 respectively where $T_1 < T_2$. Since there are now two dates at which equity holders may choose to default the DG model thus involves a compound option pricing approach (Sundaran and Das, 2010).

Delianedis and Geske (1998) further illustrate that at the first maturity date T_1, the firm is solvent if:

$$V_{T_1} > F_1 + B_{2,T_1} \tag{2.3.12}$$

Where, V_{T_1} is the value of the firm's assets at T_1, F_1 is the face value of the first tranche of debt at T_1 and B_{2,T_1} is the value of the second tranche at T_1. If the firm is solvent, the Delianedis and Geske (1998) model then assumes that the first tranche of debt will be refinanced with equity. The model may be implemented under the assumption that refinancing is not allowed however, this adversely affects the second tranche of debt and is less realistic (Sundaran and Das, 2010).

The condition for solvency provided by Delianedis and Geske (1998) defines a critical cut-off value V^*, for the value of the firm at T_1, which is equivalent to the strike price of the first option in a compound option. The critical cut-off value or strike price of the first option is given by:

$$V^* = F_1 + B_{2,T_1} \qquad B_{2,T_1}: \text{ Market Value of Risky Debt at } T_1$$

$$\Rightarrow V^* = F_1 + V^* - BS(V^*, \sigma_v, r, T_2, K_2,' Call')$$

$$\Rightarrow F_1 = BS(V^*, \sigma_v, r, T_2, K_2,' Call')$$

In the DG model it can be shown the holders of equity are now the holders of a compound call option on the assets of the firms as follows:

$$E_t = CO(V_t, \sigma_v . r . T_1, T_2, v^*, F_2,' Call','Call')$$

Where CO denotes the compound option. The analytical price, at time t, for the compound call on call option is given by Hull (2012) as,

$$E_t = V_t N_2(a_1, b_1, \rho) - F_2 e^{-r(T_2-t)}\, N_2(a_2, b_2, \rho) - F_1 e^{-r(T_1-t)}\, N(1-a_2)$$

Where,

$$a_1 = \frac{\log\left(\frac{V_t}{V^*}\right) + \left(r - q + \frac{1}{2}\sigma_v^2\right)(T_1 - t)}{\sigma_v\sqrt{T_1 - t}} \qquad a_2 = a_1 - \sigma_v\sqrt{T_1 - t}$$

$$b_1 = \frac{\log\left(\frac{V_t}{F_2}\right) + \left(r - q + \frac{1}{2}\sigma_v^2\right)(T_2 - t)}{\sigma_v\sqrt{T_2 - t}} \qquad b_2 = b_2 - \sigma_v\sqrt{T_2 - t}$$

The value of the combined tranches of risky debt in present value terms is,

$$D_t = V_t - E_t$$

$$D_t = V_t[1 - N_2(a_1, b_1, \rho)] + F_2 e^{-r(T_2-t)}\, N_2(a_2, b_2, \rho) + F_1 e^{-r(T_1-t)}\, N(1-a_2)$$

The set of unobservable parameters $\{V^*, V_t, \sigma_v\}$ are solved from the following set of simultaneous equations.

$$F_1 = BS(V^*, \sigma_v, r, T_2 - T_1, K_2,' Call') \qquad (2.3.13)$$

$$E_t = CO(V_t, \sigma_v . r . T_1, T_2, v^*, F_2,' Call','Call') \qquad (2.3.14)$$

$$\sigma_E = \frac{V_t}{E_t} \times \frac{\partial E_t}{\partial V_t} \times \sigma_v$$

$$\sigma_E E_t = e^{-qT_2} N_2(a_1, b_1, \rho)\, V_t \sigma_v \qquad (2.3.15)$$

From the compound option model, Delianedis and Geske (1998) provide three risk-neutral probabilities as follows:

$$\text{risk neutral short run PD} = 1 - N(\mathrm{a}_2)$$

$$\text{risk neutral long run PD} = 1 - \frac{N_2[a_2; b_2, \rho]}{N(a_2)}$$

$$\text{risk neutral total PD} = 1 - N_2[a_2; b_2; \rho]$$

The short run default probability represents the probability of default at T_1. The total default probability represents the probability of the firm defaulting at either T_1 or T_2. The long-term default probability is the probability of default at T_2 conditional on not having defaulted at T_1 and is thus also refered to as the forward default probability (Chen, 2013).

The DG model has the appealing feature of being able to simultaneously capture short-term and long-term default characteristics of the firm. Sundaran and Das (2010) argue that there are many firms with poor quality yet, conditional on survival of initial financial difficulty, have reasonable longer term financial prospects and that the forward default probability of the DG model is likely to reflect these key features.

Although the model appears to be a relatively simply extension of the Merton framework, considerable additional complexity arises in solving for the unobservable parameters of process for value of the firm. The procedure for estimating these unobservable parameters is subject to the same weaknesses as with the case of Merton. The equity and asset volatility relationship is still instantaneous as described by Crosbie and Bohn (2003), additionally there is the added complexity of a third unobservable variable V^*, the cut-off value, in the estimation procedure.

2.4 STRUCTURAL MODELS OF FIRM VALUE AND EXPECTED RETURN

"… Options are specialized and relatively unimportant financial securities …" – Robert Merton Nobel Prize winner for work on option pricing – in 1974 seminal paper on option pricing.

Vassalou and Xing (2004) is the first study that uses Merton's (1974) option pricing model to computed default measures for individual firms and assess the effect of default risk on equity returns. The formulation proposed here is significantly different in terms of exploring the possibility of applying Merton (1974) models to capture forward looking expectations of return for individual stocks from traded debt and equity.

2.4.1 Structural models of firm asset value

Using the contingent claims framework of Merton (1974), Crouhy *et al.* (1999) considers a business that is financed in such a way that the expected rate of return on equity equals some pre-specified value. The authors show that while it is possible to pick a capital structure so as to achieve a required rate of return on equity, the probability of default will change as the volatility of the rate of return on the firm's assets changes. Crouhy *et al.* (1999) further show that while it is possible to pick a capital structure so that the probability of default equals some pre-specified level, the expected rate of return on equity will change as the volatility of the rate of return on the firm's assets changes.

The insights of Crouhy *et al.* (1999) demonstrate how the contingent claims framework can be used to describe the relationship between default, capital structure and the expected rate of return. Presupposing the Merton model or structural class of models accurately describes the default risk associated with a firm's listed debt, an easy extension of logic would suggest then that the process of the firm's value must be reasonably described within the model framework. Where the Merton/structural model framework accurately encompasses default risk, the methodology provides a unique frame work for firm valuation.

Alternatively the maximum likelihood methodology, outlined in Duane (1994), provides the additional benefit of an explicit estimate of μ (The growth rate of firm value). The drift rate μ in a GBM also serves as the expected growth rate for the process, suggesting that maximum likelihood estimate of μ can be adjusted to be a reasonable proxy for the expected return to be earned on the equity security. The asset returns can be leveraged in order to provide an estimate of the return on equity from the DuPont analysis $ROE = ROA \times EM$.

The assumption that the Merton model accurately describes default risk on corporate debt is far from realistic and there is no evidence to suggest that this should automatically be the case. Fortunately, this far reaching assumption is easily circumnavigated within the Merton model framework. Opposed to Crouhy *et al* (1999) framework the probability of default is now an exogenous variable describing relationship between capital structure and expected return.

Default probabilities obtainable from the reduced form class of models provides the basis for extending the Merton model to estimate the firms value from market observable credit spreads. The probability of default is then a known constant provided from the reduced form model. The Merton model can then be reformulated with equity or firm value being used as the subject of the formula. Once again the Merton model provides a unique estimate of the firm's value based on current market information. The expected return on equity is then estimated from market credit spreads using individual capital structure and traded equity information.

2.4.2 Default probabilities from credit spreads

Under the assumption that the yield spread on the corporate bond is only owing to the compensation for the possibility of default, Hull (2012) shows that the hazard rate or default intensity can be estimated from bond prices as follows:

$$\bar{\lambda} = \frac{s}{1-R} \tag{2.4.1}$$

Where, s the yield is spread of the corporate bond over similar risk free bond and R is the expected recovery rate. The assumption is far from realistic as in practice many other factors contribute to the credit spread such as liquidity, embedded options and tax treatments of the instrument (Huang and Huang, 2003).

A key determinant of default probabilities from bond prices is the meaning of the risk-free rate or risk free bond against which the credit or yield spread is determined. Duffee (1996) notes that the treasury rate is lower than similar very low credit risk rates for a variety of factors and that the treasury rate no longer provided a suitable proxy for the risk-free rate. The tendency of treasury rates to be lower than other rates has led many market participants to regard the swap rate as an improved proxy for the risk-free rate (Hull, 2012).

The Credit Default Swap (CDS) market provides a manner in which the benchmark risk-free rate used by participants in credit markets can be estimated. CDS are considered less influenced by non-default factors and thus able to provide a good proxy of the risk-free rate when analysing default risk (Wang, 2006). The other key variable in determining default probabilities from bond prices is the expected recovery rate. The expected recovery rate for a bond is usually expressed as the bond's market value shortly after defaulting, as a percentage of its face value (Hull, 2012: 523). The expected recovery rate is thus the percentage of the original investment that an investor expects to receive in the event of default.

There are varieties of factors that influence the expected recovery rate for a bond however. Fons (1994) argues that the chief determinant of the expected recovery rate is the bond's seniority within the capital structure of the firm. Moody's estimates the recovery rates of bonds by seniority, based on bond prices one month after default. The estimation of default probabilities from bond prices and yield spreads thus requires some form of a subjective or historical estimate for the expected recovery rate.

In most studies surrounding the extracting of default probabilities from bond prices and credit spreads, such as the works of Jarrow and Turnbull (1995) along with Duffie and Singleton (1999), only plain vanilla bonds are considered in the study. Inferring default probabilities from bonds with embedded options or floating rates become significantly more complex. Estimating default probabilities for a firm from the bonds it has issued, becomes problematic for firms that issue a variety of types of bonds in addition to the plain vanilla type bonds.

Another difficulty encountered by this approach is the inability to easily separate the portion of the credit spread owing to default and the part owing to the rate of recovery. Furthermore, the findings of Elton *et al.* (2001) along with Delianedis and Geske (2001) indicate that default risk only accounts for a small proportion of the yield spread and that the greater part of the credit spread can be attributed to fiscal and systematic risk effects. This is consistent with the reasoning for the significant difference between actual default probabilities and risk-neutral default probabilities described in the previous section.

2.4.3 Credit implied equity values

Whether the probability of default is estimated from CDS spreads or bond prices, an estimate of the recovery rate is also required. Credit spreads are also affected by additional factors such as tax differences, liquidity and recovery rates (Hayne, 2004). The insights from reduced form models of default suggest that the translation of credit spreads to default probabilities may induce information losses or require additional assumptions. Thus proposed that firm value can also be solved directly from observed credit spread on debt in the structural model framework since this requires no assumption of the recovery rate and includes the possibility of capturing the creditors' view of the firm's intrinsic asset value.

Consider the same simplifying assumptions as presented in the Merton framework of the previous section. The mathematical ontology of unique estimates of firm value in this framework is now finally presented. The notation remains much the same as before however it is now shown that for an observable probability of default or credit spread the structural model can provide an alternative methodology to firm valuation.

2.4.3.1 Discount

Within this research paper the risk-neutral approach to reverse engineering firm valuations from default or credit spreads is labelled as the 'Discount' approach. Consider first the case of a market observable default probability for the single outstanding debt issue of the firm. The exogenous default probability then provides an alternate set of simultaneous equations in order to solve for the parameters of the hidden firm asset value process. The firm asset value and volatility $\{V_t, \sigma_V\}$ may now be solved from the following system of equations:

$$PD = N\left(-d + \sigma\sqrt{T - t}\right) \quad (1)$$

$$\sigma_E = \sigma V_T \frac{N(d)}{E_T} \tag{2}$$

The term PD represents the external estimate of the firms default probability in the system of equations above. Instead of using the probability of default, the Merton model also defines the credit spread or spread over the risk-free rate on corporate debt as:

$$D_t = Fe^{-(r+S)*(T-t)}$$

$$S = -\frac{\log\left(\frac{D_t}{F}\right)}{T} - r$$

The alternate expression for the value of risky debt suggests how risky debt is priced in terms of asset value and volatility.

$$D_t = V_t - E_t = V_t - BSCall(V_t, F, \sigma_V, r, T)$$

In the Merton framework value is always preserved as evidenced in the equation above. The hidden asset value process parameters $\{V_t, \sigma_V\}$ may now be solved for in the following system of equations when using exogenous credit spreads.

$$CS = S \tag{1}$$

$$\sigma_E = \sigma V_T \frac{N(d)}{E_T} \tag{2}$$

The variable CS in the system above represents the exogenous or market observable credit spread and S is the measure of the credit spread in the structural model framework as outlaid above. Then the value of equity is the value of a call option on the asset of the firm with strike price equal to the face value of outstanding debt.

$$E_T = BSCall(V_T, F, \sigma_V, r, T)$$

The value of the call option can then be regarded as the bond holder's view of the fundamental value of the assets of the firm. The expected return is then determined as discount or the premium at which the equity value E is traded in the market relative to the debt implied value of E_T.

$$Discount = \frac{E_T - E}{E}$$

The implicit assumption here is that the stock price will move towards the debt implied value of the firm. In the evaluation of the 2008 credit crisis, Hong and Sraer (2013) demonstrate that debt bubbles are quiet, high price comes with low volume. Further they illustrate that there is less scope for disagreement around the belief of the fundamental value of debt contracts. This would suggest that the market price of risk contained within debt prices may reveal a less noisy view of the value of the assets owned by the firm.

There is another reason to believe that the debt holder's valuation of the firm's asset may be more reasonable. In the structural models of default debt holders are assumed to hold absolute priority in the event of default. In the event of liquidation the debt holder's valuation of the assets is reasonably assumed to provide a decent proxy for the market value of the assets of the firm.

2.4.3.2 Return on Equity (ROE)

The methodology can be extended further when assuming that the external default probability or credit spread is a real-world measure and not a risk-neutral one. Within this research paper the risk-neutral approach to reverse engineering firm valuations from default or credit spreads is labelled as the 'ROE' approach. The expected drift rate on the asset value, μ can then also be estimated as part of the hidden asset value process of the firm. The firm's real-world hidden asset value process parameters $\{V_t, \sigma_V, \mu\}$ can be estimated from the following system of equations:

$$PD = N\left(-d + \sigma\sqrt{T-t}\right) \mid CS = S \tag{1}$$

$$\sigma_E = \sigma V_T \frac{N(d)}{E_T} \tag{2}$$

$$E_T = BSCall(V_T, F, \sigma_V, \mu, T) \tag{3}$$

The excess asset drift (EAD) is the defined as the asset drift rate above the risk free rate.

$$EAD = \mu - r$$

This provides an estimate of the excess returns expected to be earned on the assets of the firm for a given PD/CS and traded share price. The asset returns can be leveraged to provide an estimate of the expected return on equity (ROE) as per the DuPont analysis. DuPont analysis provides that $ROE = ROA * EM$ where EM is the equity multiplier, the proportion of assets financed through equity.

$$ROE = EAD * \left(\frac{V}{E}\right)$$

The ROE variable as defined above is not the same as traditional measures of return on equity in the accounting framework. Rather the shorthand for an estimate of the excess return above the risk free rate to be earned by the firm in the structural model re-appropriation.

The Delianedis & Geske (1998) model could also be used to estimate the value of the firm under more complex debt structures. The KMV re-iterative procedure for estimating the hidden asset value process of the firm can also be employed to a time series of credit spread observations to avoid the pitfalls of the simultaneous equations process.

2.5 Summary

The preceding literature review comprehensively demonstrated that theories of asset pricing in equilibrium offer applications to portfolio performance evaluation but little application towards performance prediction at the security level. In the economists models of equilibrium the incorporation of predictability in asset prices is assumed away on the basis that prices perfectly incorporate all future developments of the firm. A bastard child of the efficient market hypothesis is the Utopic notion that the price of equity and value of equity are the same.

The fundamental approach to theories of stock price behaviour claim that price and value are two distinctly separate quantities. It is assumed further that the price of equity is determined completely by its underlying fundamentals. However there is a much more complex reflexive relationship between price and value that is exploited in the market. The *greater fool* approach to investing offers explanations on the formations of speculative bubbles in markets but provides few tools for a measured investment strategy.

The structural approach proposes that the price and risk on traded debt can be used in conjunction with traded equity prices to uniquely determine the value of equity or the expected return of a security. The basis for the structural approach in fact rests on challenging the CAPM assumption that equity returns are independent of the firm's capital structure. Empirically it is well documented that the CAPM fails to describe risk and return in the investment context, which suggests that the assumptions of the model are not realistic.

There are significant differences between the company valuation methodologies in the discounted cash flow approach and the valuation methodology proposed here. The firm valuation under the Merton approach is arguably less subjective than discounting estimated expected future cash flows at the WACC. This is since the equity valuation under the structural model is implied from traded financial instrument prices as opposed to subjective estimates of companies' future earnings.

The formation of debt and equity bubbles advocated that there is less scope for disagreement concerning the fundamental value of debt contracts. This suggests that the debt holder's value of the assets is less noisy than the view of *intrinsic* value inferred from the holders of equity. Furthermore the debt holder's valuation of the firm's assets can reasonably be thought of as the minimum value of the assets in the event of firm liquidation. The subsequent chapter outlines the methodology followed in order to estimate expected returns for South African firms using the CAPM and Merton models.

3 RESEARCH METHODOLOGY

"Complete realism is clearly unattainable, and the question whether a theory is realistic enough can only be settled only by seeing whether it yields predictions that are good enough for the purpose in hand or that are better than predictions from alternative theories." (Friedman, 1953).

The insight provided by Friedman sets the background for the manner in which the research moves forward to evaluate whether structural models of equity valuations implied from market credit spreads can be used to predict returns on a stock. Discussed in detail will be the methodology regarding the estimation of the inputs required for each model's calculation of expected return. With specific reference to the theory of the models provided in the previous chapter.

3.1 TARGET VARIABLE AND FIRM SELECTION

In order to implement the structural model approach it is necessary that the firms have traded debt and equity in the market for which historical data may be obtained. In the South African context there is a shortage of firms for which historical data of credit spreads on traded bonds can be found. Owing to limitations and costliness of data in this regard, the following five companies have historical market data for both debt and equity:

Table 3.1.1 Sample of firms included

Abbreviation	Company	Sector
GRF	Group Five Limited	Construction
INL	Investec Ltd	Asset Management
BVC	Bidvest Group	Group Company
CAPITEC	Capitec	Retail Banking
ABSA	ABSA Group Limited	Financial Services

The movement in stock price or the return earned on the stock is the variable of interest to be predicted for these five firms. In this research experiment the target variable is chosen as the log return in excess of the risk-free rate earned by the stock over some length of time. Note that dividends are not accounted for in this measure of return and that the use of five companies is not ideal. However, the experimentation and analysis performed using data from these five companies may provide preliminary insight into whether the approach merits further pursuit.

The excess return is chosen as the measure of return to be predicted for in line with performance attribution or *returns* models often being formulated with the excess return above the risk-free rate as the dependent variable. Additionally the excess returns above the spread are already somewhat built in with two of the predictor variables. Firstly the CAPM regression is performed as a regression of excess equity returns against excess returns on the market over the risk-free rate, and secondly the credit spreads data set is already recorded as the spread above companion bond.

In the fundamentalist approach to stock price behaviour it is assumed that the stock price will tend towards its intrinsic value. However there is no specification of the length of time the market correction in price is expected to take. The excess return is thus defined for six separate time horizons in order to evaluate the predictions of expected return from various models. The aim is to evaluate predictive capabilities for 1 year, 6 months, 3 months, 1 month, 1 week, and 1 day excess returns above risk-free rate for the five stocks found above.

The terms 'return' and 'excess returns above the risk-free rate' are used interchangeably in the research paper for ease of readability. The next sections in the methodology resolves further the exact manner in which estimates of expected return shall be constructed. Presented thereafter is the fashion in which the performance of return estimates are measured and evaluated.

3.2 PREDICTOR VARIABLES

The usefulness of the predictors of expected return is better settled by which theory provides better predictions (Friedman, 1952). The discounted cash flow approach does not provide a means for creating daily view of expected return and requires more subjective estimates that cannot be inferred from market variables. The CAPM and estimates of expected returns from the Merton model are the primary predictors of return that shall be put to the sword in this research.

3.2.1 CAPM predictors

The CAPM model is used to produce estimates for the expected return on the firm for the five chosen firms by estimating the functional form of the equation below. Equation 3.2.1 represents the model form of the CAPM that is used to provide an estimate of the excess return to be earned on the firm. Equation 3.2.2 specifies the form of the statistical model used to estimate the parameters of the CAPM.

$$E(R) - R_f = \alpha + \beta\left(R_M - R_f\right) \qquad \text{CAPM equation} \qquad (3.2.1)$$

$$\hat{Y} = \hat{\alpha} + \hat{\beta}\left(R_M - R_f\right) \qquad \text{Statistical estimation} \qquad (3.2.2)$$

3.2.1.1 Excess return on market

The proxy for the risk-free rate is taken as the 1-year swap rate less ten basis points. Hull (2012) provides that many regard the swap rate as a better proxy for the risk-free rate than the treasury rate. The JSE Top40 is assumed to provide a reasonable proxy for the market index in South Africa. The excess return on the market is defined as the excess log return on the market index above the proxy for the risk-free rate.

3.2.1.2 Alpha and Beta

The alpha and beta coefficients in the CAPM model are estimated by regressing the daily excess log returns of the stock against the daily excess log returns over the preceding 400 day period. The daily log returns are used since they come closest to satisfying the independent identically distributed random variable assumptions required for ordinary linear regression. Kidd (2011) provides that betas estimated from more recent data may be more relevant in predicting future returns.

3.2.1.3 Summary of CAPM estimators.

The estimates of alpha and beta obtained from the regression of daily log return are then scaled by using the average period market excess return to produce the CAPM estimates of the required rate of return for the firm over different horizons. The lag window used in the average excess return on the market calculation is determined by the time horizon for which it is a predictor or measure. The beta estimates do not change for any of the CAPM estimators where as other variables are scaled or defined differently in order to produced estimates or return over different time horizons.

Table 3.2.1 Summary of CAPM predictors

CAPM: $E(R) - R_f = \alpha + \beta(R_M - R_f)$			
Predictor Variable		$R_M - R_f$	α
CAPM_1DER	1 Day excess return	Av daily excess return on market over previous week	
CAPM_1WER	1 Week excess return	Av weekly excess return on market over previous month	$* 5\ days$
CAPM_1MER	1 Month excess return	Av monthly excess return on market over 6 months	$* 21\ days$
CAPM_3MER	3 Months excess return	Av 3 Month excess return on market over last year	$* 63\ days$
CAPM_6MER	6 Months excess return	Av 6 Month excess return on market over previous year	$* 126\ days$
CAPM_1YER	1 Year excess return	Av 1 Year excess return on market over previous year	$* 252\ days$

Consider the CAPM estimate of the 1-year excess return on the firm's equity (CAPM_1YER). The estimate of alpha and beta are obtained from the regression of daily log returns. The estimate of alpha is multiplied by the assumed 252 trading days in the year to provide an estimate of yearly return. The estimate of beta does not change but the estimate of the excess return on the market is taken as the average 1-year excess return observed in the market over the preceding year. Table 3.2.1 above summarizes the scaling of estimated parameters for creating estimates of expected return under the CAPM framework.

3.2.2 Merton model expected return

In order to create estimates of expected return for the firm under the structural approach there are a large number of modeling decisions required in order to allow the firm to be describable under the simple contingent framework as laid out in section 2.3 and 2.4. The summary of the input requirements and simplifications used in order to produces estimates of expected return in the Merton framework is outlined below.

3.2.2.1 Estimating parameters of firm value process

The estimates of the firm value process $\{V_t, \sigma_v, \mu\}$ are estimated from the sets of simultaneous equations as outlined in section 2.4.3. The method makes use of the Newton Raphson algorithm to solve the non-linear simultaneous equations. The method requires initial guesses for the firm value V_T and volatility σ, the Newton Raphson method is very sensitive to these initial guesses. Crosbie and Bohn (2003) propose the following initial values for solving the system:

$$V_T = E_T + F_T \quad (3.2.3)$$

$$\sigma_V = \sigma_E \frac{E_T}{E_T + F_T} \quad (3.2.4)$$

In the case where there is an additional parameter $\{\mu\}$ the risk-free rate in the model is used as the initial starting guess for this parameter. Alternatively to using these system of equations, the re-iterative approach as found within the KMV methodology may be employed, however this proves challenging where no solution can be found to solve for parameters that match observed credit spreads or probabilities of default.

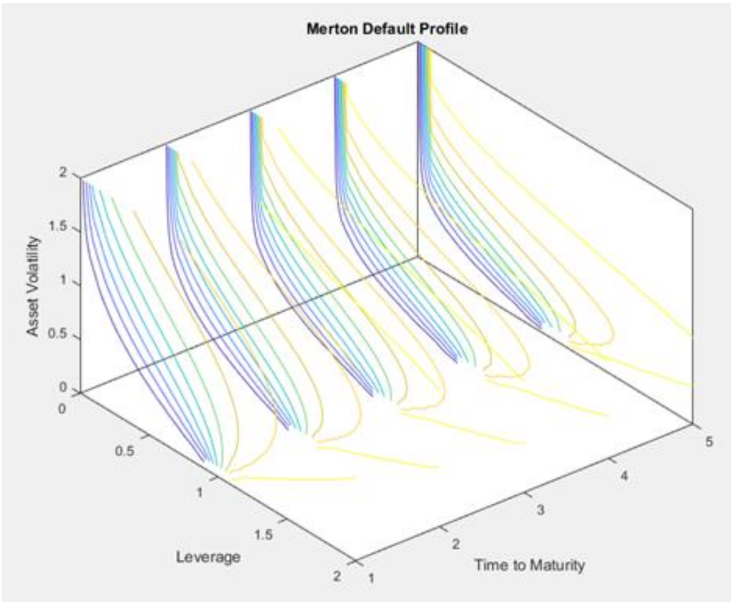
3.2.2.2 Capital structure of the firm

The outstanding debt of the firm is required to be compressed into single outstanding issue of a zero coupon bond. In this instance it is assumed that the time to maturity of outstanding debt is 5 years and that the face value of the firm's debt is captured by the total debt per share. The total debt per share is downloaded from Bloomberg and created as a daily series. The total debt per share is updated quarterly with the quarterly financial statements, although these are not the most reliable estimates. The lack of market observable amount of debt per share could be approximated by using the market capitalized value of the firm's outstanding bonds. However historical bond price series data is expensive in the South African case and is left as a suggestion for further research.

The assumption of the time to maturity of five years also seems a reasonable approximation for duration and average maturity on the debt on the firm's balance sheet. Figure 3.2.1 illustrates the default profile of the Merton model framework, graphically displaying the contour profiles of default probabilities generated under different combinations of leverage, asset volatility and time to maturity in the model framework. The graphic illustrates that under the 5 year time to maturity there is a far larger available set of unique solutions for asset value and volatility to match observed probabilities of default. This seen by the larger area under the curve corresponding to the five year maturity. The of the 5-year time to maturity on outstanding debt aids the likelihood of being able to solve for unique parameters of the firms asset value process in the simultaneous equation set up.

Alternatively to simplifying the debt structure into a single-zero coupon issue, the Delianedis and Geske model provides an alternative as shown in section 2.3.3 of the literature review**.** The Geske model is computationally expensive with bivariate normal distribution function and probability density function uses. Chen (2013) adds that there is the additional hassle of estimating the implied strike price. Moreover, the Geske model requires an intricate knowledge of the debt structure of the firm to produce two tranches of debt. The compound option model is not employed here as the aim is to minimize the required specialist assumptions regarding the firm. The increased use of specialist assumptions around the capital structure of each firm is a suggestion for areas of improvement in practical applications of the approach.

Figure 3.2.1 Merton model default profile



3.2.2.3 Market value of equity and equity volatility

The firms market value of equity is taken as the close of day price as per Bloomberg (2018). The firms daily share price is recorded for each day where there is also an observation available for the credit spread on market traded debt. The market value of equity is taken as the value of a single share as opposed to market cap value. This is since solving non-linear simultaneous equations, required in the estimation procedures of firm value and volatility, prove to be computationally more efficient when per share values are used.

Crouhy *et al* (1999) demonstrate that while it is possible to pick a capital structure so that the probability of default equals some pre-specified level, the expected rate of return on equity will change as the volatility of the rate of return on the firm's assets changes. For this reason the estimate of the volatility of the firm's equity is quite important and however unfortunately not traded in the market.

The volatility of the firm's equity returns is estimated in three different fashions. Bloomberg provides the first measure of volatility in the 360D standard deviation. The volatility of equity returns is also estimated using a GARCH $(1, 1)$ fit to the time-series of daily log returns. The GARCH $(1,1)$ model is a mean reversion model and assumes that volatility is pulled back to its long-term average at a certain rate. The GARCH $(1,1)$ model does not suffer from 'Ghost Feature' problems in volatility estimation and, hence, is expected to act as a reliable estimate for volatility of equity for the firm (Alexander, 2008).

In the long term volatility created predictor, the volatility is taken as the volatility predicted 100 days forward. Anything more than 30 days is appropriate of long term unconditional volatility estimate. We also include the GARCH $(1, 1)$ predicted 1-day ahead as a model for the firm's estimated future volatility. Estimates of volatility are then annualized to be used as inputs in the structural model framework.

3.2.2.4 Risk-free rate

The swap rate less 10 basis points is used as the proxy for the risk-free rate. This is consistent with the proxy used for the risk-free rate in the CAPM determination of excess returns for the firm.

3.2.2.5 Probability of default

The probability of default used as the market input in the structural models matches the time to maturity on the simplified debt structure. The Bloomberg proprietary model for default provides the daily time-series for the 5-year probability of default. The Bloomberg probability of default advantageously requires no assumptions regarding the expected recovery rate on the bond issue in order to obtain an estimate for default. Bloomberg is ubiquitously used by most investors so much so that the information from Bloomberg can be regarded as publicly available information.

3.2.2.6 Credit spread

The credit spread is taken as the traded yield above the bond's companion bond. The credit spreads were obtained from Inet Bridge courtesy of Mr Carel van der Merwe. The Credit Spread is also assumed to be the 5-year spread. Where there is more than one outstanding bond for the firm in question, the credit spread is then determined as the market capitalized weighted credit spreads from the bonds market capitalization values.

3.2.2.7 *Summary of predictor variables created*

The predictor variables are named such that each term in the label describes methodology and inputs of the estimator. The prefix describes the method of structural model predictor, the middle letters denote the estimate of equity volatility uses and the suffix denotes the debt market variable used. For example the 'ROE_360D_BB' estimate is constructed from the ROE method, using the 360 day standard deviation as the volatility, and finally the Bloomberg probability of default is used as the debt market variable.

Table 3.2.2 Summary of explanation of structural model predictors created: predictor label describes the predictor methodology. The prefix describes the method, middle letters denote the estimate of equity volatility and the suffix narrates the debt market variable used in the reverse engineering.

Exogenous Variables	Et:	*PX_LAST close of day trading price*
	T:	*5 Years to Maturity on outstanding Debt*
	F:	*Total Debt per share*
	r:	*Swap rate less 10 bps*
	BB 5Y PD	*Bloomberg 5Y Default Probability*
	MC Spread	*Market Capitalized weighted Credit Spread on Bond*

Summary of Predictor Variates Created under Structural Approach

Predictor Variable	**Method**		**Equity Volatility**			**Market Debt Variable**	
	Discount	*ROE*	*360D Volatility*	*GARCH(1,1) n=100*	*GARCH(1,1) n=1*	*MC Spread*	*BB 5Y PD*
Discount_360D_BB	■		■				■
Discount_360D_CS	■		■			■	
Discount_LTVol_BB	■			■			■
Discount_LTVol_CS	■			■		■	
Discount_GARCH_BB	■				■		■
Discount_GARCH_CS	■				■	■	
ROE_360D_BB		■	■				■
ROE_360D_CS		■	■			■	
ROE_LTVol_BB		■		■			■
ROE_LTVol_CS		■		■		■	
ROE_GARCH_BB		■			■		■
ROE_GARCH_CS		■			■		■

In Table 3.2.2 the acronyms of created predictor variables as they will be seen in the results are shown. The highlighted cells illustrate the method, assumptions and inputs used in each of the predictor variables as identifiable by the created label.

3.3 MODELING PROCESS

"It is a mistake to use, as some journalists and some economists do, statistics without logic, but the reverse does not hold: It Is not a mistake to use logic without statistics." Taleb (2004).

It is not incorrect to follow the rational and logic and then determine the quantification of the hypothesis. However it is incorrect to follow a modeling process without rational upon which hypothesis are falsely generated on the basis of explaining random predictive capabilities instead of meaningful causality in prediction. The basis for the use of statistical models and methods is carefully motivated here.

3.3.1 Statistical learning methods

The CAPM is not a direct estimate of the expected return of the firm on a stand-alone basis. The CAPM provides the expected return when the stock is part of a well-diversified portfolio in an efficient well-functioning securities market. In order to forecast the expected return on the stock, forecasts of beta and forecasts of the excess return on the market would be required (Kidd, 2011). The expected rate of return or required rate of return from the CAPM is only a forecast of future return where markets are perfectly efficient and future stock prices cannot be predicted under the EMH setting. The point is that in the EMH setting there is no need to distinguish past from future performance. Structural models also have a large number of simplifying assumptions in the framework, under which the estimates produced may be considered direct estimates of expected return.

The CAPM theory as well Merton's theory of the firm are theories that describe perfect worlds and should be adjusted accordingly when being implemented in vastly more complex reality. The proposal is to then account for this uncertainty by introducing a statistical learning method to see how well the structural model variables and CAPM variables can be used as a predictor of the direction of future excess returns. More generally the statistical learning method framework provides that an observed quantitative response Y is related to p potential predictors through some relationship of the general form:

$$Y = f(X) + \varepsilon \tag{3.3.1}$$

Here f is some fixed but unknown function of the predictor variables $\{X_1, \ldots, X_p\}$ and f represents the systematic information that the predictors provide about the response variable. If prediction is the chief concern of employing a statistical learning method, the exact form of the relationship between the response variable and predictor is not of much concern provided that it yields accurate predictions for Y (James *et al.*, 2015).

3.3.2 Further revision of target variable

The CAPM predictor already estimates the excess return for the share in nominal terms. This means that it is possible to measure the realized future returns against the CAPM predictions and find a mean squared error (MSE) as well as correlation between the predictions made and actual returns realized. However, MSE and correlation are average measures that are not readily comparable and do not easily lend to meaningful interpretation in the characterization of prediction performance. The MSE measures provides no indication of the types of errors made, under or over prediction etc.

In the Merton model of the firm the debt structure and time to maturity on outstanding debt has been estimated rather loosely in the procedure. These predictors are more appropriately thought of as being relative indicators of future performance as opposed to absolute nominal predictors. The excess return is further classified into a dichotomous stratification of positive and negative excess returns. Furthermore splitting the performance into an easy categorization of up and down allows for extended analogy of credit risk methodologies in the forecasting of future equity excess returns.

The response variable or target variable is now a categorical one, shifting the problem into the *classification* setting as opposed to *regression* problems in the statistical modeling framework. The classification of excess returns split around zero is reduction of the return vector into a directional focused component. If the predictors and learning methods provide good forecasts of direction of future returns then question can be extended further towards quantification of magnitude of future stock returns. It is interesting to note that with enough classes the approach is analogous to predicting quantitative nominal excess returns. Since it is the excess returns that are stratified into two classes there is actually still a nominal component even though we have moved to a categorical space.

Predicting exact future asset returns is so difficult that many economic theorists have explained it away with the EMH or random walk hypothesis. The classifications of future excess returns into categories of up and down for negative and positive would assumedly be easier to predict than the exact nominal amount. Ideally the classification of excess return brackets would correspond to decision making frameworks for asset managers. The choice of predicting the class of returns is assumed to lend itself well to practical applications for investors. Suggestion for further research is then the considerations of the extension of classes. Ideally, investment strategies should like to consider three classes to predict when individual firms will make: big losses, big gains or average returns in relation to the risk-free rate.

3.3.3 Link function

Depending on whether the ultimate goal is prediction, inference, or a combination of the two, different methods for estimating f may be appropriate. Linear models allow for relatively simple interpretable inference but may not yield as accurate predictions as some other approaches (James *et al.*, 2015). The research question does not include an exact specification or interest in the form of the relationship between the predictors of return and the return. Considerations of the exact nature of the relationship to adjust for uncertainty in the models is beyond the scope of this research and runs the risk of requiring the generation of a false hypothesis to explain.

The primary aim is to assess how well the predictor variables can be used to predict the direction of future excess stock returns. In this setting, it might be expected that using the most flexible learning method will yield the most accurate predictions at the expense of interpretability. However, James *et al*. (2015) provides that more flexible methods suffer from over fitting and may actually produce less accurate predictions than those from more inflexible statistical learning methods. There is no free lunch in statistics: no one method dominates all others over all possible data sets. In this research experiment the elementary non-parametric and parametric models for classification problems will be employed.

3.3.3.1 Logistic regression

Logistic regression may be viewed as an extension of ordinary linear regression to adjust for the case where the response variable of interest may be categorical, as opposed to continuous in the ordinary linear regression case (James *et al.*, 2015). The model is constructed upon the same assumptions as ordinary linear regression, although the response is not directly predicted. The response is instead modeled as the probability of belonging to a particular category.

According to Baesens, Rosch and Scheule (2016), simple logistic regression considers a binary categorical response variable. The logistical function is used to transform continuous response predictions into probabilities of belonging to a particular category, often referred to as the 'link' function in this setting. Ubiquitous applications of binary logistic classification approaches include fraud detection, Target Marketing and Credit Scoring to name a few. In the credit scoring environment the relevant binary target is usually whether the case (loan) defaulted or not.

Linear Discriminant Analysis (LDA) and classification trees, are popular statistical alternatives to categorical classification prediction. James *et al.* (2015) notes that, curiously the LDA procedure produces the same classifications as ordinary linear regression used to predict a binary response. In this research only the logistic regression approach is considered for parametric approaches, although it is worth noting the available alternatives.

Furthermore, Patetta (2010) provides that traditional inference can be important in predictive modeling as well. The traditional linear regression model assumptions must hold for any valid inference to be made from the logistical regression procedure. Although, in practice this is often not the case and the discovery of structure is informal and exploratory. This is since the validity of predictive models is preferably assessed by empirical performance and the ability to generalize. James *et al.* provides the following formula used to predict novel cases within the logistic regression framework.

$$P(X) = \frac{e^{\beta_0+\beta_1 X_1+\cdots+\beta_p X_p}}{1+e^{\beta_0+\beta_1 X_1+\cdots+\beta_p X_p}} \tag{3.3.2}$$

$$\log\left(\frac{p(X)}{1-p(X)}\right) = \beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p \tag{3.3.3}$$

Where $X = (X_1, \ldots, X_p)$ are p potential predictors. The model can also be represented in linear form to calculate what is commonly termed 'the Log Odds' ratio and is expressed as in Equation 3.3.3. Maximum Likelihood Estimation (MLE) is used to estimate the coefficients in Equation 3.3.2 using the chosen p predictors or explanatory variables. The coefficients can also be estimated using least squares, although James *et al.* (2015) notes that the use of MLE produces more desirable statistical properties of the estimators.

3.3.3.2 K-Nearest Neighbours (KNN)

The *K-nearest neighbours* (KNN) classifier is a non-parametric statistical learning method. The non-parametric approach makes no assumption regarding the distribution or shape of f. James *et al.* (2015) adds that these methods have the benefit of the potential to accurately fit a wider range of possible shapes for f.

Given a positive integer K and a test observation x_0, the KNN classifier first identifies the K points in the training data that are closest to x_0, represented by η_0. It then estimates the conditional probability for class j as the fraction of points in η_0 whose response values equal j:

$$P(Y = j | X = x_0) = \frac{1}{K} \sum_{i \in \eta_0} I(y_i = j). \tag{3.3.4}$$

Finally, KNN applies Bayes rule and classifies the test observation x_0 to the class with the largest probability. Despite the fact that it KNN is a very naïve approach it can offend produce classifiers that are surprisingly close to the optimal Bayes classifier.

The choice of K on the number of nearest neighbours in the KNN approach has a drastic effect on the KNN classifier obtained. The choice of neighbours is directly correlated with the flexibility of the decision boundary estimated in the KNN approach. In both the regression and classification settings, correctly choosing or identifying the level of flexibility is critical to the success of any learning models.

In this research experiment arbitrary low and high values of K are used in the KNN approach in order to find the appropriate level of flexibility.

3.3.4 Performance evaluation

Theory of the CAPM does not highlight important difference between explaining variation in past returns and predicting future returns in the implementation of the CAPM. The predictors are firstly evaluated by how well they perform as stand-alone predictors of the direction of excess return on the stock. In this since the estimates of expected return from the Merton and CAPM model are treated as pure forecasts of future return.

3.3.4.1 Training data & test prediction set

The predictors are then employed in a statistical learning model to evaluate how well they may be used to forecast the direction of future excess returns on the stock. Furthermore an initial training period and then a predict one and update process is employed. This is done for the excess return as measured and defined over the six different time horizons.

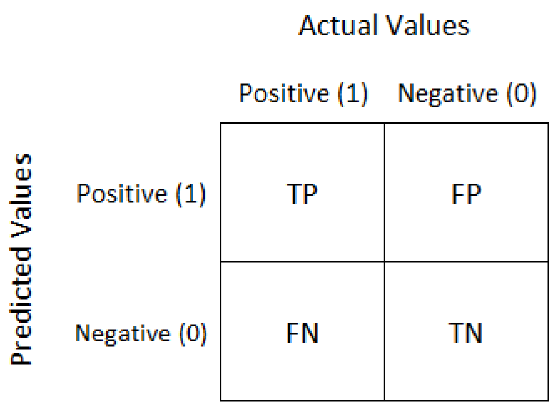
For the majority of the firms an initial training set of 1500 days is used. This is perhaps not the most accurate description of the out of sample model diagnostics procedure. This is since predictions are made for observation 1501 in the set and then updated until the end of the sample period. They key is that if a prediction is being made for the excess return to be achieved 1-year from observation 1501, then only $(1501 - 252)$ observations can have labelled outcomes for the training of such a model. That is, if we are standing at observation 1500 we cannot know what the excess yearly return was for any observation after (1500-252) in the sample set.

The models should be validated on the out of sample model diagnostics or backtesting procedures as Alexander (2008) defines. The models should not be evaluated on the training error or ability to fit the data that was used in the construction of the estimation of model fit. Model validation using measures from the fitted training sample is parallel to staring in the mirror to catch yourself blink.

3.3.4.2 Test prediction performance measures

Once the out-of-sample test predictions have been made the question still remains on how to best assess the performance of the models. One of the reasons given for moving from the regression setting to the classification setting in the statistical problem setting is that performance measures from categorical predictions provide more granularity in the characterization of prediction performance. James *et al.* (2015) provides that the confusion matrix describes the performance measurements for binary classification problems. The confusion matrix in Figure 3.3.1 illustrates the basis for performance measurements in the binary classification problem setting.

Figure 3.3.1 Confusion matrix: classification prediction performance measures



There are numerous metrics that can be spawned from the confusion matrix to evaluate the performance of test predictions. However the following measures will be included in this research:

$$FPR = \frac{FP}{TP+FP+TN+FN} \tag{3.3.5}$$

$$FNR = \frac{FN}{TP+FP+TN+FN} \tag{3.3.6}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{3.3.7}$$

$$Sensitivity = \frac{TP}{TP+FN} \tag{3.3.8}$$

$$Specificity = \frac{TN}{TN+FP} \tag{3.3.9}$$

$$Precision = \frac{TP}{TP+FP} \tag{3.3.10}$$

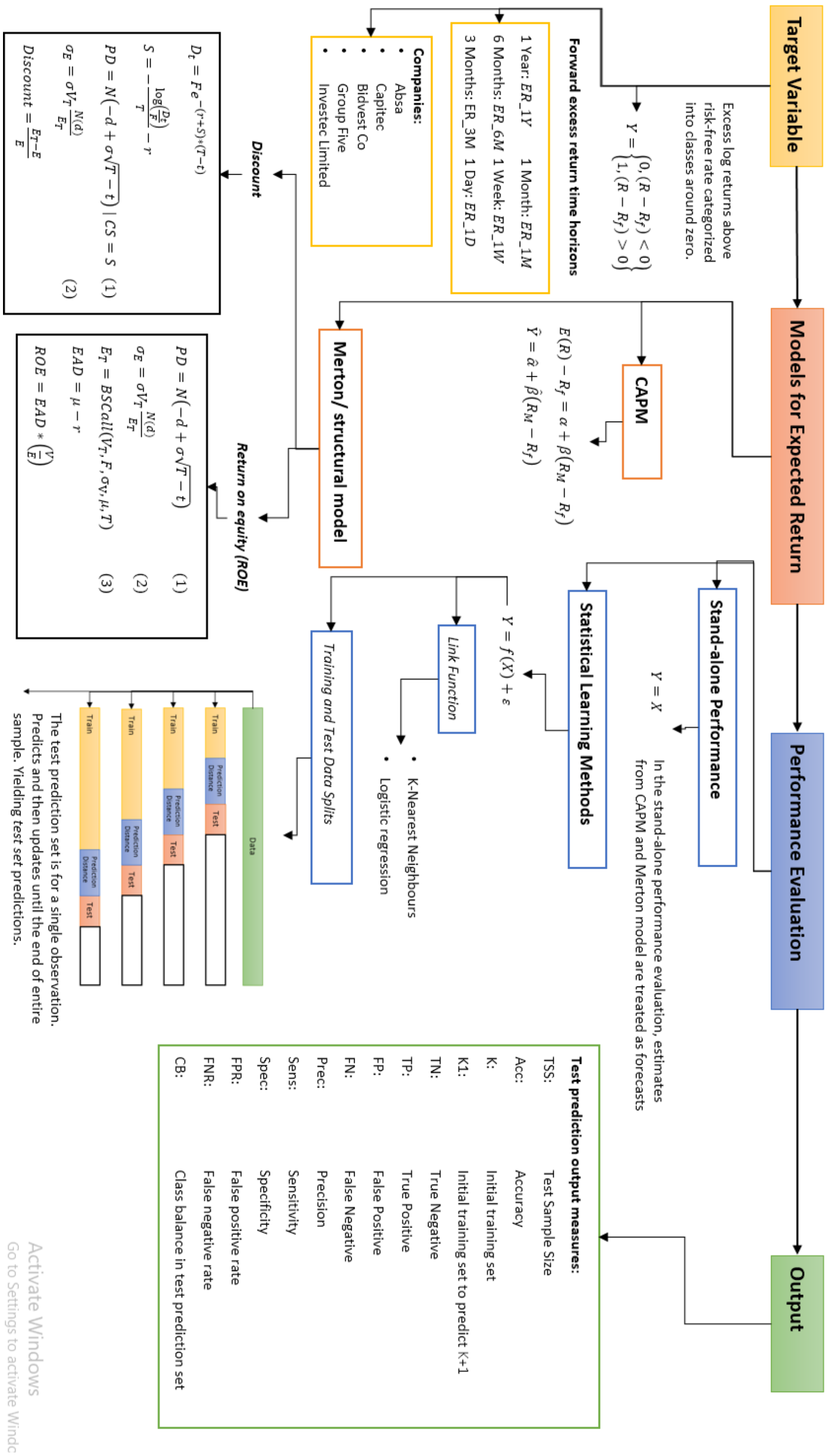
The false negative rate (FNR) and false positive rates (FPR) provide the measures of the types of classification errors made by the model in the test prediction set. Accuracy is the simplest scoring measure which simply calculates the proportion of correctly classified instances. Patetta (2010) contends that accuracy should not be the primary measure of interest where the test sample is dominated by a majority class. Sensitivity is also called recall and captures the proportion of actual positives which are correctly identified as positives by the classifier.

Specificity relates to the classifiers ability to identify negative results. Similar to recall specificity captures the proportion of actual negatives which are correctly identified as negatives by the classifier. The precision measure indicates the proportion of positive predictions made by the classifier which were in fact correct. No single performance measure will identify the best classifier model for all possible purposes. The prediction performance of a model should be judged by evaluating a combination of these measures in order to assess the appropriate usefulness of predictions.

3.4 SUMMARY OF RESEARCH METHODOLOGY

Figure 3.4.1 graphically illustrates the research methodology process and provides a comprehensive overview of the thinking process in the research. The R-code used for the methods and models discussed in presented in Appendix F. The results obtained from applying the outlined methodology and procedures along with the concurrent analysis of these results are presented in the next chapter.

Figure 3.4.1 Summary of research methodology



4 FINDINGS AND OBSERVATIONS

"*In God we trust, all others must bring data.*" William E. Deming (1900-1993)

As so concisely summarized by Deming, the only way to truly evaluate whether CAPM estimates or Merton model estimates of return can be used to predict excess returns, is to put the theory to the test. The results for predicting the direction of the excess return for the five stocks over the test samples is now presented.

4.1 ABSA GROUP LTD RESULTS

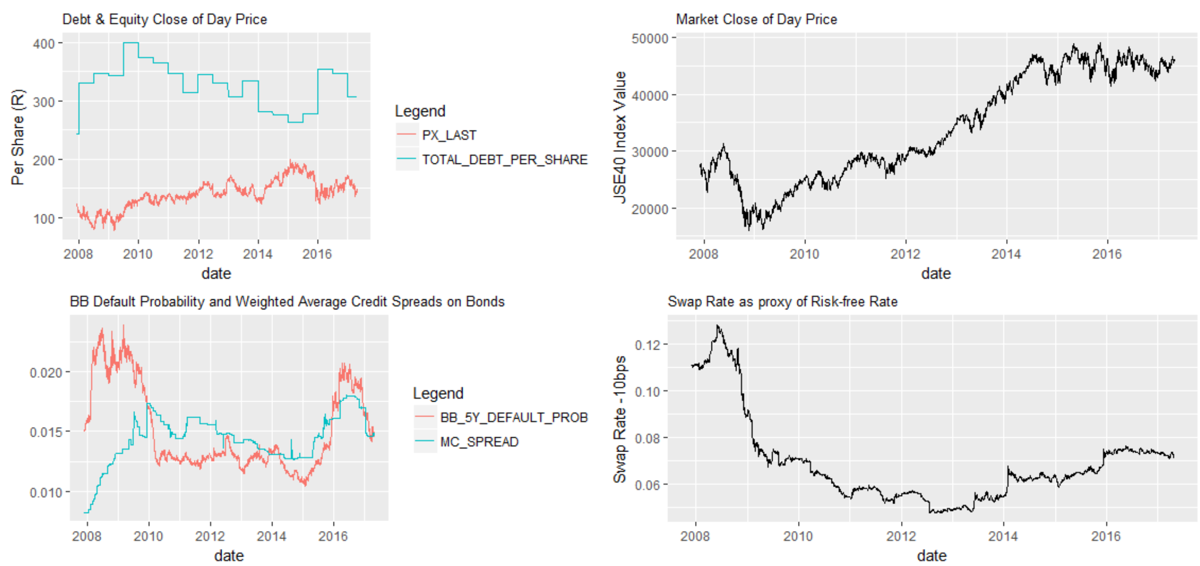
The results of predicting the direction of the excess log returns for Absa Group Ltd., hereafter referred to as Absa, is now presented in full detail. The market variable inputs along with the created predictors of excess return are first discussed briefly before moving to the results of class prediction.

4.1.1 Target and predictor variables

4.1.1.1 Input variables

Figure 4.1.1 captures the debt and equity characteristics for Absa over the period of 2008 to 2017. From the figure it is clear that the total debt per share is only updated quarterly in line with only being observable on publication of financial statement data. The highlight here is the amount of leverage being employed by Absa, the debt per share far exceeds the traded price of equity per share. Given Absa's large use of debt, priori expectations provide that debt should be a good determinant of future performance.

Figure 4.1.1 Absa market and firm input variables



4.1.1.2 Structural model predictor variables

Figure 4.1.2 graphically illustrates the predictors created under the structural approach for Absa over the period of 2008-2017. The large spikes in the graphics illustrates where the simultaneous equations used in the procedure were unable to solve for a unique value in the algorithm. The 'Discount' variables and the 'ROE' variables are on the opposite sides of zero, barring the unsolved points. This is since in the Merton framework, value is always preserved when estimating firm value from the observed credit spread. The weakness of preservation of value in the estimation procedure is the primary reason for evaluating these predictors in relative terms. Interestingly the graphic illustrates that the firm values implied from the probabilities of default do not necessarily share this crux.

Figure 4.1.2 Absa structural model predictor variables

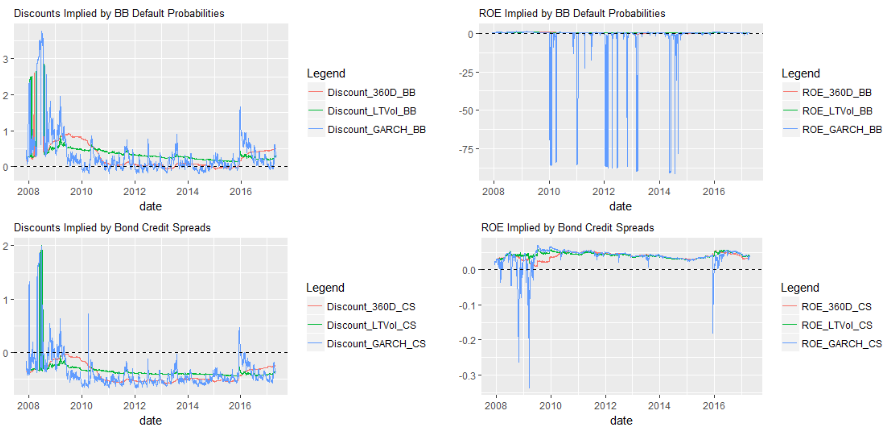


Table 4.1.1 confirms the number of data points where no unique solution was found in the sample set for Absa. The reader is referred to Appendix A for the graphical analysis of the estimates of volatility for the Absa stock price over the period. The highly erratic nature of the volatility estimates from the GARCH models are confirmed by the number of points in the data set where these inputs could not be used to solve for unique firm asset value process parameters. The exact shape of the return estimates is far better examined in conjunction with the actual returns being estimated.

Table 4.1.1 Absa convergence to solution in simultaneous equations: 1-convergence satisfied; 2-solution is uncertain; 3-no better solution found than starting point; 7 Jacobian is unusable.

ABSA TermCD Check										
	1	2	3	4	5	6	7	8	9	10
TermCD Discount_360D_BB	2245	41	59	0	0	0	0	0	0	0
TermCD Discount_LTVol_BB	2260	48	37	0	0	0	0	0	0	0
TermCD Discount_GARCH_BB	2274	38	33	0	0	0	0	0	0	0
TermCD ROE_360D_BB	2345	0	0	0	0	0	0	0	0	0
TermCD ROE_LTVol_BB	2345	0	0	0	0	0	0	0	0	0
TermCD ROE_GARCH_BB	2269	64	12	0	0	0	0	0	0	0
TermCD Discount_360D_CS	2324	2	19	0	0	0	0	0	0	0
TermCD Discount_LTVol_CS	2319	2	24	0	0	0	0	0	0	0
TermCD Discount_GARCH_CS	2314	6	25	0	0	0	0	0	0	0
TermCD ROE_360D_CS	2345	0	0	0	0	0	0	0	0	0
TermCD ROE_LTVol_CS	2345	0	0	0	0	0	0	0	0	0
TermCD ROE_GARCH_CS	2345	0	0	0	0	0	0	0	0	0

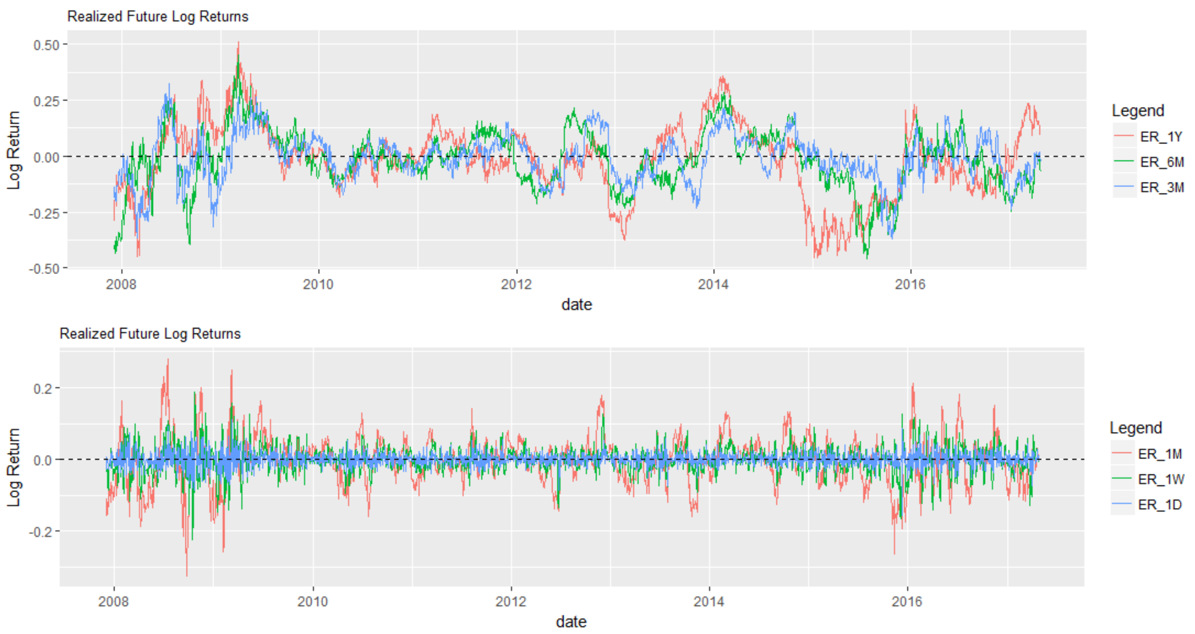
4.1.1.3 Realized forward excess log returns

Figure 4.1.3 illustrates the time-series of realized forward excess log returns defined over different horizons. The forward nature of the returns means that the 1-year excess return (ER_1Y) plotted at 2016 is interpreted as the excess return that is realized in one-year from that date.

The short term horizon excess returns are extremely noisy and it would appear extremely difficult to predict. This is in line with investor theory expectations that fluctuations in the short run are noisy but prices are driven to their equilibrium value over the longer period. The logic also provides that using issues of debt assumed to mature in 5 years implies that equity valuations under these views would also correct over the longer time horizon as opposed to instantaneously in the market.

The three and six month excess returns along with the 1-year excess returns for Absa are much further from white noise. Additionally they display the ideal property of fluctuating around zero, this means that predictive performance evaluated on prediction of direction (up or down) does yield some insight into discriminatory power of different variables and approaches.

Figure 4.1.3 Absa realized forward excess log returns over varying time horizons



4.1.2 Univariate prediction performance

Here, the set of predictions where each estimate or predictor of excess return is used in isolation is covered in the analysis presented. The analysis begins by considering how well the predictors perform as classifiers of future return without the implementation of a statistical learning method.

4.1.2.1 Indicator

The indicator analysis is essentially a classification function with a constant boundary of zero. If the predictor variable is a good estimate of nominal excess return then should also be able to forecast direction of excess log return using a fixed boundary of zero.

This applies more to CAPM than it does the other variables, since under the use of CAPM it is argued that past asset returns are a good indication of future asset returns in an efficient market and thus an acceptable manner in which to solve the portfolio selection problem. Table 4.1.2 provides the performance of predictors the class or direction of future excess returns on the Absa stock. The indicator function requires no training data since the predictor variables are not used in a statistical learning method.

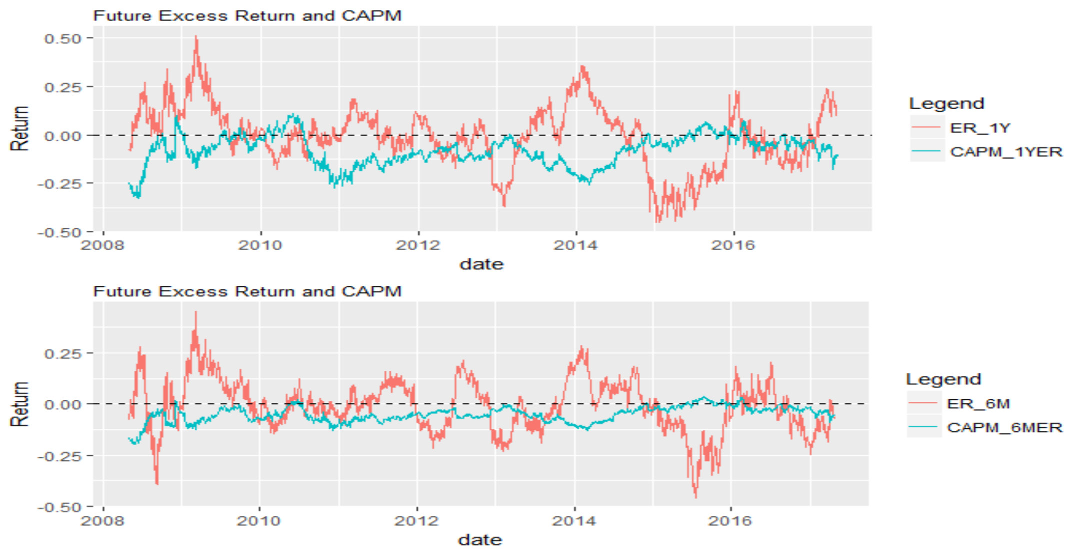
However, the results found in Table 4.1.2 correspond to predictions from after 1500 days in the sample set for each of the predictors – this is for comparability with predictions from the employment of statistical learning methods. There are 845 days for which predictions are made when using 1500 observations for the training data. Appendix A provides the full set of results for the indicator performance measures. The CAPM excess return predicts the direction of the future excess return for the Absa share only 40% of the time. This is not surprising since the CAPM model is validated on the basis of how well it explains past variations in the returns.

Table 4.1.2 Absa accuracy of expected return estimates evaluated as stand-alone predictors of future equity return class. (Class predictions are made using constant boundary of zero on the return estimate)

	1 Year	6 Months	3 Months	1 Months	1 Week	1 Day
ROE_360D_BB	67%	**70%**	67%	52%	51%	51%
Discount_GARCH_CS	65%	57%	55%	48%	49%	51%
Discount_LTVol_CS	63%	57%	55%	50%	49%	51%
Discount_360D_BB	62%	67%	63%	49%	50%	51%
Discount_360D_CS	62%	57%	54%	50%	49%	51%
CAPM	49%	44%	40%	45%	45%	48%
ROE_LTVol_BB	39%	45%	47%	50%	51%	49%
ROE_360D_CS	39%	45%	47%	50%	51%	49%
ROE_LTVol_CS	39%	45%	47%	50%	51%	49%
ROE_GARCH_CS	38%	45%	47%	51%	51%	49%
Discount_LTVol_BB	32%	39%	41%	47%	50%	48%
ROE_GARCH_BB	29%	36%	44%	50%	50%	50%
Discount_GARCH_BB	28%	35%	43%	49%	50%	50%

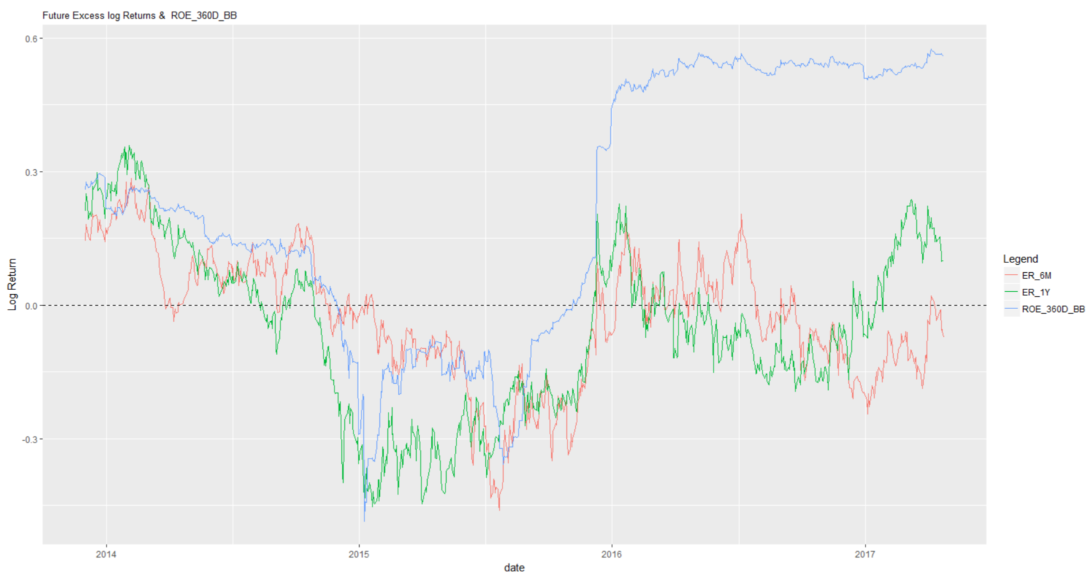
The indicator results clearly reaffirm that the past return is not necessarily a good predictor of future returns. The Figure 4.1.4 below illustrates the CAPM estimates of *1-year* and *6-month* excess returns against the *1-year* and *6-month* excess returns actually realized on the Absa stock. Examining the full sample set it becomes unambiguous that CAPM employed in its current fashion is an explanation of past returns and not a predictor of future asset returns. This is clearly visible by shifting the returns forward to where they are actually realized, or shifting CAPM backwards to show that CAPM co-integrates very well on historic return series.

Figure 4.1.4 Absa CAPM estimated return and forward realized excess returns over entire sample



In the test sample of 845 days we see that the CAPM required rate of return is in fact a dismal indicator of the future excess return. In relative terms the ROE predictor created from the Bloomberg 5-year default probability does provide a much better indication or predictor of future returns. Figure 4.1.5 illustrates that the *ROE_360D_BB* predictor of expected return and the forward 1-year and 6-months realized excess returns on the Absa stock form 2014-2017. In this particular sample space the predictor appears to do a relatively amazing job in forecasting the direction of excess returns one year into the future.

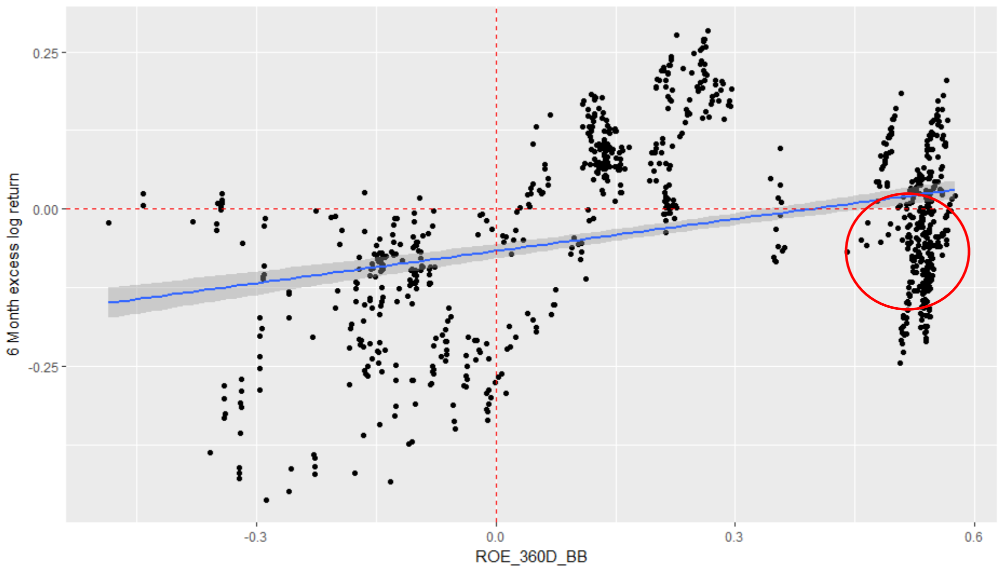
Figure 4.1.5 Absa 1-year and 6-months excess log returns with ROE_360D_BB predictor from the period 2014- 2017.



Although the initial presumption was that the structural model variables would fare better as relative indicators, the *ROE* variable appears to track the direction of the future 6 month excess return sufficiently. Especially in the period up until 2015, almost a very good nominal forecast of future return. The preservation of value for predictors of return created from credit spreads does not allow for predictors under that measure to change direction. The fact that estimates created from default probabilities track the future performance of the firm so well in bot direction is a bit surprising but also undoubtedly a win.

Figure 4.1.6 illustrates how well the predictor is able to correctly classify the direction of future excess return on the stock over the 6 month horizon. It is interesting to see that when the *ROE_360D_BB* estimate is negative the ABSA share return is almost always negative, this is evidenced by the clear left quadrant in Figure 4.1.6. The majority of classification errors are around zero and the far right end scale of the *ROE_360D_BB* predictor variable.

Figure 4.1.6 Absa: scatterplot of 6M forward excess returns vs. ROE_360D_BB predictor



The discriminatory power of the predictor variable constructed from the Merton framework show promising signs for the ability of these predictors to accurately make predictions around the future excess return on the Absa share price.

4.1.2.2 K-Nearest Neighbours

The K-Nearest Neighbours methodology is employed on a training data set of a maximum of 1500 days. Predictions are made for day 1501 till the end of the sample period. The number of neighbours is arbitrarily chosen as 99, so the non-parametric estimation is rather inflexible in this instance. Table 4.1.3 summarizes the performance of the predictions made using the different constructed predictor variables.

There are a number of interesting results in the prediction performance using the KNN approach. Firstly the ROE_360D_BB variable does worse than when the decision boundary was taken as constant around zero. This would suggest that the inflexibility of the large number of neighbours used has a poor impact on performance. The performance for the predictors created under the CAPM are significantly improved over longer time horizons. The performance of the CAPM predictors is improved far more drastically for the 1 year returns versus the 6 month returns. The difference in performance improvement might be the impact of how rolling historic average return on the market is used in the formulation CAPM predictors.

Table 4.1.3 Absa KNN (K= 99) test prediction accuracy for returns defined over various time horizons.

	1 Year	6 Months	3 Months	1 Months	1 Week	1 Day
CAPM	**68%**	**53%**	55%	52%	55%	52%
ROE_360D_BB	**67%**	**51%**	63%	49%	49%	50%
Discount_LTVol_CS	63%	44%	42%	51%	48%	50%
Discount_360D_CS	60%	51%	46%	45%	47%	51%
ROE_GARCH_BB	59%	52%	53%	49%	47%	51%
Discount_GARCH_BB	56%	41%	46%	45%	52%	50%
Discount_360D_BB	54%	67%	53%	46%	47%	50%
ROE_LTVol_CS	46%	35%	44%	55%	53%	50%
Discount_GARCH_CS	43%	38%	44%	44%	49%	52%
ROE_360D_CS	43%	39%	47%	55%	50%	48%
Discount_LTVol_BB	39%	27%	48%	44%	46%	52%
ROE_LTVol_BB	38%	40%	51%	43%	43%	50%
ROE_GARCH_CS	37%	39%	44%	49%	54%	49%

Figure 4.1.7 and Figure 4.1.8 both illustrate the poor performance of the K-Nearest Neighbours for estimating the relationship between the predictors of return and the actual return achieved by the Absa share. The predictor variables are plotted in black in the figures below whereas the excess return is plotted in different colours corresponding to the prediction classifications made.

Figure 4.1.7 Absa KNN (K=99) 1-year return predictions with CAPM predictor variable. The CAPM estimate of return is represented by the black line in the figure. The multi-coloured points represent the future 1-year returns, each colour representing the prediction performance against the actual realized returns.

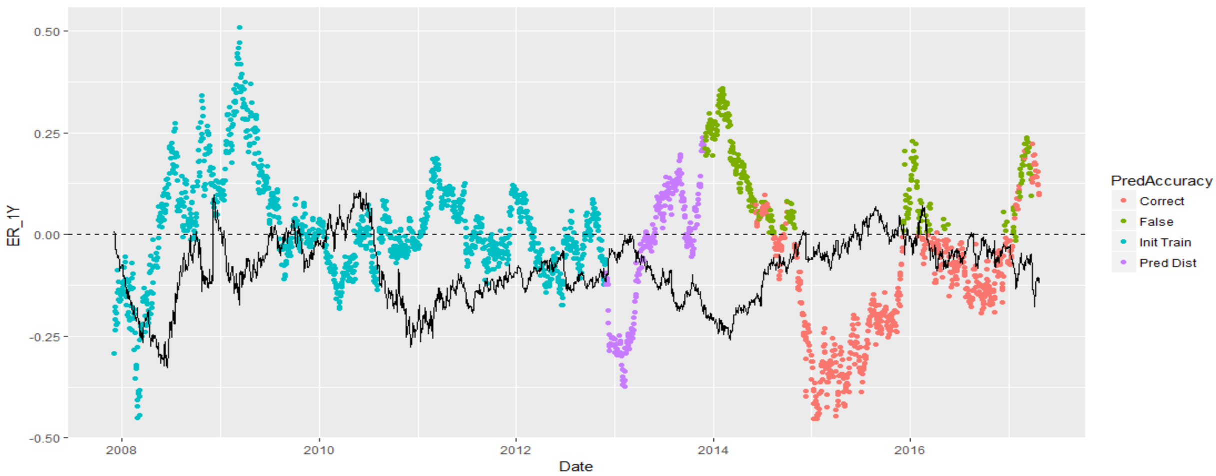
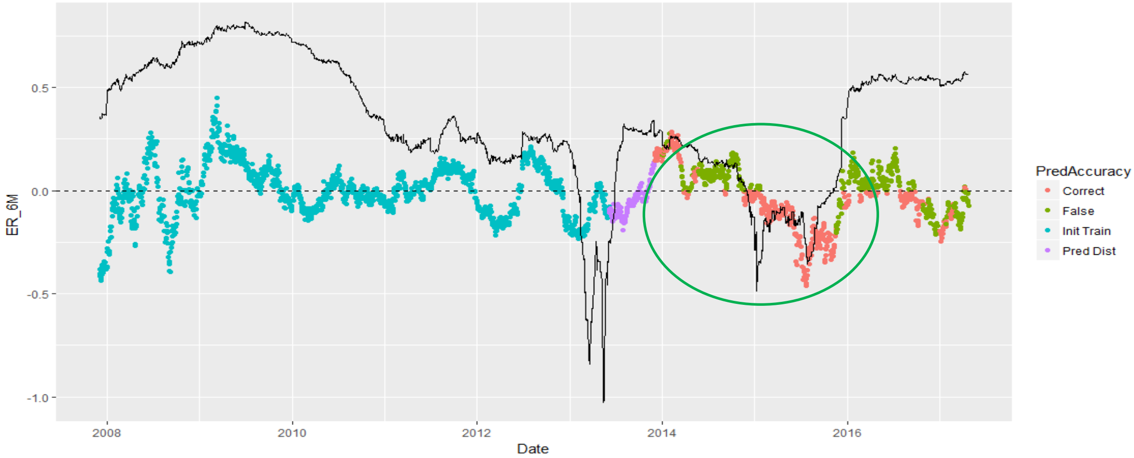


Figure 4.1.8 Absa KNN (K=99) 6-month return predictions with ROE_360D_BB as predictor variable. The ROE_360D_BB estimate of return is represented by the black line in the figure. The multi-coloured points represent the future 6-month returns, each colour representing the prediction performance against the actual realized returns



4.1.2.3 Logistical regression

Table 4.1.4 summarizes the performance of the predictions made using logistic regression to specify the relationship between predictors and the actual forward return. The predictor from the CAPM now performs quite well for predicting 1-year returns. It is also interesting to note that the performance of the ROE_360D_BB variable has drastically deteriorated again, this is indicative of the impact of the learning lag or incorrect specification of the functional form of the relationship between the target and predictor variables.

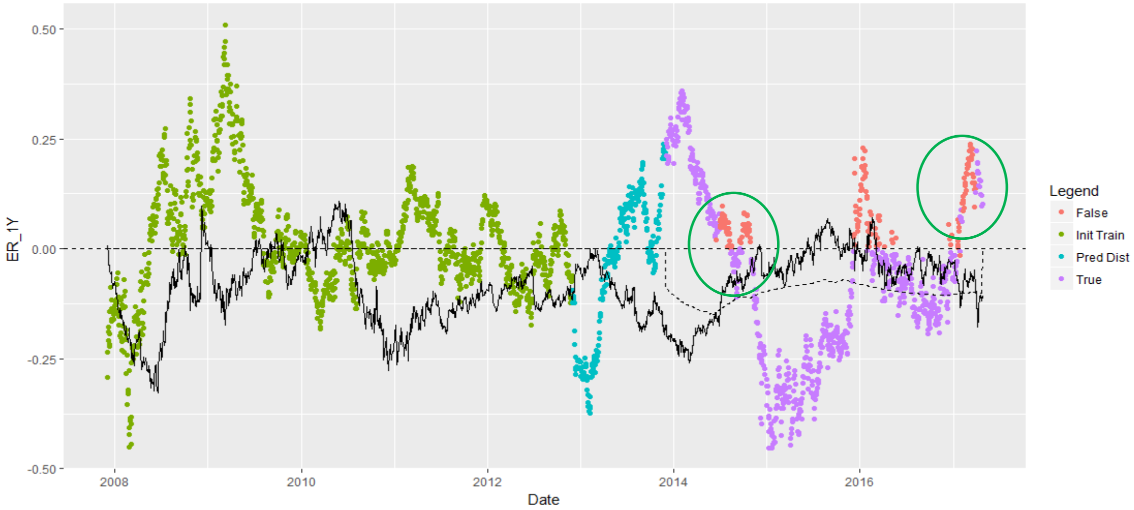
Table 4.1.4 Absa logistic regression prediction performance. Test prediction accuracy for predictions made using logistic regression with various predictor variables.

	1 Year	6 Months	3 Months	1 Months	1 Week	1 Day
CAPM	**80%**	53%	59%	59%	49%	51%
Discount_LTVol_BB	69%	62%	59%	51%	51%	52%
Discount_LTVol_CS	64%	57%	54%	43%	48%	51%
ROE_LTVol_BB	64%	57%	55%	51%	49%	51%
Discount_360D_BB	57%	53%	53%	48%	50%	52%
Discount_GARCH_BB	56%	48%	49%	46%	44%	51%
Discount_360D_CS	52%	47%	50%	43%	45%	51%
ROE_360D_BB	**51%**	48%	49%	45%	50%	51%
Discount_GARCH_CS	46%	40%	43%	47%	43%	51%
ROE_GARCH_CS	37%	29%	39%	45%	49%	51%
ROE_360D_CS	35%	39%	45%	44%	49%	51%
ROE_GARCH_BB	32%	33%	40%	46%	48%	51%
ROE_LTVol_CS	31%	53%	54%	51%	53%	51%

Figure 4.1.9 illustrates the improved nature of 1-year return predictions made using the CAPM predictor under the logistical regression link function. The CAPM predictor now correctly predicts the direction of future 1-year excess returns around 80% of the time for the 845 days. Figure 4.1.9 clearly reveals the impressive predictive capabilities as the prediction accurately captures the change in

direction of future returns to be earned a year out from around 2014. However the model fails to correctly predict the positive returns that will be realized in a year from 2016 but does capture a small part of the last turn of positive returns as seen by the small dots of purple towards the end of the test sample period.

Figure 4.1.9 Absa logistic regression 1-year return predictions with CAPM predictor variable. . The CAPM estimate of return is represented by the black line in the figure. The multi-coloured points represent the future 1-year returns, and the dotted line represents the class decision boundary estimated by the logistic regression model. .



4.1.3 Multiple predictor variables

The results for using multiple predictors as opposed to a single predictor in the statistical learning methods for predicting returns on Absa stock price are now presented.

4.1.3.1 Variable selection

The variable selection is performed very loosely in the move from a univariate predictor to multiple predictors in the statistical learning methods. The variable selection is performed in two different manners. The variable selection performed where the AIC is used as the selection criterion is done assessing the predictors in the logistic regression framework. The adjusted r-squared, BIC and Cp measures are constructed under the regular regression setting using the quantitative returns as the response variable.

The variable selection techniques are run using the full set of predictors and the training set of 1500 days. Table 4.1.5 summarizes the results obtained under the subset selection techniques as well as the results obtained by using these predictors over the test sample set of 845 days. The reader is referred to the Appendix A for the statistical outputs from the variable selection techniques. The results clearly provide that models selected by adjusted r-squared contain an un-parsimonious number of predictor variables. Most importantly the predictions from these over-fitted models perform rather poorly as expected.

Table 4.1.5 Absa predictor variable subset selection. The highlighted cells display the predictor variables included in the optimal set under the variable selection criterion. The last three rows display the prediction accuracy for predictions made on the test sample set using the chosen set of predictor variables in the logistic regression function.

ABSA Variable Selection												
	1 Year				6 Months				1 Day			
	Adjr2	Cp	Bic	Aic	Adjr2	Cp	Bic	Aic	Adjr2	Cp	Bic	Aic
CAPM												
Discount_360D_BB												
Discount_LTVol_BB												
Discount_GARCH_BB												
ROE_360D_BB												
ROE_LTVol_BB												
ROE_GARCH_BB												
Discount_360D_CS												
Discount_LTVol_CS												
Discount_GARCH_CS												
ROE_360D_CS												
ROE_LTVol_CS												
ROE_GARCH_CS												
No. Predictor Variables	13	11	11	9	11	11	10	11	13	5	1	1
1Y Prediction Accuracy	0,63	0,64	0,64	0,58	0,59	0,59	0,54	0,58	0,63	0,46	0,31	0,31
6M Prediction Accuracy	0,38	0,55	0,55	0,50	0,42	0,42	0,37	0,32	0,38	0,58	0,53	0,53
1D Prediction Accuracy	0,54	0,52	0,52	0,53	0,50	0,50	0,50	0,52	0,54	0,53	0,51	0,51

4.1.3.2 Bivariate prediction performance

The 1-year, 6 months and 3 months excess returns are concentrated on the rest of the analysis for the following reasons:

- 3 months is the minimum time for updates to the financial statements of the firm, although quarterly financial statement releases are not the most reliable. Annual financial statements published by the firm are expected to provide more consistent estimates of the total debt per share. Thus, predictor variables are not responsive enough over short-term horizons.
- The shorter-term is more influenced by noise and subjectivity and the short-term returns are more stochastic as evidenced in Figure 4.1.3.
- Predictors created from the price/risk on long term debt are more suitable predictors over longer horizons.

The variable selection experiment reveals the difficulty of selecting the optimal combination of variables in the statistical learning space. The experiment also highlighted the folly of model validation for prediction on measures of goodness of fit to the training data. In the move to the bivariate setting the combinations of variables are evaluated by performing out of sample model diagnostics for all possible pairs of predictor variables – this is computationally expensive but worthwhile.

Table 4.1.6 Absa summary of top 1-year return class predictions including pairs of predictors. The abbreviated measures in the table include; TSS: test sample size; FPR: False positive rate; False negative rate; CB: Class Balance. IND in the link function is the stand-alone class prediction with a constant class decision boundary of zero.

Link	Predictor Variables		TSS	Accuracy	Precision	Sensitivity	Specificity	FPR	FNR	CB
KNN 99	CAPM_1YER	Discount_LTVol_CS	819	80%	99%	47%	100%	0%	20%	37%
Logit		CAPM_1YER	845	80%	99%	48%	100%	0%	20%	39%
Logit	Discount_LTVol_BB	ROE_GARCH_BB	684	77%	93%	8%	100%	0%	23%	25%
Logit	CAPM_1YER	ROE_GARCH_CS	845	75%	71%	62%	84%	10%	15%	39%
Logit	CAPM_1YER	ROE_360D_CS	845	75%	68%	66%	80%	12%	13%	39%
Logit	CAPM_1YER	ROE_GARCH_BB	769	75%	96%	25%	99%	0%	25%	33%
KNN 99	CAPM_1YER	Discount_GARCH_CS	814	73%	70%	46%	89%	7%	20%	37%
Logit	Discount_360D_BB	ROE_LTVol_BB	745	72%	77%	16%	98%	2%	26%	31%
KNN 99	CAPM_1YER	Discount_GARCH_BB	774	71%	60%	44%	85%	10%	19%	34%
KNN 99	CAPM_1YER	Discount_LTVol_BB	760	71%	92%	10%	100%	0%	29%	32%
Logit		Discount_LTVol_BB	760	69%	88%	6%	100%	0%	30%	32%
KNN 99		CAPM_1YER	845	68%	97%	18%	100%	0%	32%	39%
KNN 99		ROE_360D_BB	845	67%	67%	31%	90%	6%	27%	39%
IND		ROE_360D_BB	845	67%	54%	100%	46%	33%	0%	39%
IND		Discount_GARCH_CS	814	65%	100%	4%	100%	0%	35%	37%

Table 4.1.6 highlights some of the top performing pairs of predictors and single set predictors for the 1-year excess return on Absa stock price. The victory achieved by adding the credit spread variable to the CAPM for the 1-year prediction is not a parsimonious victory. This is since the addition of a second variable has not increased performance significantly as well as the loss of sample size for where the predictor from credit spread was not able to find a unique solution. The change in sample size impacts the data set upon which the functional relationship is estimated. The undesirability of loss in sample size and small margin of improvement justify why this is a false victory.

Table 4.1.7 Absa summary of top 6-month return class predictions including pairs of predictors. The abbreviated measures in the table include; TSS: test sample size; FPR: False positive rate; False negative rate; CB: Class Balance. IND in the link function is the stand-alone class prediction with a constant class decision boundary of zero.

Link	Predictor Variables		TSS	Accuracy	Precision	Sensitivity	Specificity	FPR	FNR	CB
KNN 99	Discount_360D_BB	Discount_360D_CS	745	71%	94%	26%	99%	1%	29%	38%
IND		ROE_360D_BB	845	70%	60%	98%	48%	29%	1%	45%
Logit	Discount_GARCH_BB	Discount_GARCH_CS	773	68%	69%	36%	89%	7%	25%	40%
Logit	Discount_LTVol_BB	ROE_GARCH_BB	684	68%	71%	2%	100%	0%	32%	33%
KNN 99		Discount_360D_BB	745	67%	75%	22%	95%	3%	30%	38%
IND		Discount_360D_BB	745	67%	54%	97%	49%	32%	1%	38%
Logit	Discount_LTVol_CS	ROE_LTVol_CS	819	63%	66%	28%	89%	6%	31%	43%
Logit	ROE_GARCH_BB	Discount_LTVol_CS	743	63%	50%	2%	99%	1%	37%	37%
KNN 99	Discount_360D_BB	ROE_360D_CS	745	62%	52%	5%	97%	2%	37%	38%
Logit		Discount_LTVol_BB	760	62%	71%	2%	100%	0%	38%	39%
KNN 99	CAPM_6MER	Discount_360D_BB	745	62%	50%	13%	92%	5%	33%	38%
KNN 99	Discount_360D_BB	ROE_LTVol_CS	745	62%	46%	5%	97%	2%	37%	38%
Logit		Discount_LTVol_CS	819	57%	50%	1%	99%	1%	42%	43%
KNN 99		CAPM_6MER	845	53%	46%	28%	73%	15%	32%	45%

Once again the pole position found in Table 4.1.7 is not truly accurate on its own. This is because the pair of predictors makes 100 less predictions owing to missing data. Surprisingly the constant boundary around zero with *ROE_360D_BB* is the best predictor for the 6 months forward excess log returns on Absa over the test sample period. The performance of predictor variables from CAPM are less dazzling over the 6 month return horizon. This is particularly interesting when considering the class balance has not shifted that drastically in the sets of returns to impact performance. The indicators good performance with *ROE_360D_BB* variable indicates that variable can be used as good predictor of future excess return, specifying the link function remains a challenge in this area for now.

4.1.3.3 Predictor variable efficiency

In this subsection the concept of debt variable efficiency is introduced. Here the analysis is aimed at providing an informal evaluation of whether information from market traded debt on Absa was used efficiently or beneficially in the prediction of returns on Absa stock. Put more plainly, did converting the observed credit spreads or default probabilities into predictors through the structural models aid the prediction performance? Or would it have been better to use the raw observed market variables in the prediction process?

The efficiency of the predictor variables created from CAPM is also considered by evaluating the use the inputs of the model to directly predict stock prices as opposed to compressing the information through the relationships expressed by theories. The efficiency of predictor variables is examined by using logistic regression to make return predictions with the raw market data as the set of predictor variables.

Table 4.1.8 Absa predictor variable efficiency in 1-year return class predictions. Table displays test prediction performance for: CS: Credit Spreads; BB: Bloomberg PD; All: all raw market data; BBCS: combination of CS and PD raw market data; CAPMCS: combination of credit spread and market index data; CAPMBB: combination of CAPM & Bloomberg PD raw data.

	TSS	Accuracy	Precision	Sensitivity	Specificity	FPR	FNR	CB
CS	845	66%	90%	15%	99%	1%	33%	39%
BB	845	69%	95%	22%	99%	1%	31%	39%
CAPM	845	50%	36%	34%	61%	24%	26%	39%
ALL	845	51%	40%	54%	49%	31%	18%	39%
BBCS	845	64%	64%	19%	93%	4%	32%	39%
CAPMCS	845	52%	41%	54%	51%	30%	18%	39%
CAPMBB	845	55%	42%	41%	63%	22%	23%	39%

Table 4.1.9 Absa predictor variable efficiency in 6-month return class predictions Table displays test prediction performance for: CS: Credit Spreads; BB: Bloomberg PD; All: all raw market data; BBCS: combination of CS and PD raw market data; CAPMCS: combination of credit spread and market index data; CAPMBB: combination of CAPM & Bloomberg PD raw data.

	TSS	Accuracy	Precision	Sensitivity	Specificity	FPR	FNR	CB
CS	845	61%	70%	23%	92%	5%	34%	45%
BB	845	68%	87%	32%	96%	2%	30%	45%
CAPM	845	55%	50%	37%	70%	17%	28%	45%
ALL	845	55%	50%	54%	56%	24%	21%	45%
BBCS	845	58%	53%	39%	72%	15%	27%	45%
CAPMCS	845	63%	59%	51%	72%	16%	22%	45%
CAPMBB	845	68%	73%	46%	86%	8%	24%	45%

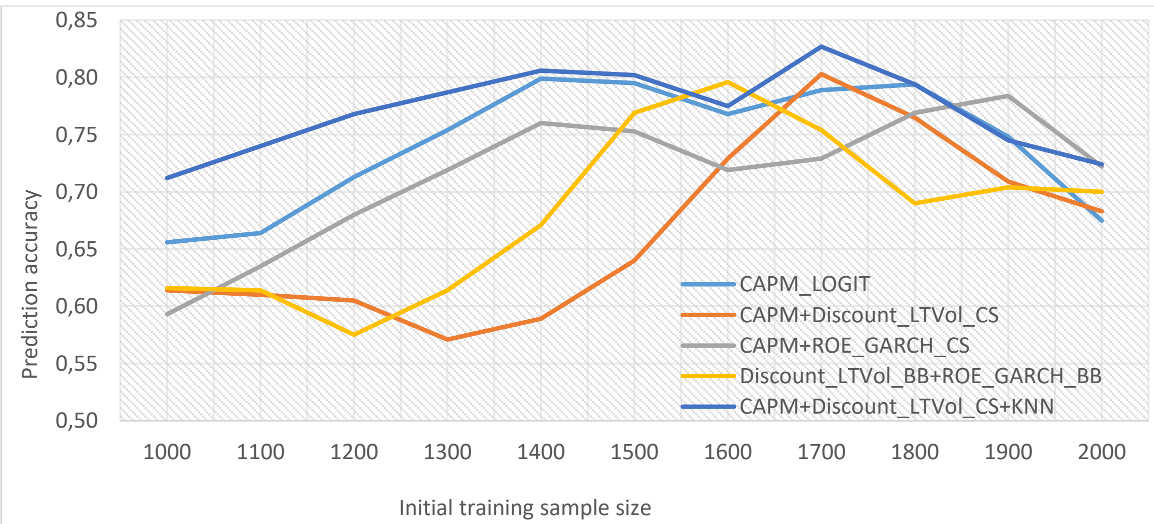
The models under this approach do not yield significant improvements in prediction performance, suggesting that the predictor creation process was beneficial or at least not detractive from the information garnered from these market observable variables. Admittedly the informal experimentation on predictor variable efficiency provides no insight with regards to what is the optimal use of information.

4.1.4 Varying training and test samples

The predictive performance has been measured using the starting point of predicting from observation 1500 in the sample set. However this choice is arbitrary and performance may vary depending on where it is assumed that the test predictions begin in the process. Figure 4.1.10 illustrates the performance for top predictors and pairs of predictors for the 1-year return on the Absa share price for varying starting points of test prediction. The varying of test and training splits allows for a stimulating view on how the models and predictors perform over larger test prediction range and how performance changes as the model feeds in more training data in the learning method.

We should reasonably expect that with more training data the statistical learning methods should produce more accurate predictions. It is interesting to note that for just about all learning methods and predictors there is a dip in performance when the initial training sample size is increased from 1600-1800 observations. This would suggest that the models are now being trained on test prediction observations that were previously well predicted and the remainder of test predictions are not predicted as well as those now included in the training sample.

Figure 4.1.10 Absa 1-year return prediction performance for varying training & test sample sizes. The graphic illustrates the test prediction accuracy for the top performing models and sets of predictor variables.



At the 1-year level the predictor from the CAPM appears to provide the best predictions using logistic regression as the link function in the statistical learning method as evidenced in Figure 4.1.10.

The case for the 6 month excess returns on the Absa stock price is a very different story though. Figure 4.1.11 illustrates the poor predictive capabilities of the CAPM created predictor variable for the 6-month excess returns. Furthermore the pairs of 'Discount' variables from the Merton model predictors appears to perform a reasonable job in predicting the class of future 6 month excess return on the stock.

Figure 4.1.11 Absa 6-month return predictions for varying training and test sample sizes. The graphic illustrates the test prediction accuracy for the top performing models and sets of predictor variables.

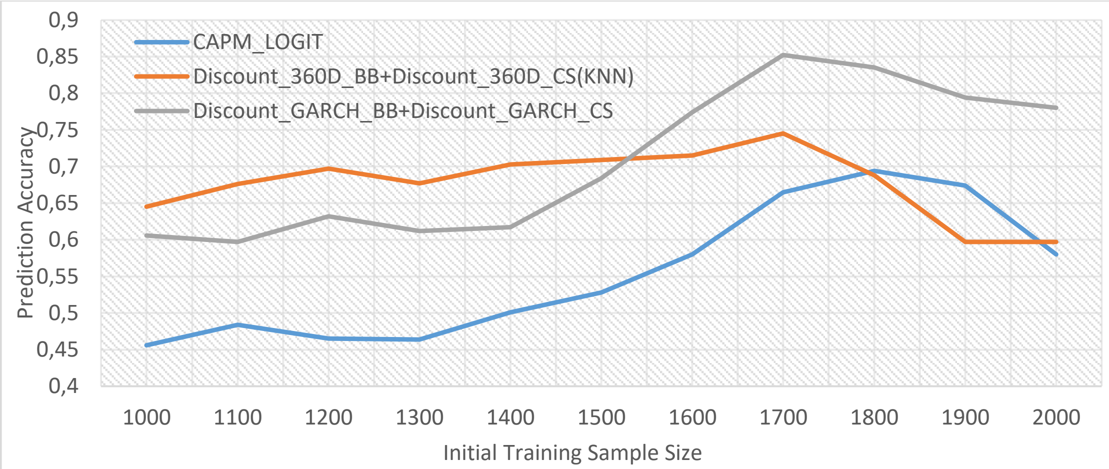


Figure 4.1.12 Absa 6-month return prediction KNN (K=99) with Discount_360D_BB & Discount_360D_CS as predictor variables. The graphic displays the time-series plot of the forward 6-month returns on the Absa stock price, the colours in the plot indicate the prediction performance for that return observation.

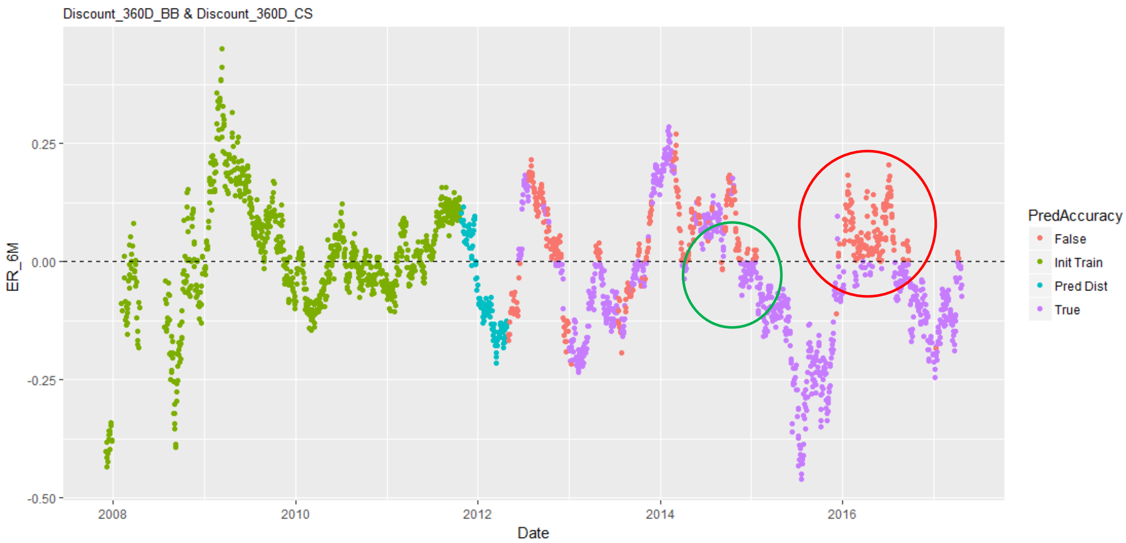
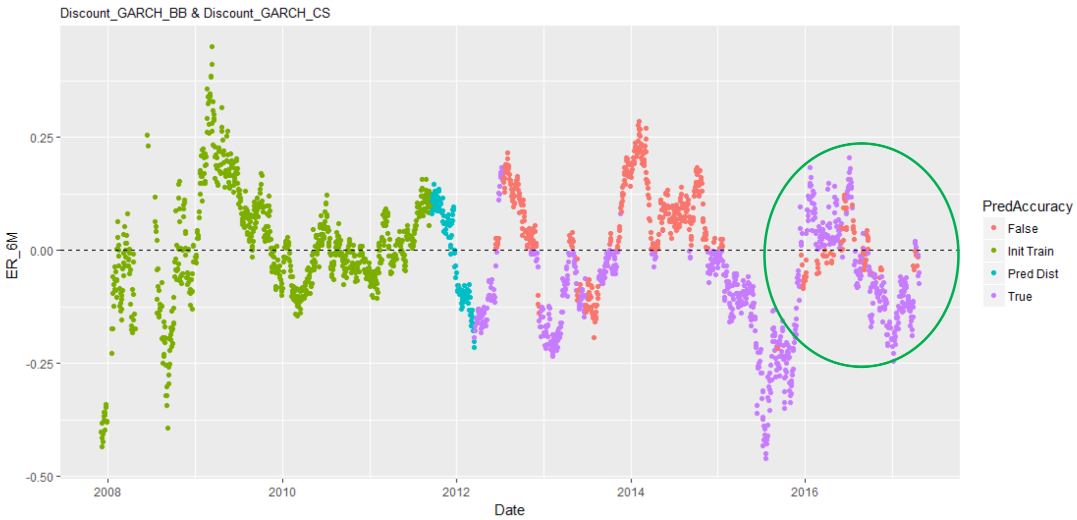


Figure 4.1.12 illustrates the predictions made using the KNN non-parametric approach for the link function with 99 neighbours and the Discount_360D_BB and Discount_360D_CS pair of predictor variables. The pair of variables fails to predict the positive returns that will be earned in 6 months from 2016, and capture small elements of the positive return of 2014.

Figure 4.1.13 illustrates the predictions made using the *Discount_GARCH_BB* and *Discount_GARCH_CS* pair of predictor variables under the logistical regression model. The latter combination of debt predictor variables accurate predicts the change in 6 month future returns around 2016 however performs very poorly in predicting prior returns.

Figure 4.1.13 Absa 6-month return predictions from logistic regression with Discount_GARCH_BB & Discount_GARCH_CS as predictor variables. The graphic displays the time-series plot of the forward 6-month returns on the Absa stock price, the colours in the plot indicate the prediction performance for that return observation.



4.2 INVESTEC GROUP LTD. RESULTS

The results for Investec Group Ltd. INL hereafter is now presented in detail.

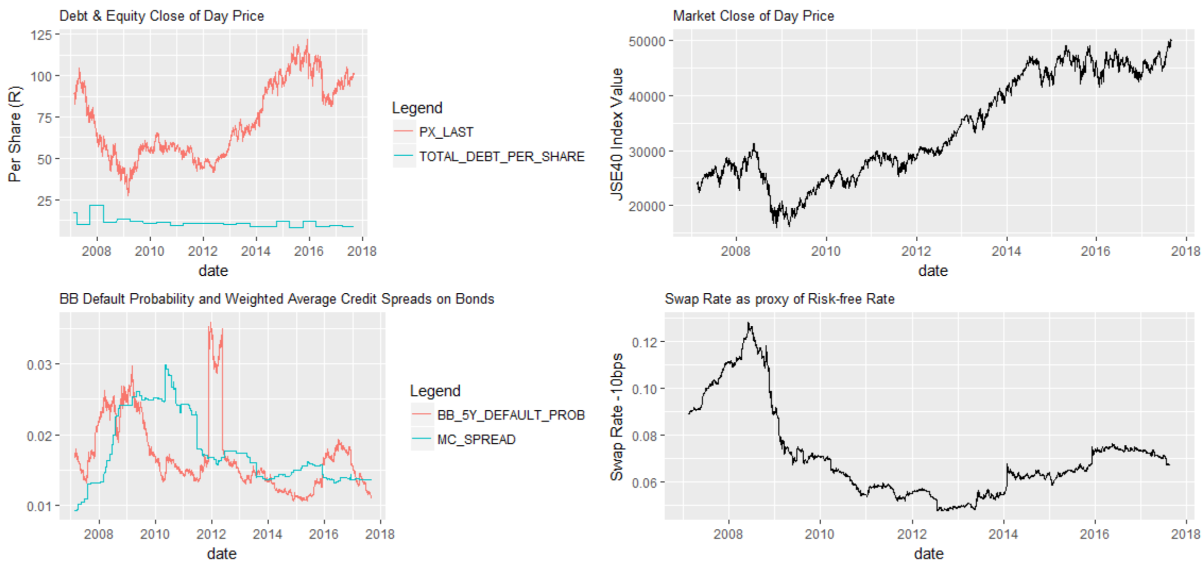
4.2.1 Predictor and target variables

A brief discussion regarding the input, predictor and target variable for INL is presented.

4.2.1.1 Initial input variables

Figure 4.2.1 illustrates that INL does not use a significant amount of leverage in the company's capital structure. The relationship between the share price and the market is not stable as evidenced by the fluctuations in estimates of alpha and beta found in the Appendix B**.** Furthermore, market rating of INL publicly traded debt appears to be well inversely related to share price performance. High spreads on corporate debt appears to be well co-integrated with decreasing stock prices and low spreads on debt is well correlated with increasing stock prices.

Figure 4.2.1 INL debt, equity and market input variables



4.2.1.2 Structural model predictor variables

The harsh large spikes in Figure 4.2.2 again signpost the greatest weakness of the current methodology for creating predictors of return under the Merton framework. The weakness of course is that the simultaneous equations for estimating the parameters of firms hidden asset value process cannot always find unique solutions. Table 4.2.1 confirms that the erratic volatility estimates from the $GARCH$ (1,1) forecast of 1-day ahead volatility produces large sets of sample points where the algorithm cannot find a unique solution to the simultaneous equations. Most encouragingly for most of the structural model predictors created, unique solutions were found for the full sample of 2748 days. The predictors created from credit spreads display the preservation of value property in the Merton framework and consistently lie juxtapose around zero.

Figure 4.2.2 INL structural model predictor variables

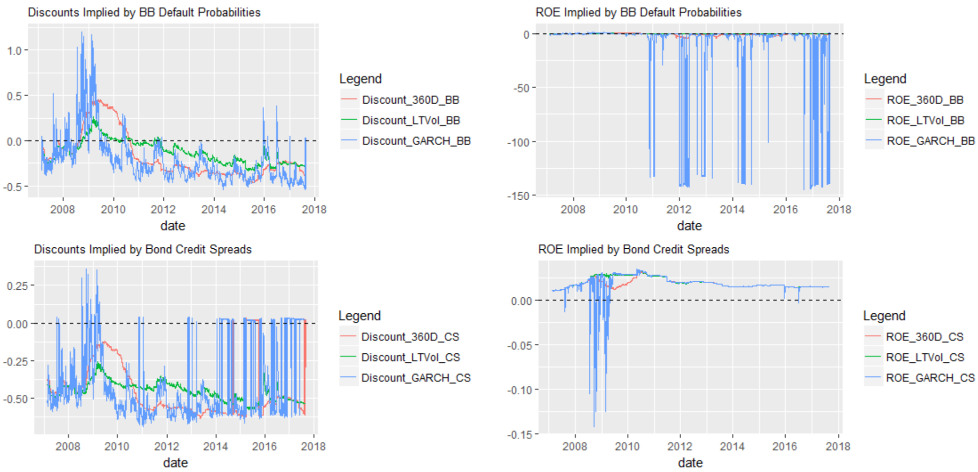


Table 4.2.1 INL convergence to solution in simultaneous equations for structural model return predictors. 1-convergence satisfied; 2-solution is uncertain; 3-no better solution found than starting point; 7 Jacobian is unusable.

INL TermCD Check										
	1	2	3	4	5	6	7	8	9	10
TermCD Discount_360D_BB	2748	0	0	0	0	0	0	0	0	0
TermCD Discount_LTVol_BB	2748	0	0	0	0	0	0	0	0	0
TermCD Discount_GARCH_BB	2738	0	10	0	0	0	0	0	0	0
TermCD ROE_360D_BB	2748	0	0	0	0	0	0	0	0	0
TermCD ROE_LTVol_BB	2748	0	0	0	0	0	0	0	0	0
TermCD ROE_GARCH_BB	2519	179	50	0	0	0	0	0	0	0
TermCD Discount_360D_CS	2604	1	143	0	0	0	0	0	0	0
TermCD Discount_LTVol_CS	2748	0	0	0	0	0	0	0	0	0
TermCD Discount_GARCH_CS	2286	13	449	0	0	0	0	0	0	0
TermCD ROE_360D_CS	2748	0	0	0	0	0	0	0	0	0
TermCD ROE_LTVol_CS	2748	0	0	0	0	0	0	0	0	0
TermCD ROE_GARCH_CS	2748	0	0	0	0	0	0	0	0	0

4.2.1.3 Realized forward excess log returns

Figure 4.2.3 INL realized forward excess log returns 2008-2017

The forward looking nature of the returns found in Figure 4.2.3 once again dictate that these should be interpreted as the excess log return that will be realized in n days from the data at which the point is plotted. In other words the 1-year excess log return plotted at 2016 is the excess log return that will be realized in 1-year from that date.

4.2.2 Univariate prediction performance

The analysis of predictability of the return on the share price for INL begins once again with an examination of each predictors stand-along predictive performance of the direction of excess returns realized over the next n period. The initial training sample size is taken as 1500 in this case so there is a maximum test sample set of 1248 observations in the sliding prediction methodology.

4.2.2.1 Indicator

The results here are dreadful. The CAPM predictor of expected return only predicts the direction of excess returns correctly 43% of the time. Some of the structural model estimates are constrained above or below zero and thus more indicative of the class balance as opposed to predictive capabilities in this sense.

Table 4.2.2 INL accuracy of expected return estimates evaluated as stand-alone predictors of future equity return class. (Class predictions are made using constant boundary of zero on the return estimate)

	1 Year	6 Months	3 Months	1 Month	1 Week	1 Day
ROE_GARCH_CS	54%	58%	57%	54%	55%	50%
ROE_360D_CS	54%	57%	57%	54%	55%	50%
ROE_LTVol_CS	54%	57%	57%	54%	55%	50%
ROE_GARCH_BB	48%	42%	43%	47%	46%	50%
Discount_360D_BB	46%	43%	43%	46%	45%	50%
ROE_360D_BB	46%	43%	43%	46%	45%	50%
Discount_LTVol_CS	46%	43%	43%	46%	45%	50%
Discount_LTVol_BB	46%	43%	43%	46%	45%	50%
ROE_LTVol_BB	46%	43%	43%	46%	45%	50%
Discount_GARCH_BB	44%	41%	42%	45%	45%	50%
CAPM	**43%**	**44%**	**50%**	**47%**	**48%**	**48%**
Discount_360D_CS	39%	38%	41%	46%	44%	51%
Discount_GARCH_CS	34%	38%	42%	47%	45%	51%

4.2.2.2 K-Nearest Neighbours

The full set of prediction results under the KNN methodology can be found in Appendix B. Examining the 1-year return predictions in Table 4.2.3 it is evident that the performance of the CAPM predictor variables have decreased significantly. The over-flexibility of the KNN classifier boundary in this instance lends itself to poor prediction results.

Table 4.2.3 INL KNN (K= 5) test prediction accuracy for returns defined over various time horizons.

	1 Year	6 Months	3 Months	1 Month	1 Week	1 Day
ROE_360D_BB	65%	52%	49%	52%	52%	51%
ROE_360D_CS	64%	58%	55%	48%	53%	48%
ROE_LTVol_CS	54%	46%	55%	52%	49%	49%
ROE_GARCH_CS	54%	49%	51%	49%	46%	49%
Discount_360D_BB	53%	51%	50%	47%	50%	47%
Discount_360D_CS	42%	45%	47%	51%	49%	50%
ROE_GARCH_BB	40%	44%	49%	51%	49%	50%
Discount_GARCH_BB	36%	44%	48%	47%	50%	50%
Discount_GARCH_CS	36%	47%	50%	47%	53%	52%
ROE_LTVol_BB	33%	47%	53%	46%	52%	49%
CAPM	**29%**	**38%**	**44%**	**46%**	**51%**	**51%**
Discount_LTVol_BB	26%	43%	46%	45%	49%	47%
Discount_LTVol_CS	22%	42%	46%	47%	49%	50%

4.2.2.3 Logistical regression

The logistical regression approach shows some improvement for the CAPM predictor variables however still very poor overall performance. The full set of results for the logistic regression predictions can be found in Appendix B. At this point the predictability of the returns on the INL share price seems unlikely.

Table 4.2.4 INL logistic regression prediction performance. Test prediction accuracy for predictions made using logistic regression with various predictor variables.

	1 Year	6 Months	3 Months	1 Month	1 Week	1 Day
ROE_360D_BB	51%	42%	41%	45%	47%	50%
CAPM	**48%**	**40%**	**43%**	**46%**	**47%**	**50%**
ROE_GARCH_BB	41%	42%	42%	45%	51%	50%
Discount_LTVol_CS	36%	42%	42%	45%	43%	50%
Discount_GARCH_BB	30%	36%	34%	38%	47%	50%
Discount_360D_BB	30%	30%	39%	41%	48%	50%
Discount_GARCH_CS	30%	38%	42%	47%	41%	50%
Discount_LTVol_BB	25%	34%	39%	45%	43%	50%
ROE_LTVol_BB	24%	32%	38%	44%	43%	50%
ROE_360D_CS	24%	40%	44%	46%	50%	50%
Discount_360D_CS	20%	33%	35%	43%	48%	51%
ROE_LTVol_CS	19%	22%	33%	41%	46%	50%
ROE_GARCH_CS	18%	42%	43%	41%	48%	50%

4.2.3 Multiple predictor variables

Extension of the application of statistical learning methods to multiple predictor variables.

4.2.3.1 Variable selection

Table 4.2.5 once again demonstrates the poor predictive capabilities of un-parsimonious optimal groups of predictors chosen from measures of goodness of fit to the training sample data. There is no performance that is worth mentioning or elaborating on any further here. The full set of statistical results for variable section techniques is found in Appendix B.

Table 4.2.5 INL predictor variable subset selection. The highlighted cells display the predictor variables included in the optimal set under the variable selection criterion. The last three rows display the prediction accuracy for predictions made on the test sample set using the chosen set of predictor variables in the logistic regression function.

INL Variable Selection												
	1 Year				6 Months				1 Day			
	Adjr2	Cp	Bic	Aic	Adjr2	Cp	Bic	Aic	Adjr2	Cp	Bic	Aic
CAPM	■	■	■	■	■	■	■	■	■			
Discount_360D_BB	■	■	■	■	■	■	■	■				
Discount_LTVol_BB	■	■	■	■	■	■	■	■	■			
Discount_GARCH_BB	■	■	■	■				■	■			
ROE_360D_BB				■	■	■		■				
ROE_LTVol_BB				■	■	■	■	■	■			■
ROE_GARCH_BB				■	■	■						
Discount_360D_CS	■	■	■	■	■	■	■	■	■			
Discount_LTVol_CS	■	■	■	■	■	■	■	■		■	■	
Discount_GARCH_CS	■	■	■	■	■	■	■	■				
ROE_360D_CS	■	■	■	■	■	■	■	■				
ROE_LTVol_CS	■	■	■	■	■	■	■	■				
ROE_GARCH_CS	■	■	■	■	■	■	■					
No. Predictor Variables	10	10	10	13	12	12	10	11	5	1	1	1
1Y Prediction Accuracy	30%	30%	30%	47%	39%	39%	35%	43%	23%	36%	36%	24%
6M Prediction Accuracy	43%	43%	43%	58%	57%	57%	53%	54%	39%	42%	42%	32%
1D Prediction Accuracy	49%	49%	49%	50%	50%	50%	49%	51%	50%	50%	50%	50%

4.2.3.2 *Bivariate prediction performance*

Table 4.2.6 INL summary of top 1-year return class predictions including pairs of predictors. The abbreviated measures in the table include; TSS: test sample size; FPR: False positive rate; False negative rate; CB: Class Balance. IND in the link function is the stand-alone class prediction with a constant class decision boundary of zero.

Link	Predictor Variables		TSS	Accuracy	Precision	Sensitivity	Specificity	FPR	FNR	CB
LOGIT	ROE_360D_BB	Discount_360D_CS	1104	72%	75%	83%	56%	17%	11%	61%
LOGIT	Discount_360D_BB	ROE_360D_BB	1248	72%	70%	83%	58%	19%	9%	54%
KNN 99	ROE_360D_BB	Discount_360D_CS	1104	68%	76%	71%	64%	14%	18%	61%
KNN 5		ROE_360D_BB	1248	65%	64%	83%	43%	26%	9%	54%
KNN 5		ROE_360D_CS	1248	64%	68%	62%	66%	16%	21%	54%
KNN 99	ROE_360D_BB	Discount_GARCH_CS	786	63%	72%	73%	43%	19%	18%	66%
LOGIT	ROE_360D_BB	ROE_360D_CS	1248	56%	56%	87%	19%	37%	7%	54%
LOGIT		ROE_360D_BB	1248	51%	54%	68%	30%	32%	18%	54%
LOGIT		CAPM_1YER	1248	48%	61%	11%	92%	4%	49%	54%
IND		ROE_GARCH_CS	1248	54%	54%	100%	0%	46%	0%	54%
IND		ROE_360D_CS	1248	54%	54%	100%	0%	46%	0%	54%

The results for the prediction performance on the INL stock price returns have been rather dismal. The inclusion of pairs of predictor variables re-invigorates the investigation. The summary of the top results for predictors and pairs of predictors under the different link functions for 1-year return predictions are displayed in Table 4.2.6. The combination of predictors created from the Bloomberg probability of default are able to correctly predict the class of future 1-year returns around 70% of the time without loss of any days in the sample set. The predictors generated from CAPM perform extremely poorly in this instance which suggests that debt pricing plays an important role in returns on the INL stock price. In the three and six month returns sets it is evident that while predictive performance is poor, debt variables provide the best predictions of future return class.

4.2.3.3 Variable efficiency

The results in Table 4.2.7 do not suggest that the transformation of information in predictor creation was counterproductive. This segmented admittedly provides no indication of whether this was the best possible use of information from the observed market and financial statement variables.

Table 4.2.7 INL predictor variable efficiency in 1-year return class predictions. Table displays test prediction performance for: CS: Credit Spreads; BB: Bloomberg PD; All: all raw market data; BBCS: combination of CS and PD raw market data; CAPMCS: combination of credit spread and market index data; CAPMBB: combination of CAPM & Bloomberg PD raw data.

	TSS	Accuracy	Precision	Sensitivity	Specificity	FPR	FNR	CB
CS	1248	52%	100%	12%	100%	0%	48%	54%
BB	1248	26%	24%	16%	39%	28%	46%	54%
CAPM	1248	41%	41%	17%	70%	14%	45%	54%
ALL	1248	21%	20%	16%	27%	34%	46%	54%
BBCS	1248	28%	24%	15%	43%	26%	46%	54%
CAPMCS	1248	46%	51%	16%	82%	8%	46%	54%
CAPMBB	1248	19%	19%	15%	24%	35%	46%	54%

4.2.4 Varying training and test samples

The results for varying the training and test samples sizes in the prediction procedure yields some very interesting results as shown in Figure 4.2.4. The prediction performance of the CAPM variables and debt variables appear to be negatively correlated, suggesting a disconnection between the influence of the price of debt and equity on the firm's returns. The changing influence of predictability over the different periods is reflective of the learning capabilities of the statistical models but also suggests non-stationarity.

Figure 4.2.4 INL 1-year return prediction performance for varying training & test sample sizes. The graphic illustrates the test prediction accuracy for the top performing models and sets of predictor variables.

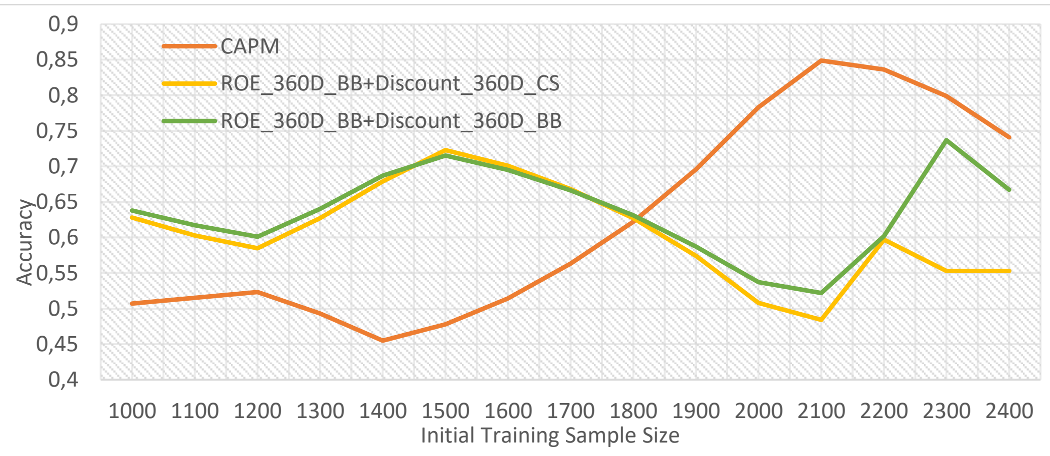


Figure 4.2.5, Figure 4.2.6 and Figure 4.2.7 further illustrate the nature of the predictions made by the top three performing predictor sets as well as provide clear explanations and insights for the shapes of the prediction accuracy performance curves found in Figure 4.2.4. Figure 4.2.5 illustrates the predictions and returns made using the CAPM predictor variable and logistic regression link function. The black line in the figure represents the CAPM estimates of return over the sample and the dotted line represents the classifier boundary estimated for test predictions over the sample. The graphic clearly illustrates the poor predictions made with smaller amounts of training data but also shows how the model correctly predicts the change in classes when using more training data.

Figure 4.2.5 INL logistic regression 1-year return predictions with CAPM predictor variable. The CAPM estimate of return is represented by the black line in the figure. The multi-coloured points represent the future 1-year returns, and the dotted line represents the class decision boundary estimated by the logistic regression model.

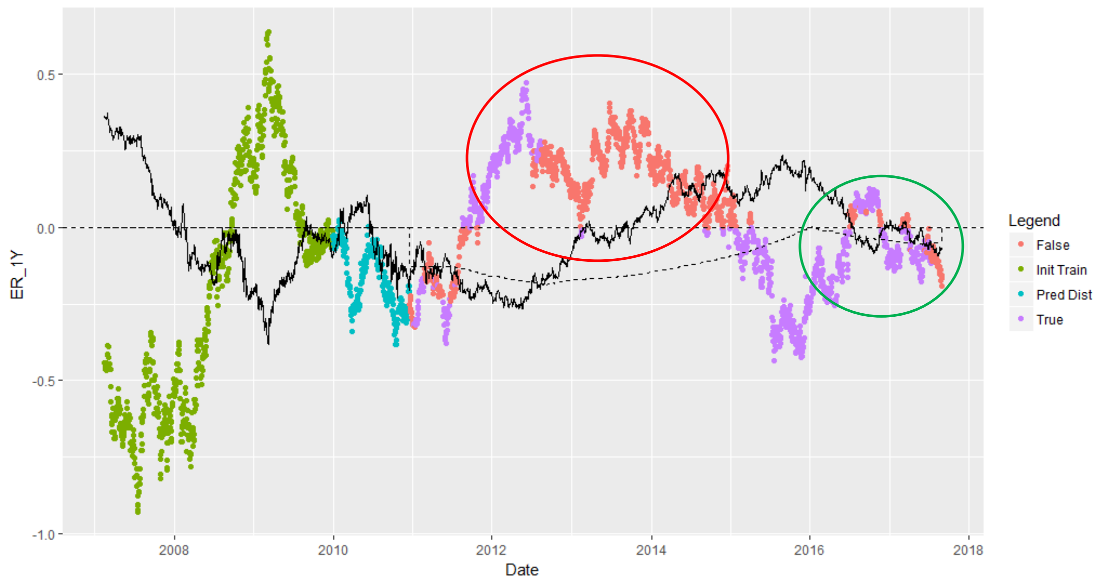


Figure 4.2.6 INL 1-year return predictions from logistic regression with ROE_360D_BB & Discount_360D_CS as predictor variables. The graphic displays the time-series plot of the forward 6-month returns on the INL stock price, the colours in the plot indicate the prediction performance for that return observation.

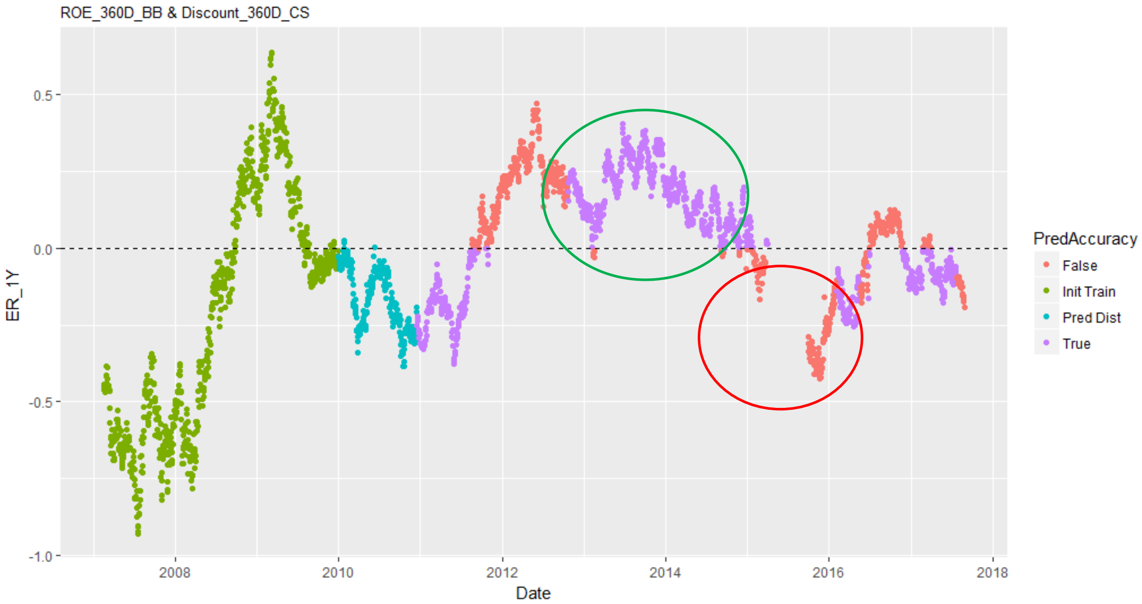
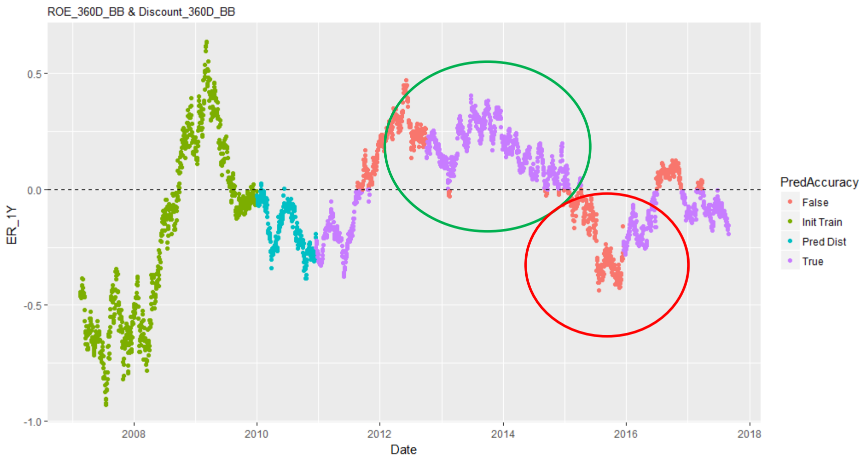


Figure 4.2.7 INL 1-year return predictions from logistic regression with ROE_360D_BB & Discount_360D_BB as predictor variables. The graphic displays the time-series plot of the forward 6-month returns on the INL stock price, the colours in the plot indicate the prediction performance for that return observation.



The predictions from the pairs of structural models of debt predictors correctly predict the class of 1-year forward excess returns from 2012-2014. However the impact of the learning lag is very noticeable in the prediction performance from these pairs. Figure 4.2.6 also illustrates the impact of the missing variables in the sample set from the 'Discount' predictor variable constructed from the credit spread on the bond.

4.3 Group Five Ltd results

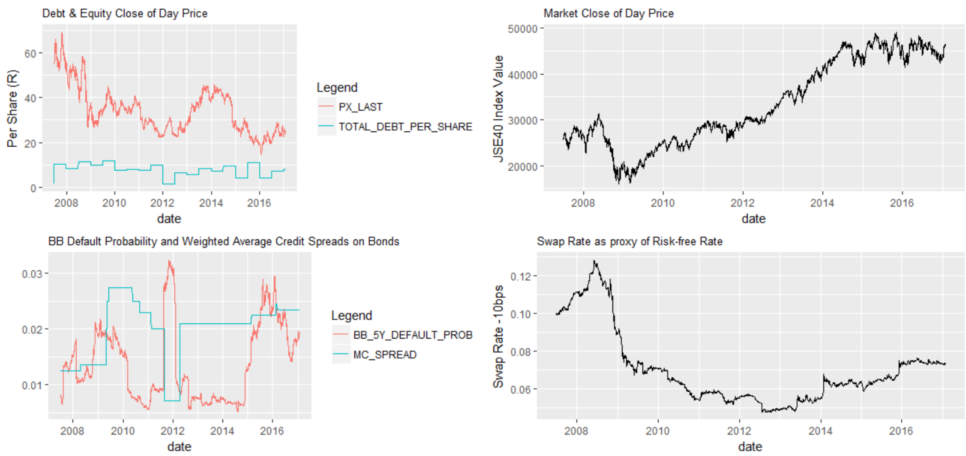
The results for predicting returns on the Group Five Ltd stock, hereafter GRF, is presented. GRF is a construction company and thus its employment of debt in its capital structure is perhaps slightly more simplistic than the financial services companies that have just been covered.

4.3.1 Predictor and target variables

4.3.1.1 Initial input variables

The GRF share price shown in the left quadrant of Figure 4.3.1, displays a downward trend in price over the sample period. This is noteworthy since the previous share price behaviours to be predicted were largely upward moving.

Figure 4.3.1 GRF debt, equity and market variable time series



4.3.1.2 Structural model predictor variables

The view of the estimates of return or predictors of return created from the structural model approach is once again hindered by the outliers for unsolved values in the case of predictors created from credit spreads and estimates created from models with $GARCH$ (1,1) inputs of 1-day volatility forecasts. The reader is referred to Appendix C for the estimates of equity volatility as well as estimates of alpha and beta over the sample period for GRF.

The estimates obtained from default probabilities provide both negative and positive estimates. Predictors obtained from credit spreads are constrained by the preservation of value in the solving of simultaneous equations for the unobservable parameters. The newton Raphson method was able to find unique estimates for the unobservable parameters under all but 3 of the variations employed in the creation of predictor variables as shown in Table 4.3.1. For most of the predictor variables there is fortunately no impact on the sample size due to missing data or observations where the methodology could not solve for unique parameters.

Figure 4.3.2 GRF structural models of debt predictor variables

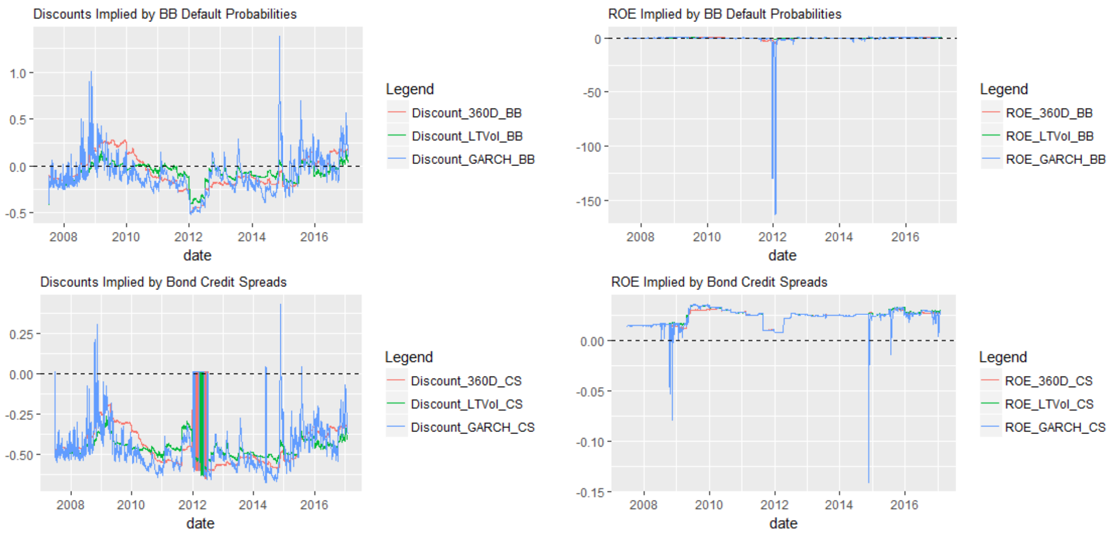
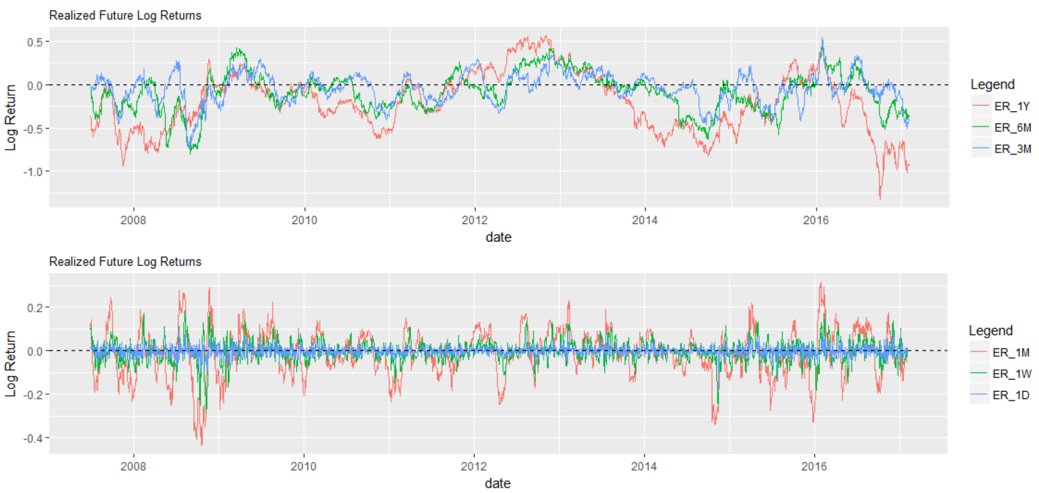


Table 4.3.1 GRF convergence to solution in simultaneous equations for structural model return predictors. 1-convergence satisfied; 2-solution is uncertain; 3-no better solution found than starting point; 7 Jacobian is unusable.

GRF TermCD Check										
	1	2	3	4	5	6	7	8	9	10
TermCD Discount_360D_BB	2398	0	0	0	0	0	0	0	0	0
TermCD Discount_LTVol_BB	2398	0	0	0	0	0	0	0	0	0
TermCD Discount_GARCH_BB	2398	0	0	0	0	0	0	0	0	0
TermCD ROE_360D_BB	2398	0	0	0	0	0	0	0	0	0
TermCD ROE_LTVol_BB	2398	0	0	0	0	0	0	0	0	0
TermCD ROE_GARCH_BB	2387	5	6	0	0	0	0	0	0	0
TermCD Discount_360D_CS	2295	0	103	0	0	0	0	0	0	0
TermCD Discount_LTVol_CS	2392	0	6	0	0	0	0	0	0	0
TermCD Discount_GARCH_CS	2307	1	90	0	0	0	0	0	0	0
TermCD ROE_360D_CS	2398	0	0	0	0	0	0	0	0	0
TermCD ROE_LTVol_CS	2398	0	0	0	0	0	0	0	0	0
TermCD ROE_GARCH_CS	2398	0	0	0	0	0	0	0	0	0

4.3.1.3 Realized forward excess log return

Figure 4.3.3 GRF realized forward excess log returns from 2008-2016



The short-term excess returns on the stock are very noisy as expected, this is clearly seen in Figure 4.3.3. Moreover the 1-year and 6-month returns series display both negative and positive returns, thus accurately predicting the class of return is a reasonable discriminator of performance.

4.3.2 Univariate prediction performance

4.3.2.1 Indicator

The analysis begins by examining the predictors' classification accuracy with a constant boundary around zero. The drastic difference in performance between the '*Discount*' and '*ROE*' predictors created from the credit spreads as evidenced in Table 4.3.2, is indicative of the class balance and the classic preservation of value constraint in the estimation.

Table 4.3.2 GRF summary of accuracy of expected return estimates evaluated as stand-alone predictors of future equity return class. (Class predictions are made using constant boundary of zero on the return estimate).

	1 Year	6 Months	3 Months	1 Months	1 Week	1 Day
Discount_LTVol_CS	84%	76%	68%	61%	55%	55%
Discount_360D_CS	83%	76%	69%	62%	55%	55%
Discount_GARCH_CS	82%	75%	67%	63%	55%	55%
ROE_LTVol_BB	80%	64%	52%	51%	52%	53%
Discount_LTVol_BB	80%	64%	52%	51%	52%	53%
Discount_GARCH_BB	73%	69%	58%	58%	55%	53%
ROE_GARCH_BB	73%	69%	59%	58%	55%	53%
ROE_360D_BB	71%	71%	55%	55%	55%	53%
Discount_360D_BB	71%	71%	55%	55%	55%	53%
CAPM	**46%**	**46%**	**51%**	**52%**	**50%**	**50%**
ROE_GARCH_CS	17%	26%	34%	40%	46%	46%
ROE_360D_CS	16%	25%	33%	40%	46%	46%
ROE_LTVol_CS	16%	25%	33%	40%	46%	46%

The class imbalance in the test prediction set artificially inflates the predictive performance at first glance. However what is once again evident is that estimates of return from the CAPM framework are very poor predictors of future return. Figure 4.3.4 illustrates the surprisingly good co-integration of the *ROE_360D_BB* created predictor and the 1-year and 6-month forward realized excess return on the GRF share. The '*Discount_LTVol_BB*' predictor variable also performs above expectations as can be seen in Figure 4.3.5 and further illustrates that estimates obtained from default probabilities are not constrained to a single class as is the case with those created from credit spreads.

Figure 4.3.4 GRF 1-year and 6-months excess log returns with ROE_360D_BB predictor from the period 2014- 2017.

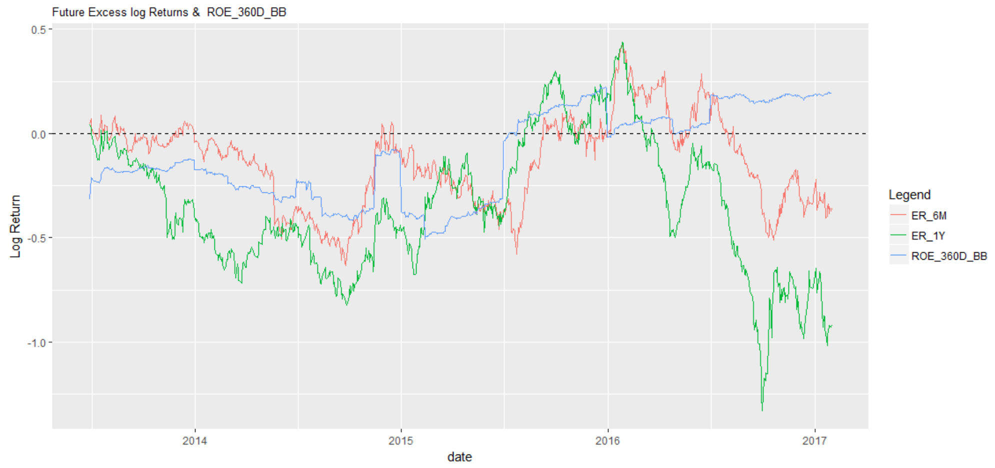
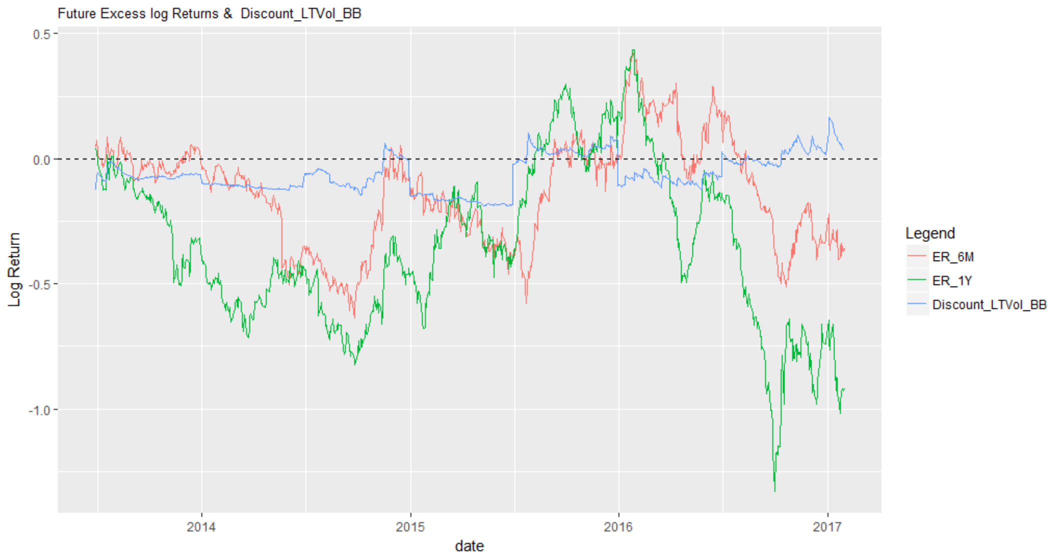


Figure 4.3.5 GRF 1-year and 6-months excess log returns with Discount_LTVol_BB predictor from the period 2014- 2017.



4.3.2.2 K-Nearest Neighbours

Table 4.3.3 GRF KNN (K= 5) test prediction accuracy for all predictor variables for returns defined over various time horizons.

	1 Year	6 Months	3 Months	1 Months	1 Week	1 Day
ROE_360D_BB	68%	62%	48%	50%	51%	48%
Discount_360D_CS	68%	58%	44%	56%	52%	50%
ROE_GARCH_BB	68%	60%	55%	51%	50%	49%
Discount_LTVol_CS	65%	60%	51%	49%	48%	51%
Discount_GARCH_CS	64%	51%	52%	50%	49%	51%
Discount_GARCH_BB	64%	53%	45%	49%	53%	52%
CAPM	**62%**	**61%**	**51%**	**48%**	**53%**	**50%**
Discount_LTVol_BB	60%	49%	44%	47%	50%	48%
Discount_360D_BB	56%	45%	49%	50%	51%	49%
ROE_LTVol_BB	52%	44%	46%	45%	48%	51%
ROE_360D_CS	46%	47%	56%	51%	53%	51%
ROE_LTVol_CS	46%	31%	45%	55%	50%	52%
ROE_GARCH_CS	43%	36%	47%	51%	50%	54%

The model is trained on a minimum of 1500 observations with a remaining potential maximum predictions for 898 observations. The K-Nearest Neighbours method is employed with a choice of $K = 5$ neighbours. This is an arbitrary choice of neighbours although the small number of neighbours is expected to produce a relatively flexible model. Table 4.3.3 confirms the counterintuitive point raised around poor predictive performance of highly flexible models and over fitting. KNN as a link function doesn't appear to be a great predictor, although CAPM has increased in performance.

4.3.2.3 Logistic regression

The results in the Table 4.3.4 show that the predictive capabilities for the CAPM are vastly improved with the aid of the logistical regression link function. The full set of results for logistic regression prediction in Appendix C reveals that impressively the model correctly predicts future stock returns in both directions using only the CAPM predictor variable.

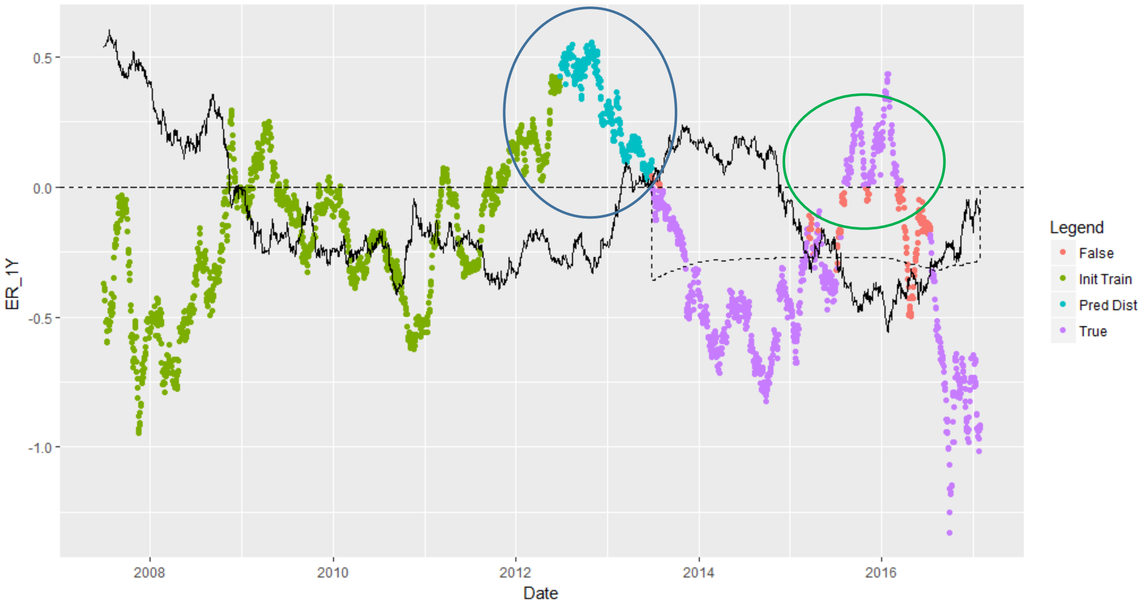
Table 4.3.4 GRF summary of logistic regression prediction performance. Test prediction accuracy for predictions made using logistic regression with various predictor variables.

	1 Year	6 Months	3 Months	1 Months	1 Week	1 Day
CAPM	**85%**	**82%**	**66%**	**56%**	**54%**	**54%**
Discount_LTVol_CS	84%	70%	66%	60%	55%	55%
ROE_360D_BB	84%	76%	68%	60%	55%	54%
Discount_360D_BB	84%	76%	68%	60%	55%	54%
ROE_LTVol_BB	84%	75%	68%	60%	55%	54%
ROE_360D_CS	84%	76%	69%	60%	55%	54%
ROE_LTVol_CS	84%	76%	68%	60%	55%	54%
Discount_LTVol_BB	**84%**	**74%**	**68%**	**60%**	**55%**	**54%**
Discount_GARCH_BB	84%	66%	65%	57%	52%	54%
Discount_360D_CS	83%	76%	63%	62%	55%	55%
ROE_GARCH_CS	82%	76%	69%	60%	54%	54%
Discount_GARCH_CS	80%	66%	63%	60%	54%	55%
ROE_GARCH_BB	76%	76%	66%	57%	53%	55%

The successful predictions and capturing of the different classes of future excess returns suggests that the logistical regression approach does not struggle to make predictions in the case of class imbalances, provided that the predictor variables contain sufficient discriminatory power. The contra positive is also true, where logistic regression struggles with a class imbalance in the training and test sets, the predictor variables might not provide sufficient powers of discrimination.

Figure 4.3.6 illustrates the impressive performance made by the logistical regression model 1-year return predictions with the CAPM predictor variable. The first impressive feature is that the model correctly predicts the first set of negative returns even where the last training point at this time is where returns are at an all-time high. The second impressive feature of the predictions made by this model is the accurate predictions of the change in return class evidenced around 2016 in the graphic.

Figure 4.3.6 GRF logistic regression 1-year return predictions with CAPM predictor variable. The CAPM estimate of return is represented by the black line in the figure. The multi-coloured points represent the future 1-year returns, and the dotted line represents the class decision boundary estimated by the logistic regression model.



4.3.3 Multiple predictor variables

4.3.3.1 Variable selection

Table 4.3.5 GRF predictor variable subset selection. The highlighted cells display the predictor variables included in the optimal set under the variable selection criterion. The last three rows display the prediction accuracy for predictions made on the test sample set using the chosen set of predictor variables in the logistic regression function.

GRFVariable Selection												
	1 Year				6 Months				1 Day			
	Adjr2	Cp	Bic	Aic	Adjr2	Cp	Bic	Aic	Adjr2	Cp	Bic	Aic
CAPM	■	■	■	■	■	■	■	■				
Discount_360D_BB	■	■	■	■	■	■	■	■	■			
Discount_LTVol_BB	■	■		■	■	■	■					
Discount_GARCH_BB	■	■	■	■								
ROE_360D_BB	■	■		■	■	■			■			
ROE_LTVol_BB				■	■				■			■
ROE_GARCH_BB				■	■			■				
Discount_360D_CS	■	■	■	■	■	■	■		■			■
Discount_LTVol_CS	■	■		■	■	■	■	■	■			
Discount_GARCH_CS	■	■	■	■	■	■	■	■				
ROE_360D_CS	■	■	■	■	■	■	■	■	■	■		
ROE_LTVol_CS	■	■	■	■	■	■	■	■	■			■
ROE_GARCH_CS				■	■	■	■	■	■	■	■	■
No. Predictor Variables	10	10	7	13	12	10	9	8	8	2	1	4
1Y Prediction Accuracy	65%	65%	76%	65%	66%	66%	69%	68%	62%	83%	82%	80%
6M Prediction Accuracy	58%	58%	61%	60%	59%	58%	59%	71%	54%	75%	76%	73%
1D Prediction Accuracy	54%	54%	54%	54%	54%	55%	54%	54%	53%	54%	54%	54%

Table 4.3.5 reiterates the difficulty of variable selection and predictive performance. The large numbers of predictors included by the models do not provide more useful predictions than those from univariate predictor variable approaches. Highlights the pitfalls of r-squared and model evaluation on in-sample model diagnostics. The reader is referred to Appendix C for the full results from variable selection techniques. Throughout the research experimentation, the adjusted r-squared criterion consistently selects the largest number of variables possible in the optimal subset.

4.3.3.2 Bivariate predictor performance

The summary of the top results for predictors and pairs of predictors under the different link functions for 1-year return predictions are displayed in Table 4.3.6. The results provide that predictors from CAPM perform quite well under the logistical regression approach yet can be improved further with the inclusion of created debt variables. The combination of the CAPM and the *ROE_360D_BB* predictor variables provides the most accurate predictions without the loss of sample size.

Table 4.3.6 GRF summary of top 1-year return class predictions including pairs of predictors. The abbreviated measures in the table include; TSS: test sample size; FPR: False positive rate; False negative rate; CB: Class Balance. IND in the link function is the stand-alone class prediction with a constant class decision boundary of zero.

LINK	Predictor Variates		TSS	Accuracy	Precision	Sensitivity	Specificity	FPR	FNR	CB
LOGIT	CAPM_1YER	Discount_360D_CS	795	94%	78%	91%	95%	4%	2%	18%
LOGIT	CAPM_1YER	ROE_360D_BB	898	93%	87%	66%	98%	2%	6%	16%
LOGIT	CAPM_1YER	ROE_GARCH_BB	887	90%	62%	93%	89%	9%	1%	16%
LOGIT	CAPM_1YER	ROE_360D_CS	898	86%	56%	69%	90%	9%	5%	16%
LOGIT	CAPM_1YER	ROE_LTVol_CS	898	86%	54%	66%	89%	9%	6%	16%
LOGIT		CAPM_1YER	898	85%	52%	96%	83%	14%	1%	16%
LOGIT	CAPM_1YER	Discount_GARCH_BB	898	85%	52%	96%	83%	15%	1%	16%
LOGIT	CAPM_1YER	ROE_GARCH_CS	898	84%	51%	90%	83%	14%	2%	16%
LOGIT	Discount_LTVol_BB	Discount_LTVol_CS	892	84%	NaN	0%	100%	0%	16%	16%
LOGIT		Discount_LTVol_CS	892	84%	NaN	0%	100%	0%	16%	16%
IND		Discount_LTVol_CS	892	84%	NaN	0%	100%	0%	16%	16%
LOGIT		Discount_360D_BB	898	84%	NaN	0%	100%	0%	16%	16%
IND		Discount_LTVol_BB	898	80%	41%	63%	83%	15%	6%	16%
KNN 5	CAPM_1YER	Discount_GARCH_BB	898	72%	29%	50%	76%	20%	8%	16%
KNN 5	Discount_GARCH_BB	ROE_GARCH_BB	887	69%	20%	30%	77%	19%	11%	16%
KNN 5		ROE_360D_BB	898	68%	2%	2%	81%	16%	16%	16%
KNN 5		Discount_360D_CS	795	68%	0%	0%	83%	14%	18%	18%
KNN 5	CAPM_1YER	Discount_GARCH_CS	807	68%	33%	88%	64%	30%	2%	17%

The poor performance of the KNN approach illustrates that the choice of neighbours has a large bearing on performance and that the choice of neighbours here may not have done justice to the KNN approach. Figure 4.3.7 and Figure 4.3.8 demonstrate the impact of the missing data points on the training and test predictions under the statistical learning methodology. Figure 4.3.7 clearly highlights the impact of the missing observations contained within the *'Discount_360D_BB'* predictor variable as well as the improved prediction accuracy in comparison to predictions in Figure 4.3.6. The methodology is constructed in such a way that any combination of variables will have the same number of observations in the initial training data set, resulting in varying test prediction sample sets. The change in training and test prediction sets is clearly visible when comparing Figure 4.3.8 and Figure

4.3.7. The impact of missing data observations changing the training and test sample splits should be kept in mind when comparing performance of these models.

Figure 4.3.7 GRF 1-year return predictions from logistic regression with CAPM & Discount_360D_CS as predictor variables. The graphic displays the time-series plot of the forward 1-year returns on the GRF stock price, the colours in the plot indicate the prediction performance for that return observation.

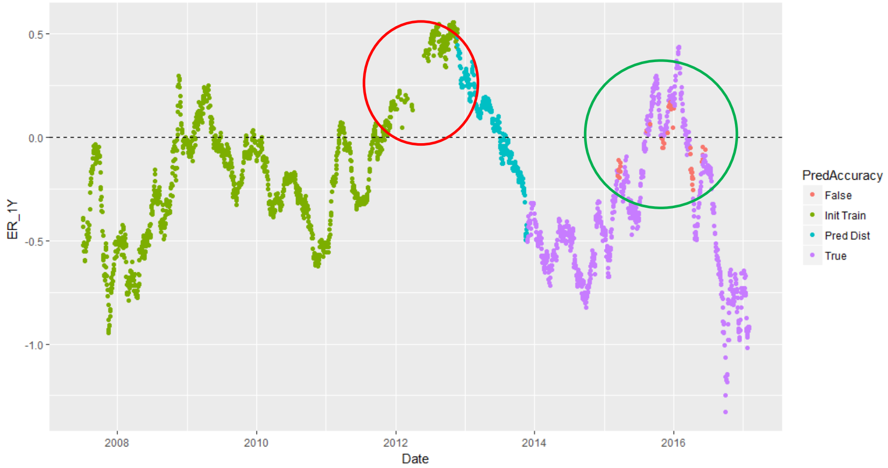
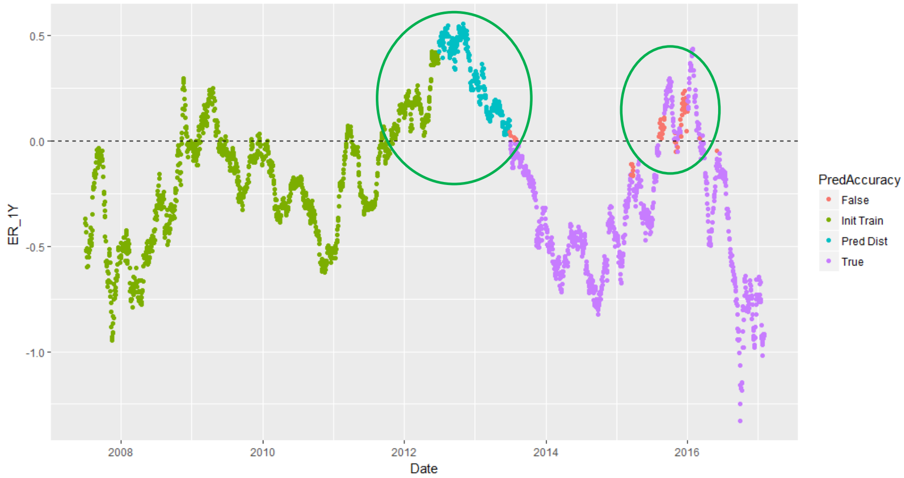


Figure 4.3.8 GRF 1-year return predictions from logistic regression with CAPM & ROE_360D_BB as predictor variables. The graphic displays the time-series plot of the forward 1-year returns on the GRF stock price, the colours in the plot indicate the prediction performance for that return observation.



The addition of the *ROE_360D_BB* predictor variables to the CAPM predictor variable yields significant improvements in prediction accuracy for 1-year returns without the loss of any sample size as evidenced in Figure 4.3.8. The 6-month return predictions are also aided by the incorporation of predictor pairs in the statistical learning model. Table 4.3.7 reaffirms that the addition of structural model predictors to the CAPM predictors vastly enhances the predictive accuracy in the 6-month future excess log returns. The combination of 'CAPM' and '*ROE_360D_CS*' as predictor variables yields prediction accuracy of 85% for 6-month return predictions over a test sample of 898 days. Impressively the model is able to provide accurate predictions for both classes of return despite the class imbalance in the 6-month returns sample. Even at the three month return horizon there appears to be some ability to forecast the class of return using the combination of CAPM and structural model predictor variables, as evidenced in Table 4.3.8.

Table 4.3.7 GRF summary of top 6--month return class predictions including pairs of predictors. The abbreviated measures in the table include; TSS: test sample size; FPR: False positive rate; False negative rate; CB: Class Balance. IND in the link function is the stand-alone class prediction with a constant class decision boundary of zero.

LINK	Predictor Variates		TSS	Accuracy	Precision	Sensitivity	Specificity	FPR	FNR	CB
LOGIT	CAPM_6MER	ROE_360D_CS	898	85%	71%	64%	91%	7%	9%	25%
LOGIT	CAPM_6MER	ROE_LTVol_CS	898	84%	68%	64%	90%	7%	9%	25%
LOGIT		CAPM_6MER	898	82%	61%	70%	86%	11%	7%	25%
LOGIT	CAPM_6MER	ROE_LTVol_BB	898	82%	61%	67%	86%	10%	8%	25%
IND		Discount_LTVol_CS	892	76%	NaN	0%	100%	0%	24%	24%
LOGIT		ROE_GARCH_BB	887	76%	NaN	0%	100%	0%	24%	24%
IND		Discount_360D_BB	898	71%	44%	76%	69%	23%	6%	25%
KNN 5	ROE_360D_BB	ROE_360D_CS	898	67%	36%	46%	73%	20%	13%	25%
KNN 5	ROE_360D_BB	ROE_LTVol_CS	898	66%	34%	41%	74%	20%	15%	25%
KNN 5		ROE_360D_BB	898	62%	24%	26%	74%	20%	18%	25%
KNN 5		CAPM_6MER	898	61%	16%	15%	76%	19%	21%	25%

Table 4.3.8 GRF summary of top 3--month return class predictions including pairs of predictors. The abbreviated measures in the table include; TSS: test sample size; FPR: False positive rate; False negative rate; CB: Class Balance. IND in the link function is the stand-alone class prediction with a constant class decision boundary of zero.

LINK	Predictor Variates		TSS	Accuracy	Precision	Sensitivity	Specificity	FPR	FNR	CB
LOGIT	CAPM_3MER	ROE_GARCH_BB	887	70%	53%	52%	78%	15%	15%	32%
LOGIT	Discount_LTVol_BB	ROE_360D_CS	898	69%	72%	6%	99%	1%	31%	33%
LOGIT	Discount_LTVol_BB	ROE_GARCH_CS	898	69%	82%	5%	100%	0%	31%	33%
LOGIT		ROE_360D_CS	898	69%	68%	7%	98%	1%	30%	33%
LOGIT		ROE_GARCH_CS	898	69%	75%	5%	99%	1%	31%	33%
IND		Discount_360D_CS	795	69%	NaN	0%	100%	0%	31%	31%
IND		Discount_LTVol_CS	892	68%	NaN	0%	100%	0%	32%	32%
LOGIT		Discount_360D_BB	898	68%	NaN	0%	100%	0%	33%	33%
IND		Discount_GARCH_CS	807	67%	0%	0%	98%	1%	32%	32%
KNN 5	CAPM_3MER	ROE_LTVol_CS	898	57%	37%	46%	63%	25%	18%	33%
KNN 5	ROE_GARCH_BB	ROE_LTVol_CS	887	56%	37%	54%	58%	29%	15%	32%
KNN 5		ROE_360D_CS	898	56%	39%	61%	54%	31%	13%	33%
KNN 5		ROE_GARCH_BB	887	55%	34%	43%	61%	27%	18%	32%

4.3.3.3 *Efficiency of predictor variable creation*

The results in Table 4.3.9 and Table 4.3.10 below quickly verify that creating predictor variables aided the predictive capabilities in the employment of statistical learning methods. The case for efficient use of information in predictor creation is well supported in the case of the prediction accuracy obtained for the 6-month returns using created predictors in comparison to raw variables.

Table 4.3.9 GRF predictor variable efficiency in 1-year return class predictions. Table displays test prediction performance for: CS: Credit Spreads; BB: Bloomberg PD; All: all raw market data; BBCS: combination of CS and PD raw market data; CAPMCS: combination of credit spread and market index data; CAPMBB: combination of CAPM & Bloomberg PD raw data.

	TSS	Accuracy	Precision	Sensitivity	Specificity	FPR	FNR	CB
CS	898	69%	25%	46%	73%	22%	9%	16%
BB	898	62%	29%	96%	55%	38%	1%	16%
CAPM	898	54%	26%	96%	46%	45%	1%	16%
ALL	898	71%	18%	22%	81%	16%	13%	16%
BBCS	898	74%	31%	46%	80%	17%	9%	16%
CAPMCS	898	72%	21%	27%	81%	16%	12%	16%
CAPMBB	898	77%	41%	95%	73%	23%	1%	16%

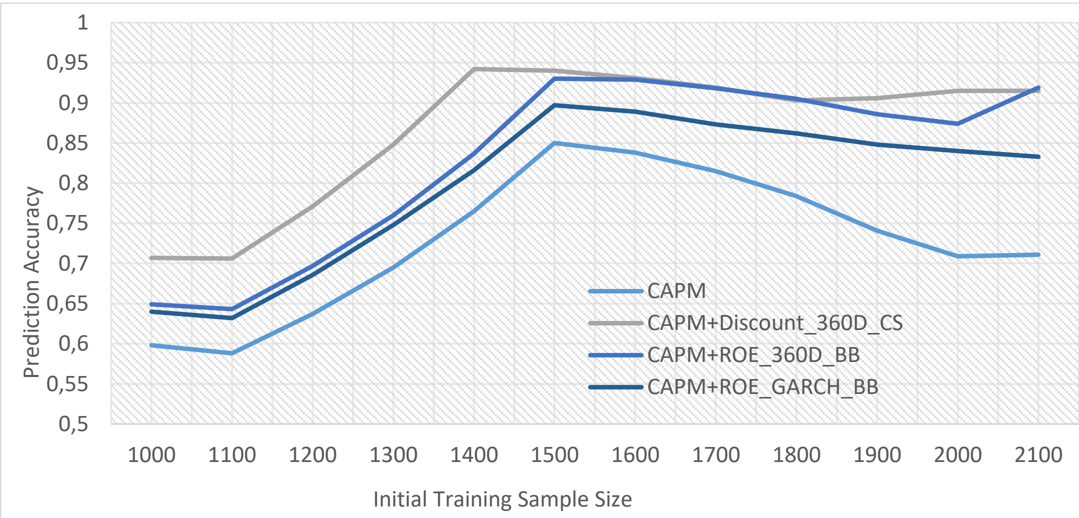
Table 4.3.10 GRF predictor variable efficiency in 6-month return class predictions. Table displays test prediction performance for: CS: Credit Spreads; BB: Bloomberg PD; All: all raw market data; BBCS: combination of CS and PD raw market data; CAPMCS: combination of credit spread and market index data; CAPMBB: combination of CAPM & Bloomberg PD raw data.

	TSS	Accuracy	Precision	Sensitivity	Specificity	FPR	FNR	CB
CS	898	53%	22%	37%	58%	32%	15%	25%
BB	898	55%	30%	61%	53%	35%	10%	25%
CAPM	898	63%	25%	27%	74%	20%	18%	25%
ALL	898	61%	5%	3%	80%	16%	24%	25%
BBCS	898	49%	16%	26%	57%	33%	18%	25%
CAPMCS	898	59%	4%	3%	77%	18%	24%	25%
CAPMBB	898	61%	5%	3%	80%	15%	24%	25%

4.3.4 Varying training and test samples

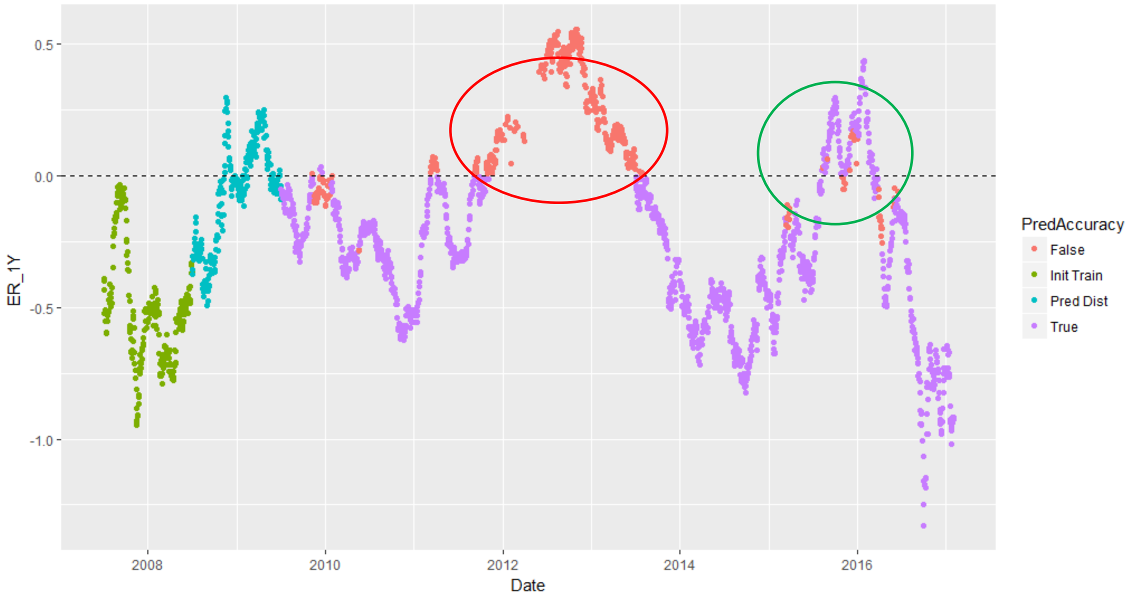
The models and predictor performance were considered for an arbitrary determination of 1500 observations in the initial training data sample. We see that CAPM and '*Discount_360D_CS*' as a predictor pair consistently provide the most accurate predictions although the impact of missing observations makes the outright comparison of model performance challenging.

Figure 4.3.9 GRF 1-year return prediction performance for varying training & test sample sizes. The graphic illustrates the test prediction accuracy for the top performing models and sets of predictor variables.



The shapes of the prediction accuracy for varying training and test data splits found in Figure 4.3.9, is most easily understood when viewed in conjunction with Figure 4.3.10. Figure 4.3.10 clearly illustrates the manner in which the model learns and improves predictions with more training data. The graphic clearly illustrates the models challenge to predict the first major series of positive 1-year returns as the training sample at this stage has very few labelled cases of positive returns to 'learn' the relationship from. The learning improvements are evident where the second shift from negative to positive 1-year returns is accurately predicted given the increase in information in the training data.

Figure 4.3.10 GRF 1-year return predictions from logistic regression with CAPM & Discount_360D_CS as predictor variables. The graphic displays the time-series plot of the forward 1-year returns on the GRF stock price, the colours in the plot indicate the prediction performance for that return observation.



The case for the 6-month return series also reveals an inflection point in prediction performance around 1500 observations in the training data set, evidenced in Figure 4.3.11. In this instance the combination of CAPM and 'ROE_360D_CS' consistently yields the most accurate predictions without and loss of observations. The fact that no observations were lost means that the predictions from CAPM and the combination of CAPM and 'ROE_360D_CS' are being measured over the same test and training data and may be directly compared. The improvement in predictive performance with the inclusion of structural model predictors in addition to CAPM is irrefutable for the 6-month returns of GRF.

Figure 4.3.11 GRF 6-months return prediction performance for varying training & test sample sizes. The graphic illustrates the test prediction accuracy for the top performing models and sets of predictor variables.

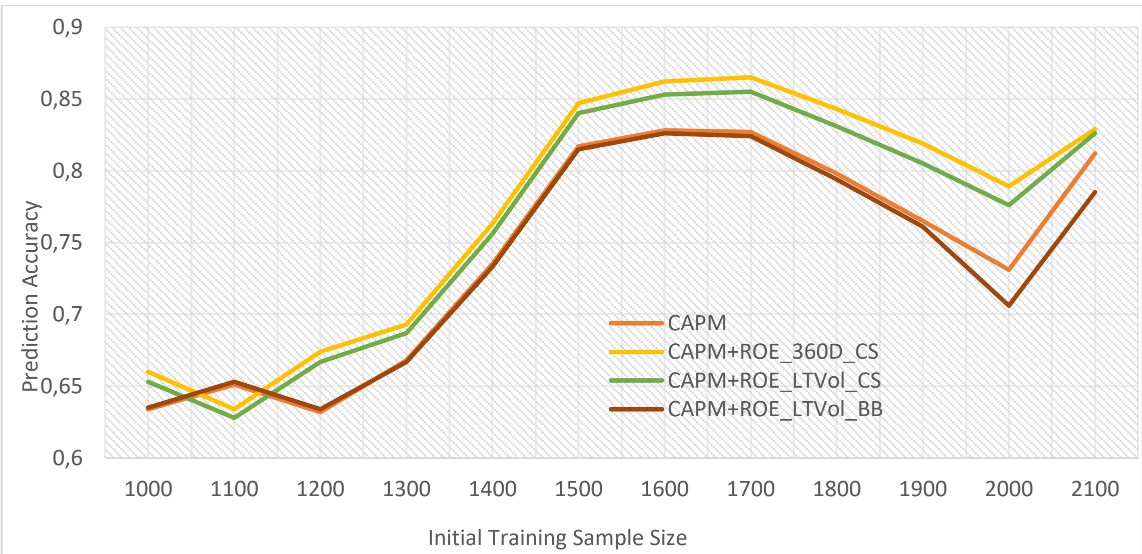


Figure 4.3.12 GRF 6-months return predictions from logistic regression with CAPM & ROE_360D_BB as predictor variables. The graphic displays the time-series plot of the forward 6-month returns on the GRF stock price, the colours in the plot indicate the prediction performance for that return observation.

The impressive nature of the predictions made for the 6-month returns are highlighted in Figure 4.3.12. The perfect prediction of the change in class of returns around 2011 going into 2012 in the figure is a real highlight for adding weight to the evidence that debt and capital structure can be used to predict future stock returns.

4.4 BIDVEST GROUP LTD. RESULTS

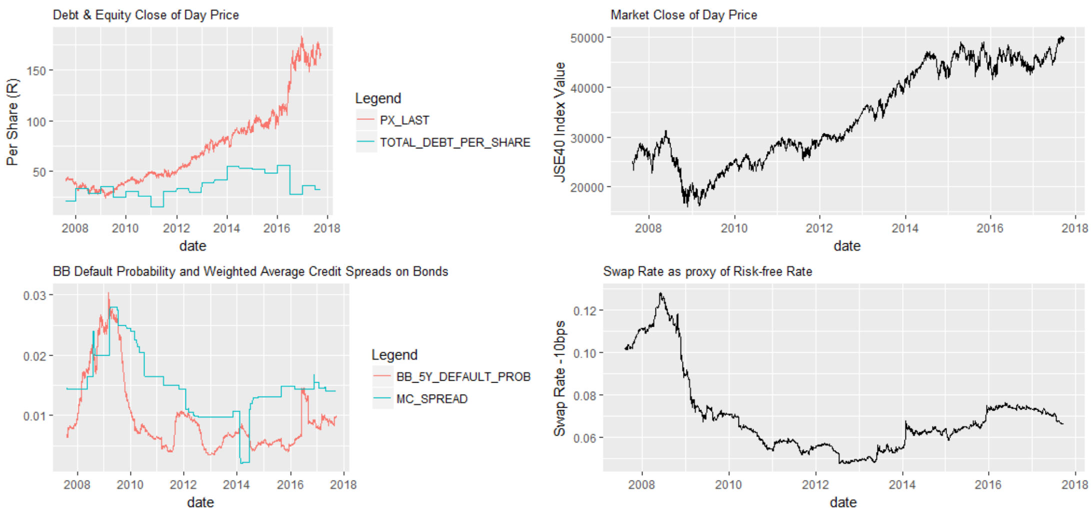
Modelling for Bidvest Co., referred to here after as BVC.

4.4.1 Predictor and target variables

4.4.1.1 Input variables

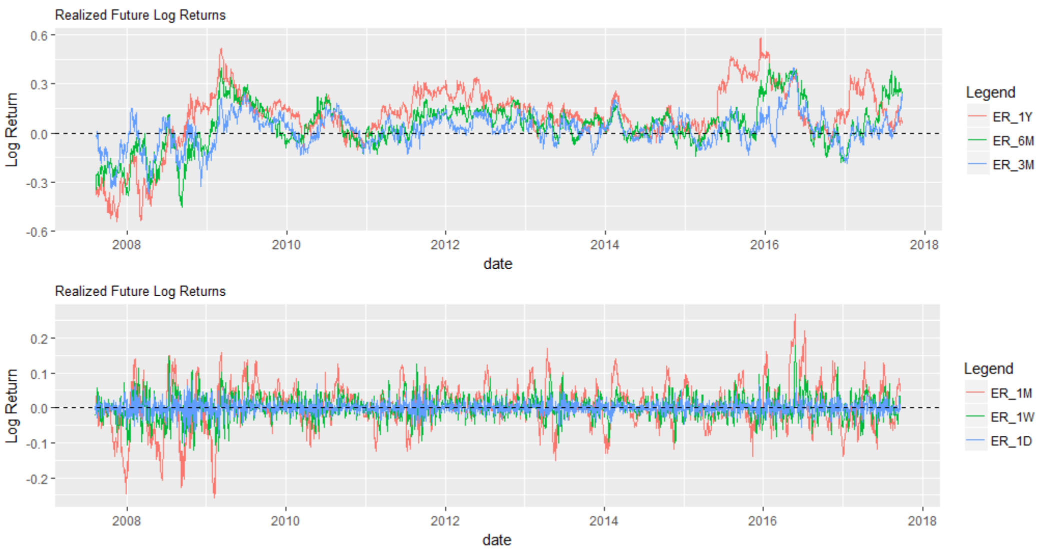
In Figure 4.4.1 illustrated below, the traded equity value and the financial statement debt per share of BVC are found in the top left graphic. The share price and market index appear to be well co-integrated although estimates of alpha and beta vary significantly over time as seen in Appendix D. The market capitalized (MC) spread also drops off around 2014 identifying a period where no data is available for publicly traded corporate bonds on BVC. Fortunately, Bloomberg estimates of five-year default probabilities are available for all days in the sample period.

Figure 4.4.1 Bidvest Co Market and Firm debt and equity time series



4.4.1.2 Realized forward excess log return

Figure 4.4.2 BVC realized forward excess log returns 2008-2017

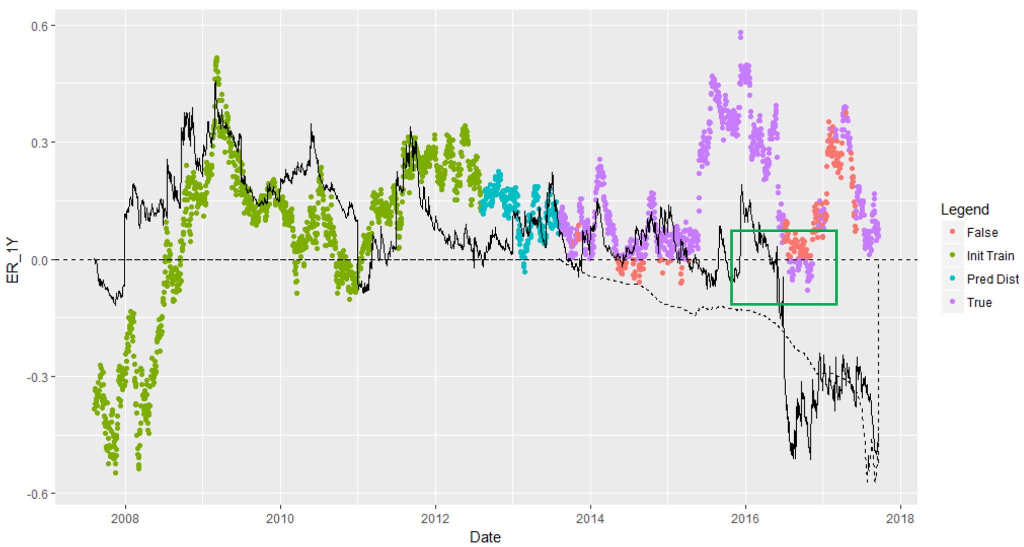


The excess log return series displayed in Figure 4.4.2 match the growth of the share price as seen in Figure 4.4.1. The 1-year, 6-month and 3-month return series then consist of a majority of positive return class observations. The evaluation of performance on class prediction does not lend itself well to discriminating between model performances when the sample consist of a majority class. For this reason the results and analysis are fast tracked for BVC.

4.4.2 Summary

In the case of BVC, more clarity on predictability and predictor variables influence on predictability is required and could be achieved by extending the number of classes or moving into the regression framework treating excess return as a continuous variable. The full set of outputs is available in Appendix D. Even with the class imbalance not providing the ideal basis for comparing predictive capabilities, the observations from the BVC analysis still provide some interesting discussion points. Figure 4.4.3 illustrates the class imbalance in the test prediction set for the 1-year returns as well as two impressive features about the '*ROE_LTVol_BB*' predictor variable. The first impressive feature is the co-integration of the future returns and the forecasted returns.

Figure 4.4.3 BVC logistic regression 1-year return predictions with ROE_LTVol_BB as predictor variable. The black line illustrates the ROE_LTVol_BB estimate of expected return over the sample. The dotted black-line plots the classification decision boundary determined by the logistic regression model. The graphic displays the time-series plot of the forward 1-year returns on the BVC stock price, the colours in the plot indicate the prediction performance for that return observation.



The second noteworthy feature of Figure 4.4.3 is both impressive and unimpressive at the same time. It is impressive that the rare negative events were successfully predicted by using the '*ROE_LTVol_BB*', however examining the results this small negative return was correctly predicted at the expense of incorrectly classifying far larger positive returns to be achieved. The magnitude of the losses are arguably relative small that the incorrect classification here is not a detriment to the model performance especially where the predictions are being used to execute a trading strategy. The disconnection between the *'ROE_LTVol_BB'* predictor variable and forward returns around 2016 correlates to a change in the total debt per share estimate used in the estimation procedure.

In the case of the 6-month returns for BVC the class imbalance in the test set is not as severe. Table 4.4.1 highlights the fact that there are a large number of unsuccessfully predicted negative returns in the test prediction set. In fact the top performing prediction results are all for models with no specificity, inability to predict the negative returns.

Table 4.4.1BVC summary of top 6--month return class predictions including pairs of predictors. The abbreviated measures in the table include; TSS: test sample size; FPR: False positive rate; False negative rate; CB: Class Balance. IND in the link function is the stand-alone class prediction with a constant class decision boundary of zero.

Link	Predictor Variables		TSS	Accuracy	Precision	Sensitivity	Specificity	FPR	FNR	CB
LOGIT	CAPM_6MER	Discount_GARCH_CS	976	72%	72%	100%	0%	28%	0%	72%
LOGIT	Discount_GARCH_CS	ROE_360D_CS	976	72%	72%	100%	0%	28%	0%	72%
LOGIT		Discount_GARCH_CS	976	72%	72%	99%	0%	28%	1%	72%
LOGIT	ROE_360D_BB	Discount_GARCH_CS	976	72%	72%	99%	0%	28%	1%	72%
LOGIT		ROE_GARCH_BB	828	71%	71%	100%	0%	29%	0%	71%
LOGIT		CAPM_6MER	1028	71%	71%	100%	0%	29%	0%	71%
IND		ROE_360D_CS	1028	71%	71%	100%	0%	29%	0%	71%
IND		ROE_LTVol_CS	1028	71%	71%	100%	0%	29%	0%	71%
IND		ROE_GARCH_CS	1028	71%	71%	100%	0%	29%	0%	71%
LOGIT		Discount_LTVol_BB	1028	66%	72%	85%	19%	24%	11%	71%
IND		CAPM_6MER	1028	64%	71%	85%	13%	25%	11%	71%
KNN 5	CAPM_6MER	ROE_GARCH_BB	828	63%	76%	70%	45%	16%	21%	71%
KNN 5	CAPM_6MER	Discount_360D_CS	1028	63%	74%	73%	37%	18%	19%	71%
KNN 5	ROE_GARCH_BB	Discount_360D_CS	828	62%	72%	77%	27%	21%	17%	71%
KNN 5		Discount_360D_CS	1028	62%	74%	72%	38%	18%	20%	71%
KNN 5		Discount_GARCH_CS	976	61%	72%	77%	22%	22%	17%	72%
KNN 5		Discount_GARCH_BB	1028	60%	71%	75%	24%	22%	18%	71%

Figure 4.4.4 BVC logistic regression 6-month return predictions with Discount_LTVol_BB as predictor variable. The black line illustrates the Discount_LTVol_BB estimate of expected return over the sample. The dotted black-line plots the classification decision boundary determined by the logistic regression model. The graphic displays the time-series plot of the forward 6-month returns on the BVC stock price, the colours in the plot indicate the prediction performance for that return observation.

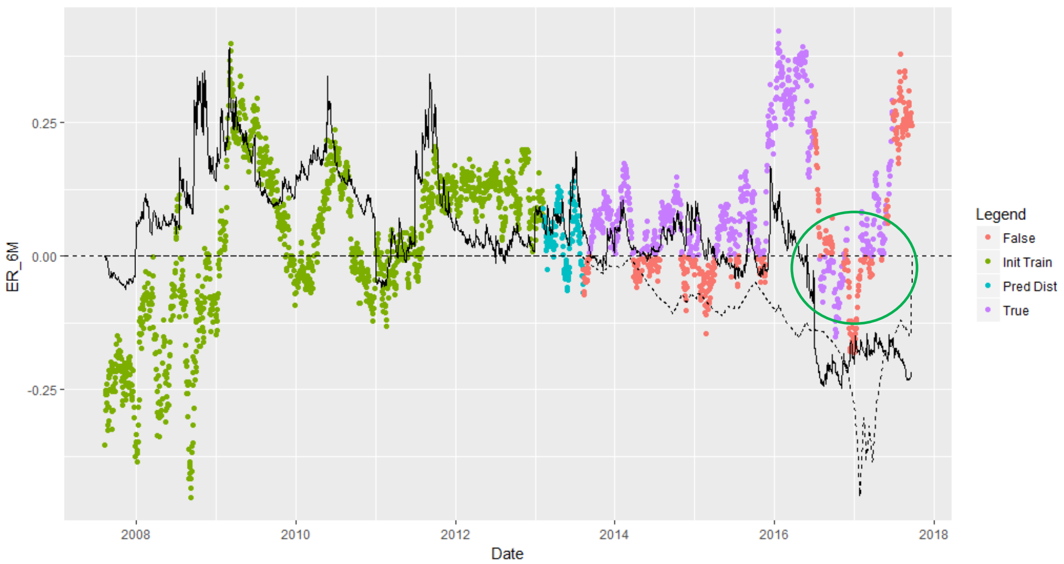


Figure 4.4.4 further illustrates that these are sufficiently large material losses worthwhile to accurately predict. Further it is not as if there is insufficient observations of both classes in the training data. The 6-month return predictions in Figure 4.4.4 provide that large negative returns were observed post 2008 although the market regime has arguably shifted post the 2008 credit crisis. The 'Discount' estimate of expected return created from the Bloomberg default probability does quite an amazing job at forecasting the actual returns earned by BVC. Barring once again the clear shift around 2016 in the 'Discount_LTVol_BB' predictor variable as seen in the graphic. The jump in the predictor estimate highlights the weakest point of the implementation of the structural approach, reliance on financial statement information.

4.5 Capitec results

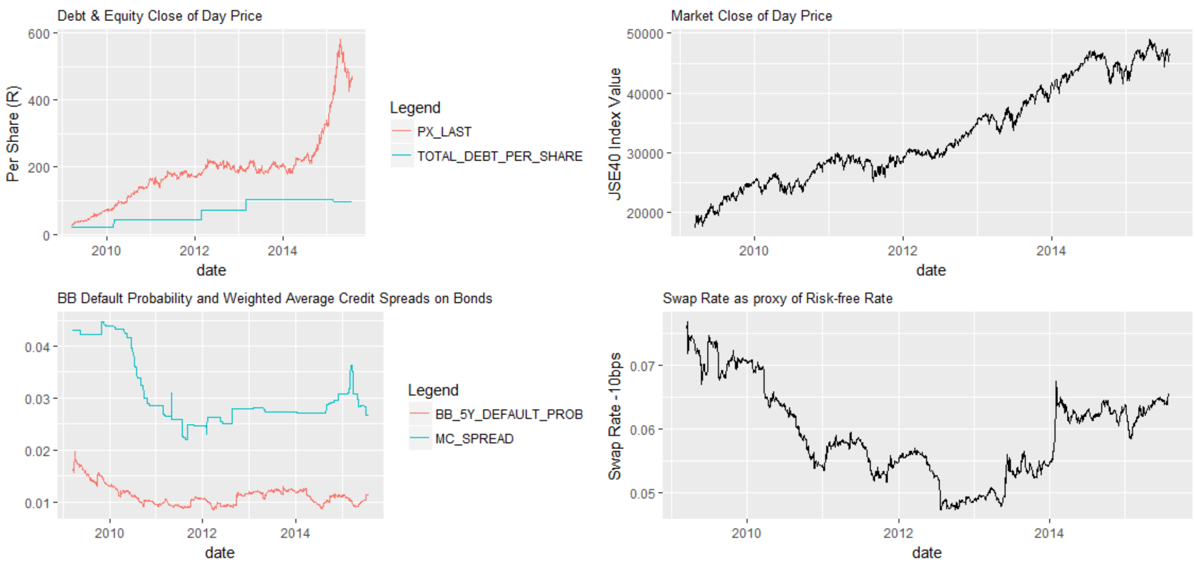
Modelling for Capitec bank is analysed in short detail.

4.5.1 Predictor and target variables

4.5.1.1 Input variables

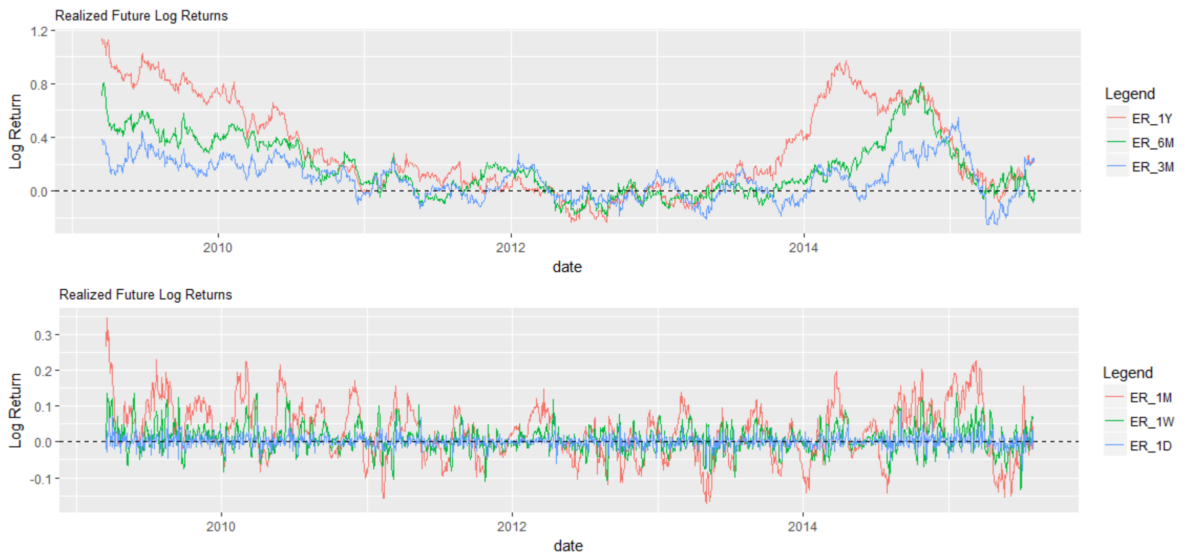
The total debt per share of Capitec appears to grow in proportion to the firm's equity price in the market. The Capitec share price also displays a very interesting spike towards the end of the sample period as seen in Figure 4.5.1.

Figure 4.5.1 Capitec firm debt, equity and market variables time series



4.5.1.2 Realized forward excess log return

Figure 4.5.2 Capitec realized forward excess log returns 2010-2015

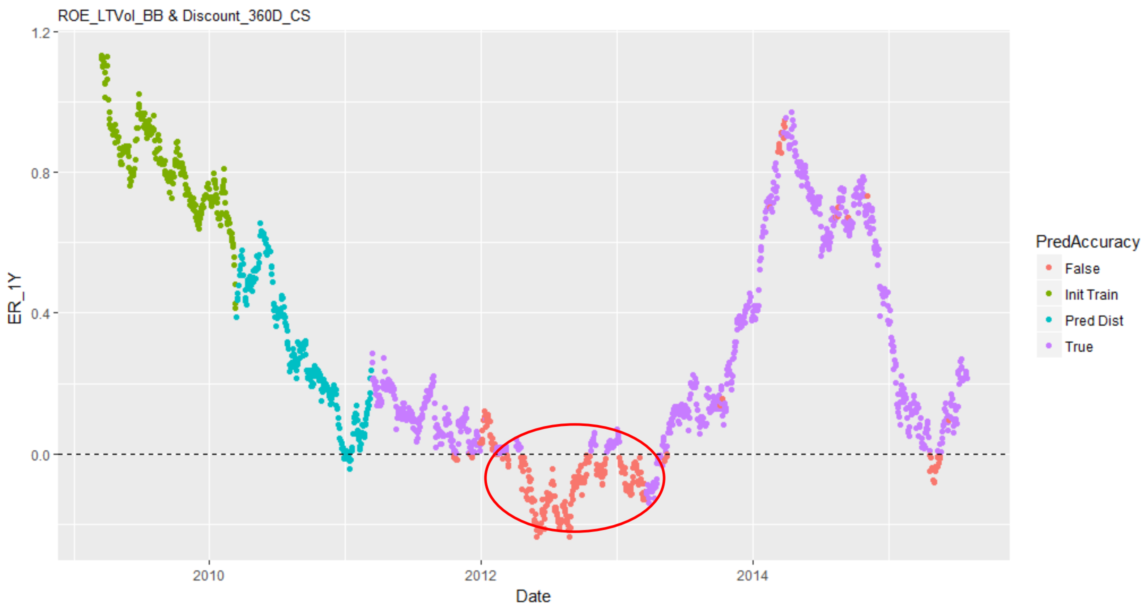


The return series for Capitec provides that test prediction samples will be dominated by a majority class and thus performance evaluation on the two class system does not provide adequate discriminatory information regarding predictor and model performance. The analysis of the prediction performance for the stock price returns for Capitec is fast tracked in a similar case to that of BVC. The full set of outputs and analysis for Capitec is found in Appendix E.

4.5.2 Summary

Figure 4.5.3 illustrates the class imbalance problem in the test prediction set. The case for Capitec is worse than for the imbalance within BVC, this is since there is only one negative return set which occurs mid-way through the test sample set. This highlights the weakness in the use of the statistical learning methods in stock price prediction – the models require sufficient training data to yield accurate predictions. Historical data for bond spreads in South Africa is not available in abundance as evidenced by the small sample size of firms and small sample of observations attainable for Capitec.

Figure 4.5.3 Capitec KNN (K=1) 1-year return predictions with ROE_LTVol_BB & Discount_360D_CS as predictor variables. The graphic displays the time-series plot of the forward 1-year returns on the BVC stock price, the colours in the plot indicate the prediction performance for that return observation.



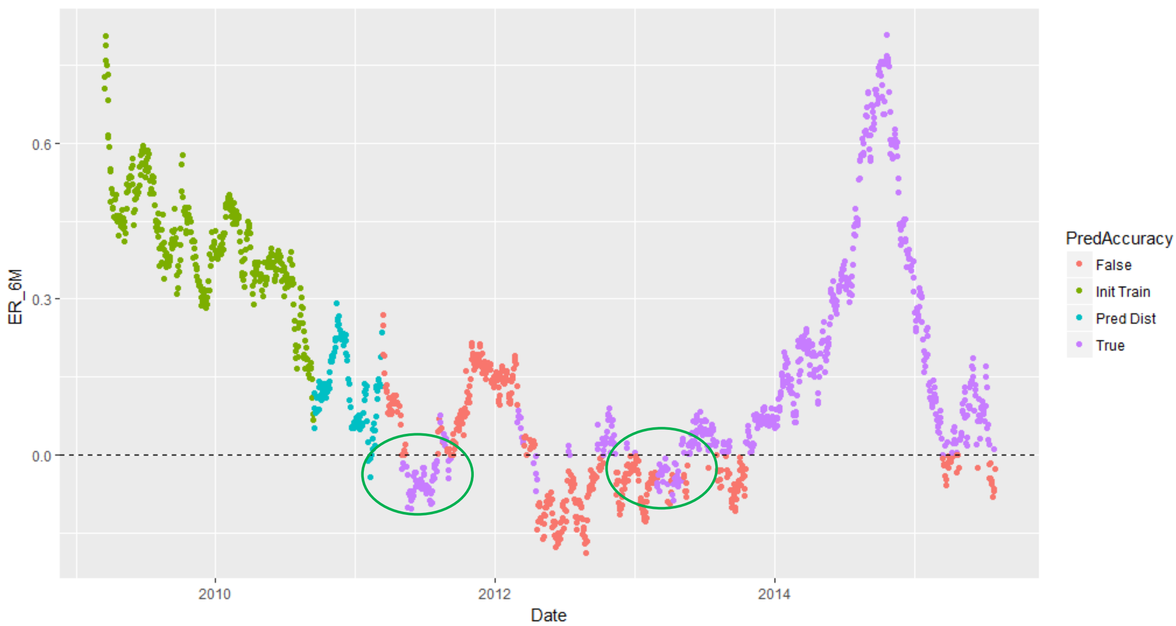
Even with the class imbalance not providing the ideal basis for comparing predictive capabilities, the observations from the Capitec analysis still provide some interesting discussion points. The CAPM historic returns are poor indicators of the direction of future excess returns. The combination of 'Discount' type predictor variables from the default probability and credit spreads is able to predict 6-month return classes for a relatively small test prediction sample of 589 days or observations.

Table 4.5.1 Capitec summary of top 6--month return class predictions including pairs of predictors. The abbreviated measures in the table include; TSS: test sample size; FPR: False positive rate; False negative rate; CB: Class Balance. IND in the link function is the stand-alone class prediction with a constant class decision boundary of zero.

Link	Predictor Variables		TSS	Accuracy	Precision	Sensitivity	Specificity	FPR	FNR	CB
LOGIT	Discount_LTVol_BB	Discount_360D_CS	589	85%	85%	100%	19%	15%	0%	82%
LOGIT	Discount_GARCH_BB	Discount_GARCH_CS	566	84%	85%	98%	12%	15%	1%	83%
LOGIT	Discount_360D_BB	Discount_GARCH_CS	566	83%	83%	100%	0%	17%	0%	83%
LOGIT		Discount_GARCH_CS	566	83%	83%	100%	0%	17%	0%	83%
LOGIT		ROE_LTVol_BB	582	83%	83%	100%	0%	17%	0%	83%
LOGIT		Discount_LTVol_CS	583	83%	83%	100%	0%	18%	0%	83%
KNN 3	ROE_LTVol_BB	Discount_LTVol_CS	582	82%	83%	99%	0%	17%	1%	83%
LOGIT		CAPM_6MER	594	81%	81%	100%	0%	19%	0%	81%
IND		ROE_360D_CS	594	81%	81%	100%	0%	19%	0%	81%
IND		ROE_LTVol_CS	594	81%	81%	100%	0%	19%	0%	81%
KNN 3		Discount_LTVol_CS	583	79%	83%	95%	7%	16%	4%	83%
KNN 3	Discount_360D_BB	ROE_LTVol_BB	582	79%	85%	92%	20%	14%	7%	83%
KNN 3		ROE_360D_BB	594	72%	82%	85%	18%	16%	12%	81%
KNN 3		Discount_GARCH_CS	566	71%	85%	79%	30%	12%	17%	83%

The summary of the 6-month return prediction performance is found in Table 4.5.1. The ability to predict novel cases is suggestive that debt structure and price of debt can be good indicators of future performance. Figure 4.5.4 displays the ability of the pair of debt predictors to successfully predict cross-overs in future class returns where there is sufficient training data. More insight can be gleamed by extending the number of classes in a meaningful way or shifting to the *regression* problem realm.

Figure 4.5.4 Capitec 6-month return predictions from logistic regression with Discount_LTVol_BB & Discount_360D_CS as predictor variables. The graphic displays the time-series plot of the forward 6-month returns on the Capitec stock price, the colours in the plot indicate the prediction performance for that return observation.



4.6 REVIEW OF FINDINGS AND OBSERVATIONS

The journey through the detailed analysis for each of the five firms was laborious and arduous. The analysis provided much food for thought for the arguments and sentiments being aired throughout the research. In this section the *take-aways* from the food for thought is served up.

Lessons from BVC and Capitec:

The imbalance in the test and training data sets for Capitec and BVC emphasised that statistical learning methods require sufficient training data in order to estimate the functional form of the relationship between the dependent variable and predictors. The class split should be adjusted where classes around zero do not provide for sufficient discrimination in predictive performance as well as proper model training.

The structural model predictors of return created from Bloomberg default probabilities for BVC displayed startling good forecasts of the future excess return on the stock. The BVC 6-month negative returns are materially large to warrant concern in lack of predictive accuracy. Although the co-integration of predicted and realized returns time series suggests that the structural approach may produce good predictors of future return, estimating the form of the relationship remains a challenge. The varying of test and training data sets revealed that the optimal set of feature variables to be used for stock return prediction is not constant.

Findings from INL:

The univariate predictor performance for INL produced very disparaging results. The move to pairs of predictors in the statistical learning methods yielded much improved results. Influence of debt on stock returns can be used to produce reasonable predictions, 70% prediction accuracy for 1248 out of sample test predictions. However, the model performance varied significantly with different test prediction and training sets, sending the reminder that there is no free lunch in statistics and no one method or set of predictors will dominate all data sets.

The INL share price changes over quite a large range yet the debt per share remains relatively consistent. The changing nature of the firm’s capital structure is connected to the *dis-connectivity* in predictor performance. In other words, changes in debt to equity structure are linked to when market or debt variables will be optimal predictors of expected return on the firm.

Insights from Absa:

Over the 1-year return horizon the CAPM predictor variable aided by the use of logistic regression provides the most accurate predictions of the future 1-year return class. The CAPM predictor variable performance is noticeably less dazzling over the 6-month returns horizon. At the 6-month level the debt predictor variables show very good capability for forecasting direction of excess returns.

The *ROE_360D_BB* variable is very interesting since it performs well with constant boundary around zero. This indicates the predictor is a good discriminator of future performance but specifying the link function remains a challenge. Further insights into predictive capabilities should be garnered from extending the analysis to the regression setting from the classification setting in lieu of such challenges.

Observations from GRF:

Predictors created from default probabilities produced reasonable forecasts of expected return over some time periods. The surprisingly good correlation between the forecasts and the actual return suggests that specifying the proper link function might yield improved accuracy in out-of-sample predictions. The performance over time for the GRF stock return predictions illustrates the nature of 'learning' in the statistical methods.

The use of CAPM predictor variable in conjunction with structural model variables showed resoundingly positive signs for the use of these predictors in stock prediction algorithms. Both the 1-year and 6-month returns showed highly accurate predictions over a decently length test sample set. Given sufficient training data the models were able to correctly predict both return classes with admirable accuracy.

Common themes:

Variable selection techniques once again emphasize the difference between in- and out-of-sample model diagnostics. In all cases the subset selection techniques yielded over-fit models and an un-parsimonious number of predictor variables in the optimal set of explanatory variables. Out of sample model diagnostics are key when evaluating predictive performance as all selected subsets produced very poor test prediction accuracy. The move to pairs of predictor variables was examined by performing test sample predictions for all possible pairs of predictors. This was extremely computationally expensive although worthwhile in terms of predictive accuracy that was shown to be possible.

The creation of the CAPM and structural model predictor variables were shown to be productive, as predictions using the raw market variable data and financial statement data yielded much poorer prediction accuracy. The CAPM estimates of expected return were shown to be very poor indicators of future return earned on the stock of the firm as stand-alone predictors. The CAPM estimate of expected return produced much improved prediction accuracy with the aid of statistical learning methods.

The deliberation and conclusions regarding the performance of the structural model predictors and the CAPM predictors in the greater scheme of things is postponed to the grand-finale in the final chapter. In the final chapter the research is reviewed as whole; the ability of the research to answer the posed research question and objectives is evaluated. The scope and limitations of the research is also presented along with recommendations for potential further research.

5 SUMMARY, CONCLUSIONS & RECOMMENDATIONS

5.1 SUMMARY

Economic models of equilibrium and asset pricing were shown to be more appropriately applied in the evaluation of past performance. The economic models of pricing in perfect competition offers little insight into forecasting future stock returns since in efficient markets the distinction between estimates of past performance and forecasts of future return are not meaningful.

In the literature review it was critically shown that estimates of future asset returns are essential to correctly solving the portfolio selection problem when operating under uncertainty. Portfolio optimization should be performed under future risk and return parameters not historical estimates. This implies that historical mean-covariance construction for portfolio selection problems are useful for evaluating past performance not selecting portfolio that will be efficient.

Estimating future asset returns for individual firms is no small feat. The discounted cash flow approach to value is one way of estimating future performance of the company's stock although the approach is hindered by the number of subjective estimates required as well as lack of market observable information from which to imply value and expected return. The CAPM can be used to predict stock price returns although requires prediction or forecasts of beta and the market, alternatively an employment of statistical learning methods.

Vast amount of credible economic theory and sound logic provided that capital structure plays an important role in future earnings of a company at the firm level. Reflexivity in stock prices suggests that both debt and equity measures of performance are useful in predicting stock price returns. The contingent claims framework of Merton (1974) or structural models was appropriated to provide an alternate framework for value of the firm. Making use of market observable information for the traded debt of the firm it was suggested that the structural model framework could provide a unique framework for firm valuation and estimating the expected return of the share.

The methodology then outlined the procedure for evaluating estimates of expected return under the proposed structural model approach and the CAPM framework. From the exploration of the basic structural models it is evident that the greatest challenge to these models lies within estimating the unobservable parameters in the firm's asset value process. The current methodology of solving from sets of simultaneous equations are shown to negatively prejudice the estimation procedure. The estimates of return obtained from credit spreads on bonds are constrained by the preservation of value in the estimation procedure.

The difference between the brilliance of theories describing an ideal world and the implementation in a vastly more complex reality was emphasised throughout the research. Statistical learning methods were proposed to adjust the models to be applied in a more complex reality. Statistical learning methods were employed to assess how well the created return estimates can be used to predict the class (positive or negative) of future excess returns. The framing of the problem as one of classification extends a further analogy between credit risk and market risk models.

5.2 Conclusions & recommendations for further research

The CAPM estimates of expected return were irrefutably shown to provide poor predictions of future firm returns for the sample evaluated in the research. The fact that average past returns did not provide decent predictions of future returns would suggest non stationarity in prices. The results of the CAPM estimates as stand-alone predictors of future returns validated the second proposition within this research – historic mean-variance constructions are not good forecasts of future return. The implication is clear that distinctions are required between historic estimates and future forecasts of stock returns, especially since markets can never be perfectly efficient.

This should send out a warning to anyone who believes in a free lunch. The law of one price and no-arbitrage provides there can be no free lunch. The laws of statistics also state there are no free lunches. Defiantly modern portfolio theory claims that diversification is the only free lunch in finance. The warning here is that diversification becomes the sole criterion to solving portfolio selection problems only when future asset returns are not predictable or price series are assumed to be stationary. Even Keynes (1939) said, *"to suppose that safety-first consists in having a small gamble in a large number of different companies where I have no information to reach a good judgement, as compared with a substantial stake in a company where one's information is adequate, strikes me as a travesty of investment policy".*

The structural models of debt predictor variables yielded useful predictions for firm returns over the longer time horizon in this research. Within the small firm sample size it was irrefutably shown that the structural model predictors of firm return provide unique explanatory variables to be used in the prediction of equity returns. The lack of easily accessible historic data for corporate debt in South Africa limits the investigation to an analysis of five firms. Albeit a small sample size, the impressive predictions of class of stock returns suggests that structural models of default can be used to capture forward looking expectations of return for individual firms and the approach warrants further investigation.

The return estimates obtained from structural models using Bloomberg default probabilities provided surprisingly accurate forecasts of firm returns in some instances. It suggests then, that exploring default probabilities obtained from reduced form models on observed credit spreads may provide further useful estimates of return under the structural model approach. There is also a nice further interplay then between the use of credit risk models in applications to market risk and equity valuation.

Only the most basic of the contingent claims framework was considered in this research and suggests that extensions of the structural model framework such as the compound option model of Delianedis and Geske (1998) can also be used to capture more complex debt structures. The DG model as well as other structural models are able to make use of more market information from multiple issues of traded corporate debt. The greatest challenge to the implementation of the structural models is the estimation of the hidden asset value process. The re-iterative approaches of the KMV model and more complex estimation procedures outlined in Vassalou and Xing (2004) are worthwhile visiting in this regard.

The use of statistical learning methods in this research was shown to add great value in the quest to predict equity price movements. The structural model estimates of return displayed startling co-integrations with future equity returns in numerous instances although specifying the link function in the statistical learning methodology remained a challenge. Combining more flexible statistical learning methods with more complex structural model approaches in a simulated investment strategy forms the basis for proposed PhD research.

The investigation surrounding the use of structural models to forecast equity prices movements yielded more questions than answers. The question of predictability in equity prices is indubitably better established by the construction of a portfolio strategy based on equity return predictions. Extended class partitions of firm returns corresponding to investment decisions are required in order to garner further insight into the predictability of equity prices.

The question of the exact nature of the reflexive relationship between price and value and how much of this can be captured through the firm's capital structure remains elusive. The study of Elton *et al.* (2001) show, much of the information in the default spread is unrelated to default risk. Vassalou and Xing (2004) conclude that independent of whether the default spread can explain, predict, or otherwise relate to equity returns, such a relation cannot be attributed to the effects that default risk may have on equities.

In conclusion: whether value is regarded as equal to price, the present value of discounted future cash flows or defined by the debt holder's claim on the assets, the relationship between equity returns and *value* is still uncertain. As Yogi Berra is famously quoted as saying *"It's tough to make predictions, especially about the future".*

6 REFERENCES

Alexander, C., 2008, *Market Risk Analysis Volume II: Practical Financial Econometrics*. New York: John Wiley & Sons.

AlHalaseh, R., Islam, A., Bakar, R., 2016, Portfolio Selection Problem: Models Review. *The Social Sciences,* 11(14), pp.3408-3417.

Ang, A., Hodrick, R.K., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. *The Journal of Finance,* 61, 259-299.

Baker, M., Bradley, B., Wurgler, J., 2011. Benchmarks as limits to arbitrage: Understanding the low-volatility anomaly. *Financial Analysts Journal*, 67, 40-54.

Arnott, R., D. 2004. "Blinded by Theory?" *Journal of Portfolio Management*. 30(5), pp. 113-123.

Bai, J., and Wu, L., 2016, Anchoring Credit Default Swap Spreads to Firm Fundamentals. *Journal of Financial and Quantitative Analysis*, 51(5), pp. 1521-1543.

Baesens, B. Rosch, B. Scheule, H., 2016, Credit Risk Analytics: Measurement Techniques, Applications, and Examples in SAS. Hoboken, New Jersey: John Wiley & Sons Ltd.

Bloomberg. (2012) Bloomberg Professional. [Online]. Available at: Subscription Service (Accessed: 09 August 2018).

Chen, R. 2013. On the Geske Compound Option Model When Interest Rates Change Randomly-With an Application to Credit Risk Modeling. Working Paper. New York: Fordham University.

Crosbie, P., & Bohn, J. (2003). Modelling Default Risk. *Moody's* KMV Working Paper.

Crouhy, M., Turnbull, S, M., Wakeman, L, M. 1999. Measuring risk-adjusted performance. *Journal of Risk,* 2(1), 5-35.

Delianedis, G., Geske, R. 1998. Credit Risk and Risk Neutral Default Probabilities: Information about Rating Migrations and Defaults. Working Paper. Los Angeles, California: Anderson School, University of California.

Delianedis, G., Geske, R. 2001.The Components of corporate credit spreads: Default, Recovery, Tax, Jumps, Liquidity and market Factors. Working Paper. Los Angeles, California: Anderson School, University of California.

Duffee, G.R. 1996. *Treasury yields and corporate bond yield spreads: An empirical analysis*. [Washington]: Federal Reserve Board.

Duffie, D. 1998. Credit Swap Valuation. Stanford, California: Graduate School of Business, Stanford University.

Duffie, D., Singleton, K., J. 1999. Modeling Term Structures of Defaultable Bonds. *The Review of Financial Studies*, 12(4), 687-720.

Dowd. K., 2005.*Measuring Market Risk (2e)*. Chichester: John Wiley & Sons.

Duan, J., C. 1994. Maximum Likelihood Estimation Using Price Data of Derivative Contract, *Mathematical Finance*, 4, 155-167.

Elton, E.J., Gruber, M.J., Agrawal, D., Mann, C. 2001. Explaining the Rate Spread on Corporate Bonds. *The Journal of Finance,* 56(1), February:247-277.

Elton, E., Gruber, M., Brown, S. & Goetzmann, W., 2011, *Modern Portfolio Theory and Investment Analysis*. Croyden: John Wiley & Sons Ltd.

Fama, E.F., and French, K.R., 1993, Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*. 33, p. 3-56.

Frazzini, A., and Pedersen, L.H., 2014. Betting against beta. *Journal of Financial Economics*. 111, p. 1-25.

Friedman, M., 1952. *The Methodology of Positive Economics, Essay in Essays on Positive Economics*. University of Chicago Press.

Fons, J.S. 1994. Using Default Rates to Model the Term Structure of Credit Risk. *Financial Analysts Journal*, 50(5), September-October:25-32.

Gregory, J., 2012, *Counterparty Credit Risk and Credit Value Adjustment: A Continuing Challenge for Global Financial Markets.* Croydon: John Wiley & Sons Ltd.

Hayne, E. L. 2004. *Predictions of Default Probabilities in Structural Models of Debt.* Working Paper. Berkeley, California: Haas School of Business, University of California.

Holman, G., Van Breda, R., Correia, C., 2011, The use of the Merton model to quantify the default probabilities of the top 42 non-financial South African firms. *The African Finance Journal,* 13(Conference Issue), pp.1-33.

Hong, H., and Sraer, D. 2013. Quiet bubbles. *Journal of Financial Economics*, 110, 596-606.

Huang, J.H., Huang, M. 2003. How Much of the Corporate-Treasury Yield Spread is Due to Credit Risk?. Working Paper. Stanford, California: Graduate School of Business, Stanford University.

Hull, J.C., 2012, *Options, Futures, and other Derivatives*. Upper Saddle River, New Jersey: Pearson Education.

James, G., Witten, D., Hastie, T., and Tibshirani, R., 2015, *An Introduction to Statistical Learning*. Upper Saddle River, New Jersey: Pearson Education.

Jarrow, R.A., Lando, D., Turnbull, S.M. 1997. A Markov Model for the Term Structure of Credit Risk Spreads. *The Review of Financial Studies*, 10(2), Summer:484-523.

Kahneman, D. and Tversky, A., 1979, Prospect Theory: An Analysis of Decision under Risk. *Econometrica 47(2)*, 263-292.

Keynes, J. M., 1939. “Memorandum for the estates committee,” Kings College, Cambridge. In Moggridge, D., *Collected writings of John Maynard Keynes, Vol XII, 66-68.*New York: Cambridge University Press, 1983.

Kidd, D. 2011. *Measures of Risk-Adjusted Return: Let’s Not Forget Treynor and Jensen*. Investment Risk and Performance Newsletter, CFA Institute.

Lintner, J. 1965. “The Valuation of Risky Assets and the selection of Risky Investments in Stock Portfolios and Capital Budgets”. *Review of Economics and Statistics.* 47, pp. 12-37.

Mandelbrot, B. 1966. Is there persistence in stock price movements? Seminar on the Analysis of Security Prices. Graduate Scholl of Business: University of Chicago

Markowitz, H. 1952. “Portfolio Selection.” Journal of Finance. 7:1, pp. 77-99.

Merton, R., 1990, *Continuous Time Finance.* Basil Blackwell: Oxford.

Modigliani, F. and Miller, M., 1958, ‘The cost of capital, corporation finance and the theory of investment’, *The American Economic Review* 68, 261-97.

Mpofu, R., De Beer, J., Mynhardt, R. & Nortje, A. Investment Management. (4e). Pretoria: Van Schaik. (2013).

Patetta, M. 2010. Predictive Modeling Using Logistic Regression Course Notes. Cary, NC: SAS Institute Inc.

R Development Core Team. 2015. R. Edition 3.2.2 [Online]. Available: http://www.r-project.org/ [2018, April 1].

Roll, R. 1977. “A Critique of the Asset Pricing Theory’s Tests Part 1: on Past and Potential Testability of the Theory”. *Journal of Financial Economics.* 4, pp. 129-176.

Ross, S.A. 1976. “The Arbitrage Theory of Capital Asset Pricing”. *Journal of Economic Theory.* 13, pp. 341-360.

Ross, S.M., 2009, *An Elementary Introduction to Mathematical Finance*. 32 Avenue of the Americas, New York: Cambridge University Press.

Sharpe, W.F. 1964. “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk”. *Journal of Finance.* 19; 3, pp. 425-442.

Smit, L., Swart, B., Van Niekerk, F. 2003. Credit risk models in the South African context. *Investment Analysts Journal*, 57, 41-46.

Soros, G., 1987, *The Alchemy of Finance*. Hoboken, Jersey: John Wiley & Sons. .

Su, Y.C., Huang, H.C, and Lin, Y.J. 2011. GJR-GARCH model in value-at-risk of financial holdings. *Applied Financial Economics,* 21(24), 1189-1829.

Sundaran, R., and Das, S., 2010, *Derivatives: Principals and Practice*. New York: McGraw-Hill Irwin.

Taleb, N., 2004, *Fooled by Randomness (1e)*. Penguin Books.

Trujllo, A. & Martin, J. (2005) Structural Models and Default Probability: Application the Spanish Stock Market. *Investment Management and Financial Innovations,* 2 (1),pp.18-25.

Varian, H. 1993. “A Portfolio of Nobel Laureates: Markowitz, Miller and Sharpe.” Journal of Economic Perspectives. 7:1, pp. 159-169.

Vassalou, M., & Xing, Y. 2004. Default Risk in Equity Returns. *Journal of Finance*, 59, 831-868.

Wang, W., Suo, W. 2006. Assessing Default Probabilities from Structural Credit Risk Models. Working Paper. Kingston, Ontario: Queen’s School of Business, Queen’s University.

Williams, J.B., 1938. *The Theory of Investment Value*. Cambridge: Harvard University Press.

Zaik, E., Walter, J., Kelling, G., and James, C. 1996. RAROC at Bank of America: From theory to practice. *Journal of Applied Corporate Finance,* 9(2), 83-92.

7 ADDENDA

A. ABSA RESULTS

i. Estimates of Alpha, Beta and equity volatility

Figure A.1 Absa estimates of alpha and beta

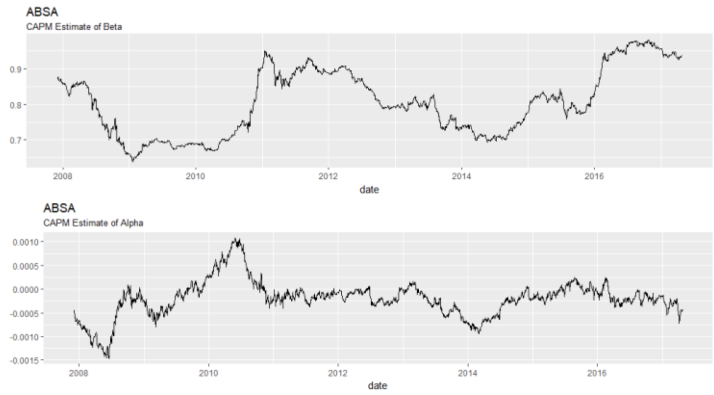
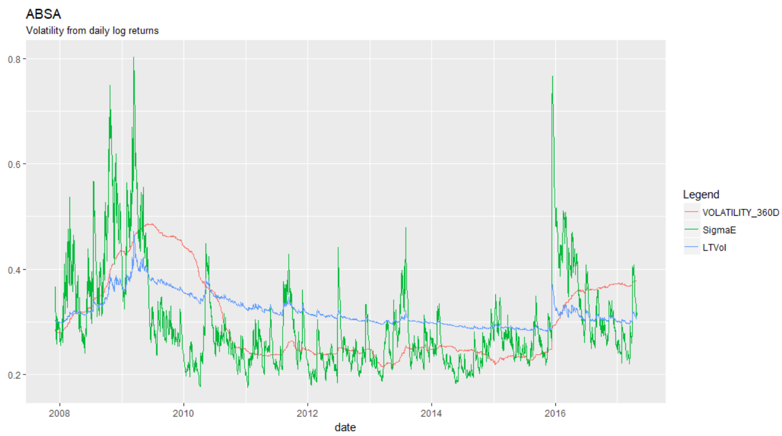


Figure A.2 Absa estimates of equity volatility



ii. Indicator results

Table A.1 Absa 1-year stand-alone prediction performance

ABSA 1 Year Indicator Results													
	K	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_360D_BB	1500	0.669	845	234	280	0	331	0.542	1.000	0.455	0.331	0.000	0.392
Discount_GARCH_CS	1500	0.647	814	514	0	287	13	1.000	0.043	1.000	0.000	0.353	0.369
Discount_LTVol_CS	1500	0.628	819	514	0	305	0	NaN	0.000	1.000	0.000	0.372	0.372
Discount_360D_BB	1500	0.624	745	234	280	0	231	0.452	1.000	0.455	0.376	0.000	0.310
Discount_360D_CS	1500	0.624	824	514	0	310	0	NaN	0.000	1.000	0.000	0.376	0.376
CAPM_1YER	1500	0.485	845	391	123	312	19	0.134	0.057	0.761	0.146	0.369	0.392
ROE_LTVol_BB	1500	0.392	845	0	514	0	331	0.392	1.000	0.000	0.608	0.000	0.392
ROE_360D_CS	1500	0.392	845	0	514	0	331	0.392	1.000	0.000	0.608	0.000	0.392
ROE_LTVol_CS	1500	0.392	845	0	514	0	331	0.392	1.000	0.000	0.608	0.000	0.392
ROE_GARCH_CS	1500	0.378	845	0	514	12	319	0.383	0.964	0.000	0.608	0.014	0.392
Discount_LTVol_BB	1500	0.324	760	0	514	0	246	0.324	1.000	0.000	0.676	0.000	0.324
ROE_GARCH_BB	1500	0.290	769	88	425	121	135	0.241	0.527	0.172	0.553	0.157	0.333
Discount_GARCH_BB	1500	0.280	774	89	425	132	128	0.231	0.492	0.173	0.549	0.171	0.336

Table A.2 Absa 6-months stand-alone prediction performance

ABSA 6 Months Indicator Results												
	K	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_360D_BB	1500	0.702	845	225	243	9	368	0.602	0.976	0.481	0.288	0.011	0.446
Discount_360D_BB	1500	0.672	745	225	235	9	276	0.540	0.968	0.489	0.315	0.012	0.383
Discount_LTVol_CS	1500	0.571	819	468	0	351	0	NaN	0.000	1.000	0.000	0.429	0.429
Discount_GARCH_CS	1500	0.571	814	460	8	341	5	0.385	0.014	0.983	0.010	0.419	0.425
Discount_360D_CS	1500	0.568	824	468	0	356	0	NaN	0.000	1.000	0.000	0.432	0.432
ROE_GARCH_CS	1500	0.449	845	7	461	5	372	0.447	0.987	0.015	0.546	0.006	0.446
ROE_LTVol_BB	1500	0.446	845	0	468	0	377	0.446	1.000	0.000	0.554	0.000	0.446
ROE_360D_CS	1500	0.446	845	0	468	0	377	0.446	1.000	0.000	0.554	0.000	0.446
ROE_LTVol_CS	1500	0.446	845	0	468	0	377	0.446	1.000	0.000	0.554	0.000	0.446
CAPM_6MER	1500	0.439	845	351	117	357	20	0.146	0.053	0.750	0.138	0.422	0.446
Discount_LTVol_BB	1500	0.387	760	0	466	0	294	0.387	1.000	0.000	0.613	0.000	0.387
ROE_GARCH_BB	1500	0.362	769	93	375	116	185	0.330	0.615	0.199	0.488	0.151	0.391
Discount_GARCH_BB	1500	0.350	774	93	375	128	178	0.322	0.582	0.199	0.484	0.165	0.395

Table A.3 Absa 3-months stand-alone prediction performance

ABSA 3 Months Indicator Results												
	K	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_360D_BB	1500	0.669	845	201	247	33	364	0.596	0.917	0.449	0.292	0.039	0.470
Discount_360D_BB	1500	0.626	745	201	246	33	265	0.519	0.889	0.450	0.330	0.044	0.400
Discount_GARCH_CS	1500	0.548	814	440	7	361	6	0.462	0.016	0.984	0.009	0.443	0.451
Discount_LTVol_CS	1500	0.546	819	447	0	372	0	NaN	0.000	1.000	0.000	0.454	0.454
Discount_360D_CS	1500	0.542	824	447	0	377	0	NaN	0.000	1.000	0.000	0.458	0.458
ROE_GARCH_CS	1500	0.472	845	7	441	5	392	0.471	0.987	0.016	0.522	0.006	0.470
ROE_LTVol_BB	1500	0.470	845	0	448	0	397	0.470	1.000	0.000	0.530	0.000	0.470
ROE_360D_CS	1500	0.470	845	0	448	0	397	0.470	1.000	0.000	0.530	0.000	0.470
ROE_LTVol_CS	1500	0.470	845	0	448	0	397	0.470	1.000	0.000	0.530	0.000	0.470
ROE_GARCH_BB	1500	0.438	769	111	334	98	226	0.404	0.698	0.249	0.434	0.127	0.421
Discount_GARCH_BB	1500	0.429	774	113	334	108	219	0.396	0.670	0.253	0.432	0.140	0.422
Discount_LTVol_BB	1500	0.412	760	0	447	0	313	0.412	1.000	0.000	0.588	0.000	0.412
CAPM_3MER	1500	0.395	845	326	122	389	8	0.062	0.020	0.728	0.144	0.460	0.470

iii. KNN prediction Results

Table A.4 Absa 1-year return predictions KNN (K=99)

ABSA 1 Year KNN Prediction Results															
	K	K1	ITD	ITU	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
CAPM_1YER	1500	1249	619	630	0.678	845	512	2	270	61	0.968	0.184	0.996	0.002	0.320	0.392
ROE_360D_BB	1500	1249	619	630	0.670	845	464	50	229	102	0.671	0.308	0.903	0.059	0.271	0.392
Discount_LTVol_CS	1500	1249	645	604	0.634	819	456	58	242	63	0.521	0.207	0.887	0.071	0.295	0.372
Discount_360D_CS	1500	1249	640	609	0.600	824	407	107	223	87	0.448	0.281	0.792	0.130	0.271	0.376
ROE_GARCH_BB	1500	1249	648	601	0.586	769	383	130	188	68	0.343	0.266	0.747	0.169	0.244	0.333
Discount_GARCH_BB	1500	1249	664	585	0.557	774	372	142	201	59	0.294	0.227	0.724	0.183	0.260	0.336
Discount_360D_BB	1500	1249	667	582	0.537	745	298	216	129	102	0.321	0.442	0.580	0.290	0.173	0.310
ROE_LTVol_CS	1500	1249	619	630	0.458	845	274	240	218	113	0.320	0.341	0.533	0.284	0.258	0.392
ROE_360D_CS	1500	1249	619	630	0.434	845	274	240	238	93	0.279	0.281	0.533	0.284	0.282	0.392
Discount_GARCH_CS	1500	1249	637	612	0.434	814	281	233	228	72	0.236	0.240	0.547	0.286	0.280	0.369
Discount_LTVol_BB	1500	1249	679	570	0.386	760	243	271	196	50	0.156	0.203	0.473	0.357	0.258	0.324
ROE_LTVol_BB	1500	1249	619	630	0.382	845	275	239	283	48	0.167	0.145	0.535	0.283	0.335	0.392
ROE_GARCH_CS	1500	1249	619	630	0.368	845	221	293	241	90	0.235	0.272	0.430	0.347	0.285	0.392

Table A.5 Absa 6-month return predictions with KNN (k=99)

ABSA 6 Months KNN Prediction Results																
	K	K1	ITD	ITU	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
Discount_360D_BB	1500	1375	743	632	0.674	745	439	21	222	63	0.750	0.221	0.954	0.028	0.298	0.383
CAPM_6MER	1500	1375	697	678	0.530	845	343	125	272	105	0.457	0.279	0.733	0.148	0.322	0.446
ROE_GARCH_BB	1500	1375	709	666	0.515	769	268	200	173	128	0.390	0.425	0.573	0.260	0.225	0.391
Discount_360D_CS	1500	1375	718	657	0.513	824	285	183	218	138	0.430	0.388	0.609	0.222	0.265	0.432
ROE_360D_BB	1500	1375	697	678	0.509	845	359	109	306	71	0.394	0.188	0.767	0.129	0.362	0.446
Discount_LTVol_CS	1500	1375	723	652	0.436	819	204	264	198	153	0.367	0.436	0.436	0.322	0.242	0.429
Discount_GARCH_BB	1500	1375	735	640	0.408	774	237	231	227	79	0.255	0.258	0.506	0.298	0.293	0.395
ROE_LTVol_BB	1500	1375	697	678	0.399	845	224	244	264	113	0.317	0.300	0.479	0.289	0.312	0.446
ROE_360D_CS	1500	1375	697	678	0.388	845	239	229	288	89	0.280	0.236	0.511	0.271	0.341	0.446
ROE_GARCH_CS	1500	1375	697	678	0.385	845	251	217	303	74	0.254	0.196	0.536	0.257	0.359	0.446
Discount_GARCH_CS	1500	1375	712	663	0.376	814	187	281	227	119	0.298	0.344	0.400	0.345	0.279	0.425
ROE_LTVol_CS	1500	1375	697	678	0.354	845	199	269	277	100	0.271	0.265	0.425	0.318	0.328	0.446
Discount_LTVol_BB	1500	1375	743	632	0.267	760	152	314	243	51	0.140	0.173	0.326	0.413	0.320	0.387

Table A.6 Absa 3-month return predictions with KNN (k=99)

ABSA 3 Months KNN Prediction Results																
	K	K1	ITD	ITU	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_360D_BB	1500	1438	716	722	0.631	845	394	54	258	139	0.720	0.350	0.879	0.064	0.305	0.470
CAPM_3MER	1500	1438	716	722	0.553	845	392	56	322	75	0.573	0.189	0.875	0.066	0.381	0.470
ROE_GARCH_BB	1500	1438	737	701	0.532	769	309	136	224	100	0.424	0.309	0.694	0.177	0.291	0.421
Discount_360D_BB	1500	1438	737	701	0.530	745	367	80	270	28	0.259	0.094	0.821	0.107	0.362	0.400
ROE_LTVol_BB	1500	1438	716	722	0.507	845	266	182	235	162	0.471	0.408	0.594	0.215	0.278	0.470
Discount_LTVol_BB	1500	1438	751	687	0.484	760	209	238	154	159	0.401	0.508	0.468	0.313	0.203	0.412
ROE_360D_CS	1500	1438	716	722	0.465	845	290	158	294	103	0.395	0.259	0.647	0.187	0.348	0.470
Discount_GARCH_BB	1500	1438	761	677	0.464	774	247	200	215	112	0.359	0.343	0.553	0.258	0.278	0.422
Discount_360D_CS	1500	1438	737	701	0.461	824	293	154	290	87	0.361	0.231	0.655	0.187	0.352	0.458
Discount_GARCH_CS	1500	1438	743	695	0.441	814	226	221	234	133	0.376	0.362	0.506	0.271	0.287	0.451
ROE_GARCH_CS	1500	1438	716	722	0.440	845	218	230	243	154	0.401	0.388	0.487	0.272	0.288	0.470
ROE_LTVol_CS	1500	1438	716	722	0.436	845	244	204	273	124	0.378	0.312	0.545	0.241	0.323	0.470
Discount_LTVol_CS	1500	1438	742	696	0.424	819	189	258	214	158	0.380	0.425	0.423	0.315	0.261	0.454

iv. Logistic Regression Results

Table A.7 Absa 1-year logistic regression prediction results

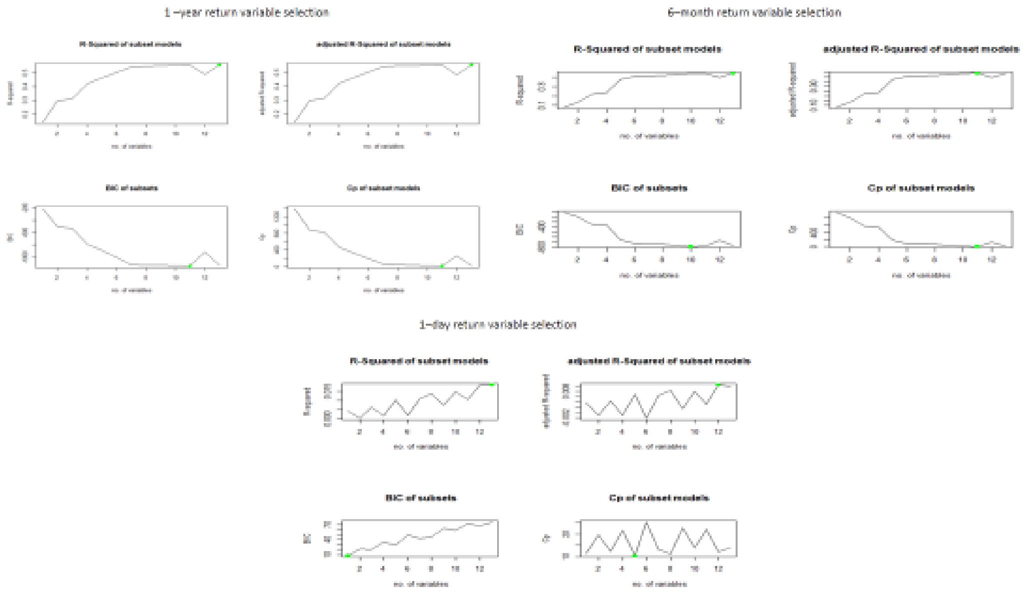
ABSA 1 Year Logistic Regression Prediction Results																	
	K	K1	ITD	ITU	IT	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
CAPM_1YER	1500	1249	619	630	-0.088	0.795	845	513	1	172	159	0.994	0.480	0.998	0.001	0.204	0.392
Discount_LTVol_BB	1500	1249	679	570	0.428	0.693	760	512	2	231	15	0.882	0.061	0.996	0.003	0.304	0.324
ROE_LTVol_BB	1500	1249	619	630	0.565	0.640	845	502	12	292	39	0.765	0.118	0.977	0.014	0.346	0.392
Discount_LTVol_CS	1500	1249	645	604	-0.328	0.640	819	508	6	289	16	0.727	0.052	0.988	0.007	0.353	0.372
Discount_360D_BB	1500	1249	667	582	0.416	0.573	745	298	216	102	129	0.374	0.558	0.580	0.290	0.137	0.310
Discount_GARCH_BB	1500	1249	664	585	0.327	0.556	774	356	158	186	74	0.319	0.285	0.693	0.204	0.240	0.336
Discount_360D_CS	1500	1249	640	609	-0.339	0.524	824	303	211	181	129	0.379	0.416	0.589	0.256	0.220	0.376
ROE_360D_BB	1500	1249	619	630	0.472	0.511	845	303	211	202	129	0.379	0.390	0.589	0.250	0.239	0.392
Discount_GARCH_CS	1500	1249	637	612	-0.384	0.462	814	292	222	216	84	0.275	0.280	0.568	0.273	0.265	0.369
ROE_GARCH_CS	1500	1249	619	630	0.041	0.367	845	178	336	199	132	0.282	0.399	0.346	0.398	0.236	0.392
ROE_360D_CS	1500	1249	619	630	0.041	0.354	845	205	309	237	94	0.233	0.284	0.399	0.366	0.280	0.392
ROE_GARCH_BB	1500	1249	648	601	1.314	0.315	769	180	333	194	62	0.157	0.242	0.351	0.433	0.252	0.333
ROE_LTVol_CS	1500	1249	619	630	0.044	0.310	845	196	318	265	66	0.172	0.199	0.381	0.376	0.314	0.392

Table A.8 Absa 6-month logistic regression prediction results

ABSA 6 Month Logistic Regression Prediction Results																	
	K	K1	ITD	ITU	IT	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
Discount_LTVol_BB	1500	1375	743	632	0.418	0.617	760	464	2	289	5	0.714	0.017	0.996	0.003	0.380	0.387
Discount_LTVol_CS	1500	1375	723	652	-0.326	0.571	819	463	5	346	5	0.500	0.014	0.989	0.006	0.422	0.429
ROE_LTVol_BB	1500	1375	697	678	0.558	0.567	845	464	4	362	15	0.789	0.040	0.991	0.005	0.428	0.446
CAPM_6MER	1500	1375	697	678	-0.094	0.528	845	437	31	368	9	0.225	0.024	0.934	0.037	0.436	0.446
ROE_LTVol_CS	1500	1375	697	678	0.044	0.528	845	315	153	246	131	0.461	0.347	0.673	0.181	0.291	0.446
Discount_360D_BB	1500	1375	743	632	0.411	0.526	745	271	189	164	121	0.390	0.425	0.589	0.254	0.220	0.383
ROE_360D_BB	1500	1375	697	678	0.441	0.480	845	267	201	238	139	0.409	0.369	0.571	0.238	0.282	0.446
Discount_GARCH_BB	1500	1375	735	640	0.403	0.478	774	270	198	206	100	0.336	0.327	0.577	0.256	0.266	0.395
Discount_360D_CS	1500	1375	718	657	-0.337	0.468	824	279	189	249	107	0.361	0.301	0.596	0.229	0.302	0.432
Discount_GARCH_CS	1500	1375	712	663	-0.351	0.403	814	231	237	249	97	0.290	0.280	0.494	0.291	0.306	0.425
ROE_360D_CS	1500	1375	697	678	0.039	0.389	845	205	263	253	124	0.320	0.329	0.438	0.311	0.299	0.446
ROE_GARCH_BB	1500	1375	709	666	0.496	0.326	769	195	273	245	56	0.170	0.186	0.417	0.355	0.319	0.391
ROE_GARCH_CS	1500	1375	697	678	0.028	0.289	845	204	264	337	40	0.132	0.106	0.436	0.312	0.399	0.446

v. Variable Selection

Figure A.3 Absa variable subset selection results



Run step Wise AIC variable Selection

```
=============================================
                    Dependent variable:
                  ---------------------------
                           ER 1Y
---------------------------------------------
Discount_GARCH_BB          96.773***
CAPM                      -17.458***
Discount_360D_BB          -88.638***
Discount_LTVol_CS         -37.194***
ROE_GARCH_CS             -216.649***
Discount_GARCH_CS        -177.197***
```

Discount_360D_CS	178.056***
ROE_LTVol_CS	229.608***
ROE_360D_CS	65.315***
Constant	-19.519***
Observations	1,500
Log Likelihood	-764.913
Akaike Inf. Crit.	1,549.826
Note:	*p<0.1; **p<0.05; ***p<0.01

	Dependent variable:
	ER 6M
ROE_360D_BB	4.780***
ROE_360D_CS	-63.210**
Discount_360D_CS	173.011***
ROE_LTVol_CS	329.090***
Discount_GARCH_BB	94.514***
Discount_GARCH_CS	-168.429***
ROE_GARCH_CS	-58.902***
Discount_360D_BB	-98.231***
ROE_LTVol_BB	-35.382***
CAPM	-17.324***
Discount_LTVol_BB	12.951***
Constant	5.227
Observations	1,500
Log Likelihood	-815.969
Akaike Inf. Crit.	1,655.937
Note:	*p<0.1; **p<0.05; ***p<0.01

	Dependent variable:
	ER 1D
ROE_LTVol_CS	18.205**
Constant	-0.828**
Observations	1,500
Log Likelihood	-1,036.788
Akaike Inf. Crit.	2,077.576
Note:	*p<0.1; **p<0.05; ***p<0.01

B. INL RESULTS

i. Estimates of alpha and Beta and equity volatility

Figure B.1 INL equity return volatility estimates

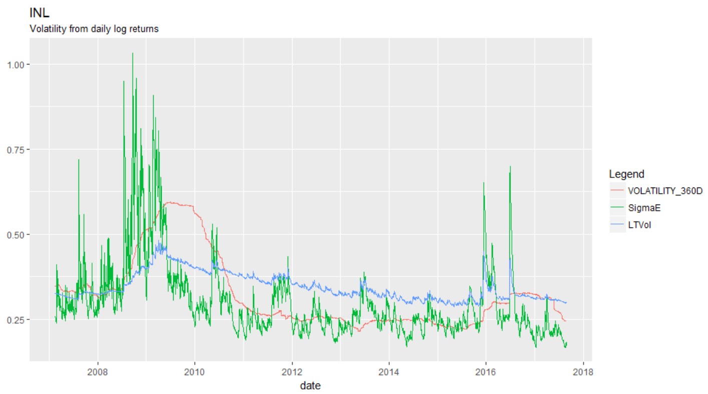
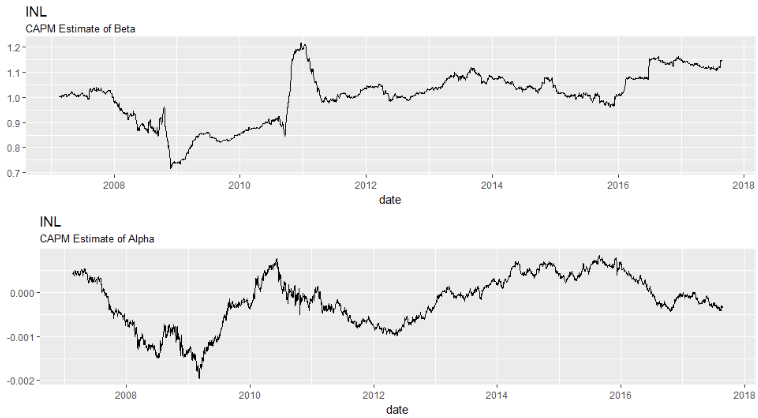


Figure B.2 INL estimates of alpha and beta



ii. Indicator Results

Table B.1INL stand-alone predictor performance for 1-year returns

INL 1 Year Indicator Results													
	K	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_GARCH_CS	1500	0.544	1248	1	569	0	678	0.544	1.000	0.002	0.456	0.000	0.543
ROE_360D_CS	1500	0.543	1248	0	570	0	678	0.543	1.000	0.000	0.457	0.000	0.543
ROE_LTVol_CS	1500	0.543	1248	0	570	0	678	0.543	1.000	0.000	0.457	0.000	0.543
ROE_GARCH_BB	1500	0.477	1019	476	23	510	10	0.303	0.019	0.954	0.023	0.500	0.510
Discount_360D_BB	1500	0.457	1248	570	0	678	0	NaN	0.000	1.000	0.000	0.543	0.543
ROE_360D_BB	1500	0.457	1248	570	0	678	0	NaN	0.000	1.000	0.000	0.543	0.543
Discount_LTVol_CS	1500	0.457	1248	570	0	678	0	NaN	0.000	1.000	0.000	0.543	0.543
Discount_LTVol_BB	1500	0.456	1248	569	1	678	0	0.000	0.000	0.998	0.001	0.543	0.543
ROE_LTVol_BB	1500	0.456	1248	569	1	678	0	0.000	0.000	0.998	0.001	0.543	0.543
Discount_GARCH_BB	1500	0.442	1238	537	23	668	10	0.303	0.015	0.959	0.019	0.540	0.548
CAPM_1YER	1500	0.431	1248	157	413	297	381	0.480	0.562	0.275	0.331	0.238	0.543
Discount_360D_CS	1500	0.390	1104	431	0	673	0	NaN	0.000	1.000	0.000	0.610	0.610
Discount_GARCH_CS	1500	0.338	786	266	0	520	0	NaN	0.000	1.000	0.000	0.662	0.662

Table B.2 INL 6-month return stand-alone prediction

INL 6 Months Indicator Results													
	K	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_GARCH_CS	1500	0.575	1248	1	531	0	716	0.574	1.000	0.002	0.425	0.000	0.574
ROE_360D_CS	1500	0.574	1248	0	532	0	716	0.574	1.000	0.000	0.426	0.000	0.574
ROE_LTVol_CS	1500	0.574	1248	0	532	0	716	0.574	1.000	0.000	0.426	0.000	0.574
CAPM_6MER	1500	0.442	1248	169	363	334	382	0.513	0.534	0.318	0.291	0.268	0.574
Discount_360D_BB	1500	0.426	1248	532	0	716	0	NaN	0.000	1.000	0.000	0.574	0.574
ROE_360D_BB	1500	0.426	1248	532	0	716	0	NaN	0.000	1.000	0.000	0.574	0.574
Discount_LTVol_CS	1500	0.426	1248	532	0	716	0	NaN	0.000	1.000	0.000	0.574	0.574
Discount_LTVol_BB	1500	0.425	1248	531	1	716	0	0.000	0.000	0.998	0.001	0.574	0.574
ROE_LTVol_BB	1500	0.425	1248	531	1	716	0	0.000	0.000	0.998	0.001	0.574	0.574
ROE_GARCH_BB	1500	0.421	1019	421	25	565	8	0.242	0.014	0.944	0.025	0.554	0.562
Discount_GARCH_BB	1500	0.410	1238	499	25	706	8	0.242	0.011	0.952	0.020	0.570	0.577
Discount_360D_CS	1500	0.384	1104	424	0	680	0	NaN	0.000	1.000	0.000	0.616	0.616
Discount_GARCH_CS	1500	0.377	786	296	0	490	0	NaN	0.000	1.000	0.000	0.623	0.623

Table B.3 INL 3-months stand-alone prediction performance

INL 3 Months Indicator Results													
	K	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_GARCH_CS	1500	0.567	1248	1	541	0	706	0.566	1.000	0.002	0.433	0.000	0.566
ROE_360D_CS	1500	0.566	1248	0	542	0	706	0.566	1.000	0.000	0.434	0.000	0.566
ROE_LTVol_CS	1500	0.566	1248	0	542	0	706	0.566	1.000	0.000	0.434	0.000	0.566
CAPM_3MER	1500	0.498	1248	216	326	300	406	0.555	0.575	0.399	0.261	0.240	0.566
Discount_360D_BB	1500	0.434	1248	542	0	706	0	NaN	0.000	1.000	0.000	0.566	0.566
ROE_360D_BB	1500	0.434	1248	542	0	706	0	NaN	0.000	1.000	0.000	0.566	0.566
Discount_LTVol_CS	1500	0.434	1248	542	0	706	0	NaN	0.000	1.000	0.000	0.566	0.566
Discount_LTVol_BB	1500	0.433	1248	541	1	706	0	0.000	0.000	0.998	0.001	0.566	0.566
ROE_LTVol_BB	1500	0.433	1248	541	1	706	0	0.000	0.000	0.998	0.001	0.566	0.566
ROE_GARCH_BB	1500	0.425	1019	422	22	564	11	0.333	0.019	0.950	0.022	0.553	0.564
Discount_GARCH_BB	1500	0.422	1238	511	22	694	11	0.333	0.016	0.959	0.018	0.561	0.569
Discount_GARCH_CS	1500	0.416	786	327	0	459	0	NaN	0.000	1.000	0.000	0.584	0.584
Discount_360D_CS	1500	0.411	1104	454	0	650	0	NaN	0.000	1.000	0.000	0.589	0.589

iii. KNN Prediction Results

Table B.4 INL 1-year return predictions with KNN

INL 1 Year KNN Prediction 5																
	K	K1	ITD	ITU	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_360D_BB	1500	1249	952	297	0.648	1248	247	323	116	562	0.635	0.829	0.433	0.259	0.093	0.543
ROE_360D_CS	1500	1249	952	297	0.637	1248	377	193	260	418	0.684	0.617	0.661	0.155	0.208	0.543
ROE_LTVol_CS	1500	1249	952	297	0.542	1248	452	118	453	225	0.656	0.332	0.793	0.095	0.363	0.543
ROE_GARCH_CS	1500	1249	952	297	0.538	1248	452	118	459	219	0.650	0.323	0.793	0.095	0.368	0.543
Discount_360D_BB	1500	1249	952	297	0.533	1248	105	465	118	560	0.546	0.826	0.184	0.373	0.095	0.543
Discount_360D_CS	1500	1249	952	297	0.418	1104	196	235	407	266	0.531	0.395	0.455	0.213	0.369	0.610
ROE_GARCH_BB	1500	1249	936	313	0.397	1019	249	250	364	156	0.384	0.300	0.499	0.245	0.357	0.510
Discount_GARCH_BB	1500	1249	952	297	0.363	1238	166	394	395	283	0.418	0.417	0.296	0.318	0.319	0.548
Discount_GARCH_CS	1500	1249	940	309	0.360	786	150	116	387	133	0.534	0.256	0.564	0.148	0.492	0.662
ROE_LTVol_BB	1500	1249	952	297	0.330	1248	184	386	450	228	0.371	0.336	0.323	0.309	0.361	0.543
CAPM_1YER	1500	1249	952	297	0.290	1248	311	259	627	51	0.165	0.075	0.546	0.208	0.502	0.543
Discount_LTVol_BB	1500	1249	952	297	0.256	1248	136	434	494	184	0.298	0.271	0.239	0.348	0.396	0.543
Discount_LTVol_CS	1500	1249	952	297	0.219	1248	158	412	563	115	0.218	0.170	0.277	0.330	0.451	0.543

Table B.5 INL 6-month return predictions with KNN

INL 6 Months KNN Prediction 5																
	K	K1	ITD	ITU	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_360D_CS	1500	1375	972	403	0.583	1248	294	238	282	434	0.646	0.606	0.553	0.191	0.226	0.574
ROE_360D_BB	1500	1375	972	403	0.522	1248	268	264	333	383	0.592	0.535	0.504	0.212	0.267	0.574
Discount_360D_BB	1500	1375	972	403	0.514	1248	237	295	311	405	0.579	0.566	0.445	0.236	0.249	0.574
ROE_GARCH_CS	1500	1375	972	403	0.494	1248	381	151	480	236	0.610	0.330	0.716	0.121	0.385	0.574
ROE_LTVol_BB	1500	1375	972	403	0.470	1248	321	211	450	266	0.558	0.372	0.603	0.169	0.361	0.574
Discount_GARCH_CS	1500	1375	960	415	0.468	786	192	104	314	176	0.629	0.359	0.649	0.132	0.399	0.623
ROE_LTVol_CS	1500	1375	972	403	0.463	1248	330	202	468	248	0.551	0.346	0.620	0.162	0.375	0.574
Discount_360D_CS	1500	1375	972	403	0.447	1104	274	150	461	219	0.593	0.322	0.646	0.136	0.418	0.616
Discount_GARCH_BB	1500	1375	972	403	0.436	1238	205	319	379	335	0.512	0.469	0.391	0.258	0.306	0.577
ROE_GARCH_BB	1500	1375	955	420	0.436	1019	254	192	383	190	0.497	0.332	0.570	0.188	0.376	0.562
Discount_LTVol_BB	1500	1375	972	403	0.427	1248	283	249	466	250	0.501	0.349	0.532	0.200	0.373	0.574
Discount_LTVol_CS	1500	1375	972	403	0.422	1248	256	276	445	271	0.495	0.378	0.481	0.221	0.357	0.574
CAPM_6MER	1500	1375	972	403	0.381	1248	308	224	548	168	0.429	0.235	0.579	0.179	0.439	0.574

iv. Logistic Regression Results

Table B.6 INL 1-year return predictions with logistic regression

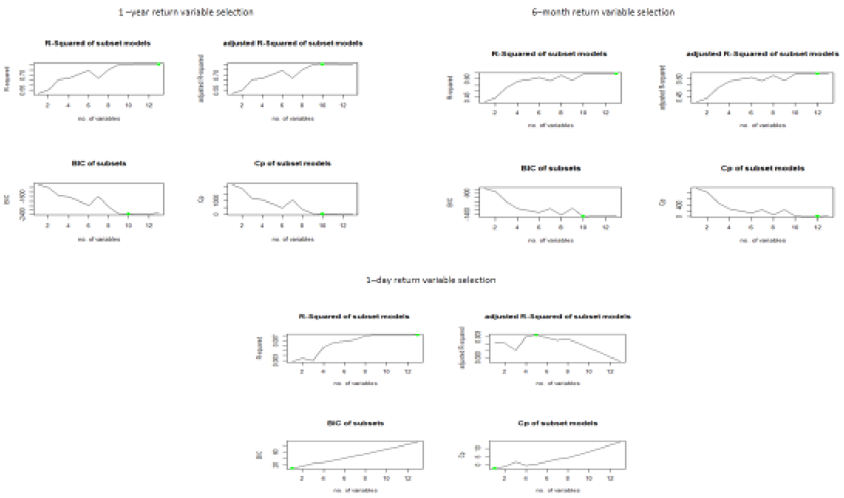
INL 1 Year Logistic Regression Prediction Results																	
	K	K1	ITD	ITU	IT	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_360D_BB	1500	1249	952	297	0.829	0.506	1248	173	397	220	458	0.536	0.676	0.304	0.318	0.176	0.543
CAPM_1YER	1500	1249	952	297	-0.178	0.478	1248	523	47	605	73	0.608	0.108	0.918	0.038	0.485	0.543
ROE_GARCH_BB	1500	1249	936	313	0.155	0.410	1019	392	107	494	26	0.195	0.050	0.786	0.105	0.485	0.510
Discount_LTVol_CS	1500	1249	952	297	-0.380	0.357	1248	435	135	668	10	0.069	0.015	0.763	0.108	0.535	0.543
Discount_GARCH_BB	1500	1249	952	297	0.189	0.300	1238	355	205	661	17	0.077	0.025	0.634	0.166	0.534	0.548
Discount_360D_BB	1500	1249	952	297	0.340	0.296	1248	147	423	455	223	0.345	0.329	0.258	0.339	0.365	0.543
Discount_GARCH_CS	1500	1249	940	309	-0.279	0.295	786	209	57	497	23	0.288	0.044	0.786	0.073	0.632	0.662
Discount_LTVol_BB	1500	1249	952	297	0.056	0.249	1248	248	322	615	63	0.164	0.093	0.435	0.258	0.493	0.543
ROE_LTVol_BB	1500	1249	952	297	0.066	0.240	1248	232	338	611	67	0.165	0.099	0.407	0.271	0.490	0.543
ROE_360D_CS	1500	1249	952	297	0.005	0.238	1248	6	564	387	291	0.340	0.429	0.011	0.452	0.310	0.543
Discount_360D_CS	1500	1249	952	297	-0.196	0.201	1104	6	425	457	216	0.337	0.321	0.014	0.385	0.414	0.610
ROE_LTVol_CS	1500	1249	952	297	0.035	0.188	1248	48	522	491	187	0.264	0.276	0.084	0.418	0.393	0.543
ROE_GARCH_CS	1500	1249	952	297	-0.009	0.175	1248	23	547	483	195	0.263	0.288	0.040	0.438	0.387	0.543

Table B.7 INL 6-month return predictions with logistic regression

INL 6 Months Logistic Regression Prediction Results																	
	K	K1	ITD	ITU	IT	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
Discount_LTVol_CS	1500	1375	972	403	-0.383	0.419	1248	522	10	715	1	0.091	0.001	0.981	0.008	0.573	0.574
ROE_GARCH_CS	1500	1375	972	403	-0.083	0.417	1248	519	13	714	2	0.133	0.003	0.976	0.010	0.572	0.574
ROE_GARCH_BB	1500	1375	955	420	-71.358	0.416	1019	398	48	547	26	0.351	0.045	0.892	0.047	0.537	0.562
ROE_360D_BB	1500	1375	972	403	-3.209	0.415	1248	332	200	530	186	0.482	0.260	0.624	0.160	0.425	0.574
CAPM_6MER	1500	1375	972	403	-0.098	0.403	1248	503	29	716	0	0.000	0.000	0.945	0.023	0.574	0.574
ROE_360D_CS	1500	1375	972	403	0.009	0.396	1248	60	472	282	434	0.479	0.606	0.113	0.378	0.226	0.574
Discount_GARCH_CS	1500	1375	960	415	-0.101	0.377	786	296	0	490	0	NaN	0.000	1.000	0.000	0.623	0.623
Discount_GARCH_BB	1500	1375	972	403	0.446	0.362	1238	444	80	710	4	0.048	0.006	0.847	0.065	0.574	0.577
Discount_LTVol_BB	1500	1375	972	403	0.062	0.337	1248	404	128	699	17	0.117	0.024	0.759	0.103	0.560	0.574
Discount_360D_CS	1500	1375	972	403	-0.232	0.331	1104	365	59	680	0	0.000	0.000	0.861	0.053	0.616	0.616
ROE_LTVol_BB	1500	1375	972	403	0.090	0.324	1248	371	161	683	33	0.170	0.046	0.697	0.129	0.547	0.574
Discount_360D_BB	1500	1375	972	403	0.294	0.296	1248	349	183	695	21	0.103	0.029	0.656	0.147	0.557	0.574
ROE_LTVol_CS	1500	1375	972	403	0.033	0.216	1248	249	283	695	21	0.069	0.029	0.468	0.227	0.557	0.574

v. Variable Selection

Figure B.3 INL variable subset selection results



	Dependent variable:
	ER 1Y
CAPM	-57.957***
ROE_360D_CS	-851.616*
ROE_360D_BB	-12.576***
Discount_360D_CS	-1,646.121***
Discount_GARCH_BB	545.248***
ROE_GARCH_CS	9.777
ROE_LTVol_BB	-57.564**
Discount_GARCH_CS	-827.830***
Discount_LTVol_CS	2,787.932***
Discount_LTVol_BB	-1,752.573***
Discount_360D_BB	1,082.092***
ROE_LTVol_CS	3,334.110***
ROE_GARCH_BB	-0.489***
Constant	46.460*
Observations	1,500
Log Likelihood	-118.372
Akaike Inf. Crit.	264.745
Note:	*p<0.1; **p<0.05; ***p<0.01

	Dependent variable:
	ER 6M
CAPM	-24.069***
ROE_360D_CS	-915.462***
ROE_LTVol_CS	656.446***
Discount_360D_CS	-455.974***
Discount_GARCH_BB	91.899***
ROE_LTVol_BB	-46.884***
Discount_GARCH_CS	-140.688***
ROE_360D_BB	0.859***

```
Discount_LTVol_CS         451.185***
Discount_360D_BB          279.066***
Discount_LTVol_BB        -235.011***
Constant                  -58.181***
---------------------------------------------
Observations                 1,500
Log Likelihood              -524.266
Akaike Inf. Crit.           1,072.532
=============================================
Note:                *p<0.1; **p<0.05; ***p<0.01


=============================================
                      Dependent variable:
                  ---------------------------
                             ER 1D
---------------------------------------------
ROE_LTVol_BB                 0.508
Constant                    -0.049
---------------------------------------------
Observations                 1,500
Log Likelihood             -1,037.273
Akaike Inf. Crit.           2,078.545
=============================================
Note:                *p<0.1; **p<0.05; ***p<0.01
```

vi. INL Summary of 6 month return predictions

Table B.8 INL summary of top 6-month return predictions

INL 6 Months Prediction Results																		
Link	Predictor Variables		K	K1	ITD	ITU	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
KNN 99	ROE_360D_BB	Discount_GARCH_CS	1500	1375	960	415	0.641	786	135	161	121	369	0.696	0.753	0.456	0.205	0.154	0.623
KNN 99	ROE_360D_BB	Discount_360D_CS	1500	1375	972	403	0.627	1104	207	217	195	485	0.691	0.713	0.488	0.197	0.177	0.616
KNN 5		ROE_360D_CS	1500	1375	972	403	0.583	1248	294	238	282	434	0.646	0.606	0.553	0.191	0.226	0.574
IND		ROE_GARCH_CS	1500				0.575	1248	1	531	0	716	0.574	1.000	0.002	0.425	0.000	0.574
IND		ROE_360D_CS	1500				0.574	1248	0	532	0	716	0.574	1.000	0.000	0.426	0.000	0.574
KNN 5		ROE_360D_BB	1500	1375	972	403	0.522	1248	268	264	333	383	0.592	0.535	0.504	0.212	0.267	0.574
LOGIT	ROE_360D_BB	ROE_360D_CS	1500	1375	972	403	0.486	1248	63	469	172	544	0.537	0.760	0.118	0.376	0.138	0.574
LOGIT	CAPM_6MER	Discount_LTVol_CS	1500	1375	972	403	0.438	1248	425	107	594	122	0.533	0.170	0.799	0.086	0.476	0.574
LOGIT		Discount_LTVol_CS	1500	1375	972	403	0.419	1248	522	10	715	1	0.091	0.001	0.981	0.008	0.573	0.574
LOGIT		ROE_GARCH_CS	1500	1375	972	403	0.417	1248	519	13	714	2	0.133	0.003	0.976	0.010	0.572	0.574

vii. INL 6 month debt variable efficiency

Table B.9 INL variable efficiency for 6-month return prediction

INL 6 Months Predictor Variable Efficiency													
	K	K1	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
CS	1500	1375	0.426	1248	532	0	716	0	NaN	0.000	1.000	0.000	0.574	0.574
BB	1500	1375	0.256	1248	286	246	682	34	0.121	0.047	0.538	0.197	0.546	0.574
CAPM	1500	1375	0.469	1248	409	123	540	176	0.589	0.246	0.769	0.099	0.433	0.574
ALL	1500	1375	0.299	1248	218	314	561	155	0.330	0.216	0.410	0.252	0.450	0.574
BBCS	1500	1375	0.260	1248	268	264	660	56	0.175	0.078	0.504	0.212	0.529	0.574
CAPMCS	1500	1375	0.464	1248	471	61	608	108	0.639	0.151	0.885	0.049	0.487	0.574
CAPMBB	1500	1375	0.296	1248	214	318	561	155	0.328	0.216	0.402	0.255	0.450	0.574

C. GRF RESULTS

i. Estimates of alpha, beta and equity volatility

Figure C.1 GRF estimates of equity volatility

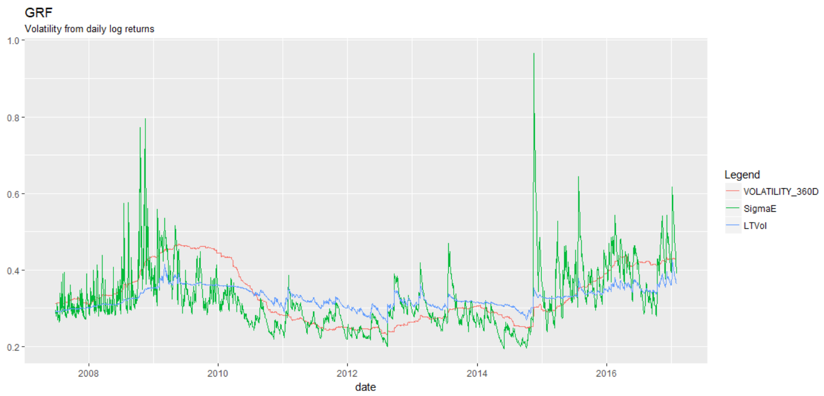
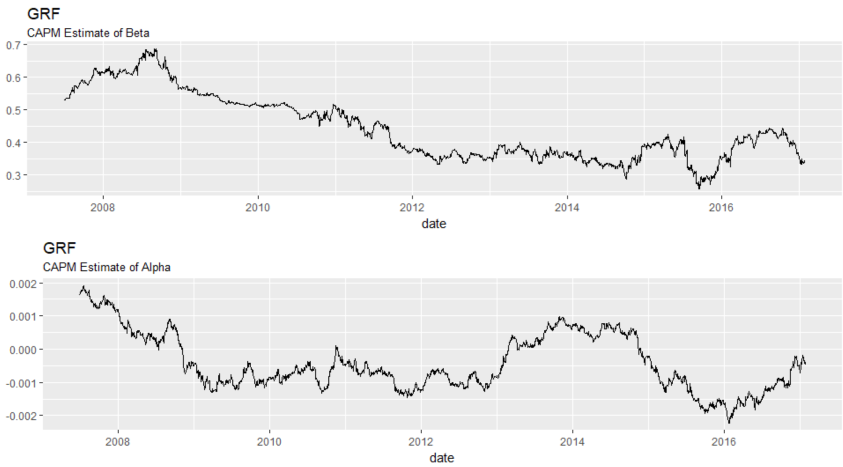


Figure C.2 GRF estimates of alpha and beta



ii. Indicator Results

Table C.1GRF 1-year return prediction performance on stand-alone basis

GRF 1 Year Indicator Performance													
	K	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
Discount_LTVol_CS	1500	0.840	892	749	0	143	0	NaN	0.000	1.000	0.000	0.160	0.160
Discount_360D_CS	1500	0.825	795	656	0	139	0	NaN	0.000	1.000	0.000	0.175	0.175
Discount_GARCH_CS	1500	0.815	807	658	10	139	0	0.000	0.000	0.985	0.012	0.172	0.172
Discount_LTVol_BB	1500	0.796	898	623	130	53	92	0.414	0.634	0.827	0.145	0.059	0.161
ROE_LTVol_BB	1500	0.796	898	623	130	53	92	0.414	0.634	0.827	0.145	0.059	0.161
Discount_GARCH_BB	1500	0.726	898	531	222	24	121	0.353	0.834	0.705	0.247	0.027	0.161
ROE_GARCH_BB	1500	0.725	887	522	222	22	121	0.353	0.846	0.702	0.250	0.025	0.161
Discount_360D_BB	1500	0.710	898	507	246	14	131	0.347	0.903	0.673	0.274	0.016	0.161
ROE_360D_BB	1500	0.710	898	507	246	14	131	0.347	0.903	0.673	0.274	0.016	0.161
CAPM_1YER	1500	0.462	898	409	344	139	6	0.017	0.041	0.543	0.383	0.155	0.161
ROE_GARCH_CS	1500	0.173	898	10	743	0	145	0.163	1.000	0.013	0.827	0.000	0.161
ROE_360D_CS	1500	0.161	898	0	753	0	145	0.161	1.000	0.000	0.839	0.000	0.161
ROE_LTVol_CS	1500	0.161	898	0	753	0	145	0.161	1.000	0.000	0.839	0.000	0.161

Table C.2 GRF 6-month return prediction performance on stand-alone basis

GRF 6 Months Indicator Performance													
	K	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
Discount_LTVol_CS	1500	0.759	892	677	0	215	0	NaN	0.000	1.000	0.000	0.241	0.241
Discount_360D_CS	1500	0.757	795	602	0	193	0	NaN	0.000	1.000	0.000	0.243	0.243
Discount_GARCH_CS	1500	0.748	807	604	10	193	0	0.000	0.000	0.984	0.012	0.239	0.239
Discount_360D_BB	1500	0.707	898	468	210	53	167	0.443	0.759	0.690	0.234	0.059	0.245
ROE_360D_BB	1500	0.707	898	468	210	53	167	0.443	0.759	0.690	0.234	0.059	0.245
ROE_GARCH_BB	1500	0.691	887	471	201	73	142	0.414	0.660	0.701	0.227	0.082	0.242
Discount_GARCH_BB	1500	0.689	898	477	201	78	142	0.414	0.645	0.704	0.224	0.087	0.245
Discount_LTVol_BB	1500	0.639	898	515	163	161	59	0.266	0.268	0.760	0.182	0.179	0.245
ROE_LTVol_BB	1500	0.639	898	515	163	161	59	0.266	0.268	0.760	0.182	0.179	0.245
CAPM_6MER	1500	0.461	898	371	307	177	43	0.123	0.195	0.547	0.342	0.197	0.245
ROE_GARCH_CS	1500	0.256	898	10	668	0	220	0.248	1.000	0.015	0.744	0.000	0.245
ROE_360D_CS	1500	0.245	898	0	678	0	220	0.245	1.000	0.000	0.755	0.000	0.245
ROE_LTVol_CS	1500	0.245	898	0	678	0	220	0.245	1.000	0.000	0.755	0.000	0.245

iii. KNN Prediction Results

Table C.3 GRF 1-year return predictions with KNN

GRF 1 Year KNN 5 Prediction Performance																
	K	K1	ITD	ITU	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_360D_BB	1500	1249	958	291	0.684	898	611	142	142	3	0.021	0.021	0.811	0.158	0.158	0.161
Discount_360D_CS	1500	1249	957	292	0.682	795	542	114	139	0	0.000	0.000	0.826	0.143	0.175	0.175
ROE_GARCH_BB	1500	1249	958	291	0.678	887	555	189	97	46	0.196	0.322	0.746	0.213	0.109	0.161
Discount_LTVol_CS	1500	1249	957	292	0.649	892	573	176	137	6	0.033	0.042	0.765	0.197	0.154	0.160
Discount_GARCH_CS	1500	1249	957	292	0.638	807	476	192	100	39	0.169	0.281	0.713	0.238	0.124	0.172
Discount_GARCH_BB	1500	1249	958	291	0.635	898	535	218	110	35	0.138	0.241	0.710	0.243	0.122	0.161
CAPM_1YER	1500	1249	958	291	0.620	898	546	207	134	11	0.050	0.076	0.725	0.231	0.149	0.161
Discount_LTVol_BB	1500	1249	958	291	0.598	898	504	249	112	33	0.117	0.228	0.669	0.277	0.125	0.161
Discount_360D_BB	1500	1249	958	291	0.559	898	502	251	145	0	0.000	0.000	0.667	0.280	0.161	0.161
ROE_LTVol_BB	1500	1249	958	291	0.523	898	426	327	101	44	0.119	0.303	0.566	0.364	0.112	0.161
ROE_360D_CS	1500	1249	958	291	0.463	898	395	358	124	21	0.055	0.145	0.525	0.399	0.138	0.161
ROE_LTVol_CS	1500	1249	958	291	0.455	898	383	370	119	26	0.066	0.179	0.509	0.412	0.133	0.161
ROE_GARCH_CS	1500	1249	958	291	0.425	898	356	397	119	26	0.061	0.179	0.473	0.442	0.133	0.161

Table C.4 GRF 6-month return predictions with KNN

GRF 6 Months KNN 5 Prediction Performance																
	K	K1	ITD	ITU	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_360D_BB	1500	1375	911	464	0.619	898	500	178	164	56	0.239	0.255	0.737	0.198	0.183	0.245
CAPM_6MER	1500	1375	911	464	0.606	898	512	166	188	32	0.162	0.145	0.755	0.185	0.209	0.245
Discount_LTVol_CS	1500	1375	905	470	0.604	892	489	188	165	50	0.210	0.233	0.722	0.211	0.185	0.241
ROE_GARCH_BB	1500	1375	900	475	0.595	887	450	222	137	78	0.260	0.363	0.670	0.250	0.154	0.242
Discount_360D_CS	1500	1375	828	547	0.580	795	429	173	161	32	0.156	0.166	0.713	0.218	0.203	0.243
Discount_GARCH_BB	1500	1375	911	464	0.533	898	404	274	145	75	0.215	0.341	0.596	0.305	0.161	0.245
Discount_GARCH_CS	1500	1375	827	548	0.513	807	330	284	109	84	0.228	0.435	0.537	0.352	0.135	0.239
Discount_LTVol_BB	1500	1375	911	464	0.488	898	382	296	164	56	0.159	0.255	0.563	0.330	0.183	0.245
ROE_360D_CS	1500	1375	911	464	0.469	898	362	316	161	59	0.157	0.268	0.534	0.352	0.179	0.245
Discount_360D_BB	1500	1375	911	464	0.453	898	367	311	180	40	0.114	0.182	0.541	0.346	0.200	0.245
ROE_LTVol_BB	1500	1375	911	464	0.440	898	323	355	148	72	0.169	0.327	0.476	0.395	0.165	0.245
ROE_GARCH_CS	1500	1375	911	464	0.355	898	245	433	146	74	0.146	0.336	0.361	0.482	0.163	0.245
ROE_LTVol_CS	1500	1375	911	464	0.305	898	235	443	181	39	0.081	0.177	0.347	0.493	0.202	0.245

Table C.5 GRF 3-month return predictions with KNN

GRF 3 Months KNN 5 Prediction Performance																
	K	K1	ITD	ITU	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_360D_CS	1500	1438	793	645	0.561	898	327	279	115	177	0.388	0.606	0.540	0.311	0.128	0.325
ROE_GARCH_BB	1500	1438	793	645	0.550	887	368	238	161	120	0.335	0.427	0.607	0.268	0.182	0.317
Discount_GARCH_CS	1500	1438	731	707	0.524	807	313	239	145	110	0.315	0.431	0.567	0.296	0.180	0.316
Discount_LTVol_CS	1500	1438	787	651	0.512	892	345	261	174	112	0.300	0.392	0.569	0.293	0.195	0.321
CAPM_3MER	1500	1438	793	645	0.507	898	364	242	201	91	0.273	0.312	0.601	0.269	0.224	0.325
Discount_360D_BB	1500	1438	793	645	0.491	898	363	243	214	78	0.243	0.267	0.599	0.271	0.238	0.325
ROE_360D_BB	1500	1438	793	645	0.476	898	393	213	258	34	0.138	0.116	0.649	0.237	0.287	0.325
ROE_GARCH_CS	1500	1438	793	645	0.468	898	264	342	136	156	0.313	0.534	0.436	0.381	0.151	0.325
ROE_LTVol_BB	1500	1438	793	645	0.463	898	277	329	153	139	0.297	0.476	0.457	0.366	0.170	0.325
Discount_GARCH_BB	1500	1438	793	645	0.451	898	287	319	174	118	0.270	0.404	0.474	0.355	0.194	0.325
ROE_LTVol_CS	1500	1438	793	645	0.451	898	264	342	151	141	0.292	0.483	0.436	0.381	0.168	0.325
Discount_LTVol_BB	1500	1438	793	645	0.440	898	302	304	199	93	0.234	0.318	0.498	0.339	0.222	0.325
Discount_360D_CS	1500	1438	737	701	0.438	795	241	304	143	107	0.260	0.428	0.442	0.382	0.180	0.314

iv. Logistic Regression Results

Table C.6 GRF 1-year return predictions with logistic regression

GRF 1 Year Logit Predictions																	
	K	K1	ITD	ITU	IT	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
CAPM_1YER	1500	1249	958	291	-0.363	0.850	898	624	129	6	139	0.519	0.959	0.829	0.144	0.007	0.161
Discount_LTVol_CS	1500	1249	957	292	-0.211	0.840	892	749	0	143	0	NaN	0.000	1.000	0.000	0.160	0.160
Discount_360D_BB	1500	1249	958	291	-0.408	0.839	898	753	0	145	0	NaN	0.000	1.000	0.000	0.161	0.161
Discount_LTVol_BB	1500	1249	958	291	-0.306	0.839	898	753	0	145	0	NaN	0.000	1.000	0.000	0.161	0.161
Discount_GARCH_BB	1500	1249	958	291	-0.827	0.839	898	753	0	145	0	NaN	0.000	1.000	0.000	0.161	0.161
ROE_360D_BB	1500	1249	958	291	-1.634	0.839	898	753	0	145	0	NaN	0.000	1.000	0.000	0.161	0.161
ROE_LTVol_BB	1500	1249	958	291	-0.509	0.839	898	753	0	145	0	NaN	0.000	1.000	0.000	0.161	0.161
ROE_360D_CS	1500	1249	958	291	0.010	0.839	898	753	0	145	0	NaN	0.000	1.000	0.000	0.161	0.161
ROE_LTVol_CS	1500	1249	958	291	0.010	0.839	898	753	0	145	0	NaN	0.000	1.000	0.000	0.161	0.161
Discount_360D_CS	1500	1249	957	292	-1.904	0.825	795	656	0	139	0	NaN	0.000	1.000	0.000	0.175	0.175
ROE_GARCH_CS	1500	1249	958	291	0.008	0.821	898	737	16	145	0	0.000	0.000	0.979	0.018	0.161	0.161
Discount_GARCH_CS	1500	1249	957	292	-0.164	0.797	807	643	25	139	0	0.000	0.000	0.963	0.031	0.172	0.172
ROE_GARCH_BB	1500	1249	958	291	-1.256	0.763	887	677	67	143	0	0.000	0.000	0.910	0.076	0.161	0.161

Table C.7 GRF 6-month return predictions with logistic regression

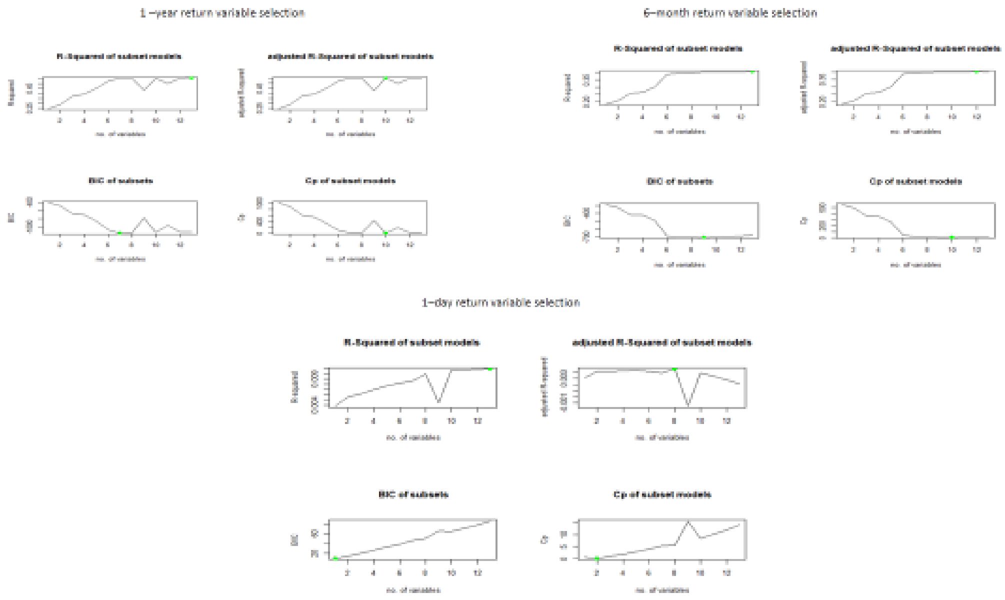
GRF 6 Months Logit Predictions																	
	K	K1	ITD	ITU	IT	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
CAPM_6MER	1500	1375	911	464	-0.156	0.817	898	580	98	66	154	0.611	0.700	0.855	0.109	0.073	0.245
ROE_GARCH_BB	1500	1375	900	475	2.316	0.758	887	672	0	215	0	NaN	0.000	1.000	0.000	0.242	0.242
Discount_360D_CS	1500	1375	828	547	-1.144	0.757	795	602	0	193	0	NaN	0.000	1.000	0.000	0.243	0.243
Discount_360D_BB	1500	1375	911	464	1.171	0.755	898	678	0	220	0	NaN	0.000	1.000	0.000	0.245	0.245
ROE_360D_BB	1500	1375	911	464	-3.351	0.755	898	678	0	220	0	NaN	0.000	1.000	0.000	0.245	0.245
ROE_360D_CS	1500	1375	911	464	-0.148	0.755	898	678	0	220	0	NaN	0.000	1.000	0.000	0.245	0.245
ROE_LTVol_CS	1500	1375	911	464	0.290	0.755	898	678	0	220	0	NaN	0.000	1.000	0.000	0.245	0.245
ROE_GARCH_CS	1500	1375	911	464	-0.066	0.755	898	678	0	220	0	NaN	0.000	1.000	0.000	0.245	0.245
ROE_LTVol_BB	1500	1375	911	464	0.222	0.751	898	674	4	220	0	0.000	0.000	0.994	0.004	0.245	0.245
Discount_LTVol_BB	1500	1375	911	464	0.093	0.735	898	659	19	219	1	0.050	0.005	0.972	0.021	0.244	0.245
Discount_LTVol_CS	1500	1375	905	470	-0.393	0.702	892	624	53	213	2	0.036	0.009	0.922	0.059	0.239	0.241
Discount_GARCH_CS	1500	1375	827	548	-0.333	0.663	807	530	84	188	5	0.056	0.026	0.863	0.104	0.233	0.239
Discount_GARCH_BB	1500	1375	911	464	0.181	0.663	898	585	93	210	10	0.097	0.045	0.863	0.104	0.234	0.245

Table C.8 GRF 3-month return prediction performance with logistic regression

GRF 3 Months Logit Predictions																	
	K	K1	ITD	ITU	IT	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_360D_CS	1500	1438	793	645	0.032	0.687	898	596	10	271	21	0.677	0.072	0.983	0.011	0.302	0.325
ROE_GARCH_CS	1500	1438	793	645	0.032	0.686	898	601	5	277	15	0.750	0.051	0.992	0.006	0.308	0.325
Discount_360D_BB	1500	1438	793	645	15.849	0.675	898	606	0	292	0	NaN	0.000	1.000	0.000	0.325	0.325
Discount_LTVol_BB	1500	1438	793	645	0.084	0.675	898	606	0	292	0	NaN	0.000	1.000	0.000	0.325	0.325
ROE_360D_BB	1500	1438	793	645	-1.546	0.675	898	606	0	292	0	NaN	0.000	1.000	0.000	0.325	0.325
ROE_LTVol_BB	1500	1438	793	645	1.300	0.675	898	606	0	292	0	NaN	0.000	1.000	0.000	0.325	0.325
ROE_LTVol_CS	1500	1438	793	645	0.032	0.675	898	606	0	292	0	NaN	0.000	1.000	0.000	0.325	0.325
Discount_LTVol_CS	1500	1438	787	651	-0.389	0.664	892	592	14	286	0	0.000	0.000	0.977	0.016	0.321	0.321
CAPM_3MER	1500	1438	793	645	-0.062	0.661	898	442	164	140	152	0.481	0.521	0.729	0.183	0.156	0.325
ROE_GARCH_BB	1500	1438	793	645	-2.072	0.661	887	586	20	281	0	0.000	0.000	0.967	0.023	0.317	0.317
Discount_GARCH_BB	1500	1438	793	645	0.218	0.646	898	574	32	286	6	0.158	0.021	0.947	0.036	0.318	0.325
Discount_GARCH_CS	1500	1438	731	707	-0.639	0.632	807	510	42	255	0	0.000	0.000	0.924	0.052	0.316	0.316
Discount_360D_CS	1500	1438	737	701	-0.488	0.626	795	482	63	234	16	0.203	0.064	0.884	0.079	0.294	0.314

v. Variable Selection

Figure C.3 GRF variable subset selection results



	Dependent variable:
	ER 1Y
CAPM	-10.413***
ROE_GARCH_CS	83.673**
Discount_360D_CS	-100.355**
Discount_GARCH_CS	0.545
Discount_LTVol_BB	-90.510***
ROE_360D_BB	5.068***
Discount_360D_BB	-11.496
Discount_GARCH_BB	17.300
ROE_LTVol_CS	5,366.876***
ROE_360D_CS	-5,703.379***
ROE_GARCH_BB	-1.403***
ROE_LTVol_BB	17.869**
Discount_LTVol_CS	98.483**
Constant	-3.755
Observations	1,500
Log Likelihood	-333.705
Akaike Inf. Crit.	695.411
Note:	*p<0.1; **p<0.05; ***p<0.01

	Dependent variable:
	ER 6M
CAPM	-13.495***
ROE_LTVol_CS	839.628***
ROE_GARCH_BB	0.475***
Discount_360D_BB	-11.545***
ROE_360D_CS	-968.803***

Discount_GARCH_CS	9.363***
Discount_LTVol_CS	-16.416***
ROE_GARCH_CS	48.952**
Constant	-3.429***
Observations	1,500
Log Likelihood	-772.556
Akaike Inf. Crit.	1,563.112
Note:	*p<0.1; **p<0.05; ***p<0.01

	Dependent variable:
	ER 1D
ROE_GARCH_CS	41.572*
ROE_LTVol_CS	-52.654**
ROE_LTVol_BB	1.318**
Discount_360D_CS	-0.949*
Constant	-0.151
Observations	1,500
Log Likelihood	-1,031.310
Akaike Inf. Crit.	2,072.619
Note:	*p<0.1; **p<0.05; ***p<0.01

D. BVC RESULTS

Structural model predictor variables

Figure D.1 BVC structural model variables

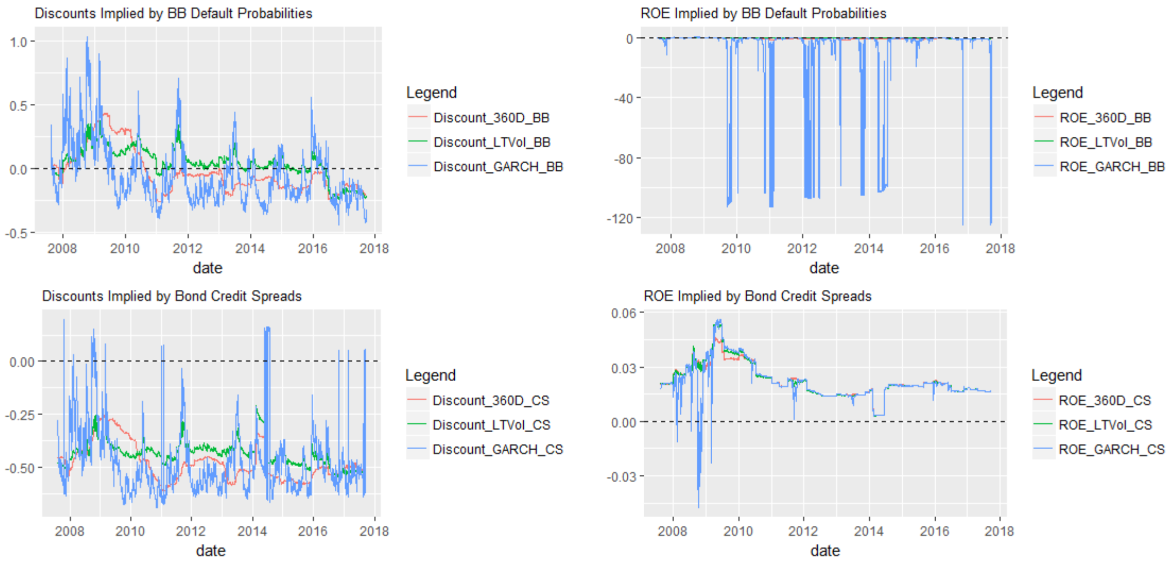


Table D.1BVC convergence to solutions in solving for asset value parameters

BVC TermCD Check										
	1	2	3	4	5	6	7	8	9	10
TermCD_M1_360D_BB	2528	0	0	0	0	0	0	0	0	0
TermCD_M1_LTVol_BB	2528	0	0	0	0	0	0	0	0	0
TermCD_M1_GARCH_BB	2528	0	0	0	0	0	0	0	0	0
TermCD_M2_360D_BB	2528	0	0	0	0	0	0	0	0	0
TermCD_M2_LTVol_BB	2528	0	0	0	0	0	0	0	0	0
TermCD_M2_GARCH_BB	2328	170	30	0	0	0	0	0	0	0
TermCD_M1_360D_CS	2528	0	0	0	0	0	0	0	0	0
TermCD_M1_LTVol_CS	2528	0	0	0	0	0	0	0	0	0
TermCD_M1_GARCH_CS	2476	2	50	0	0	0	0	0	0	0
TermCD_M2_360D_CS	2528	0	0	0	0	0	0	0	0	0
TermCD_M2_LTVol_CS	2528	0	0	0	0	0	0	0	0	0
TermCD_M2_GARCH_CS	2528	0	0	0	0	0	0	0	0	0

i. CAPM estimates of alpha and beta and equity volatility

Figure D.2 BVC estimates of alpha and beta

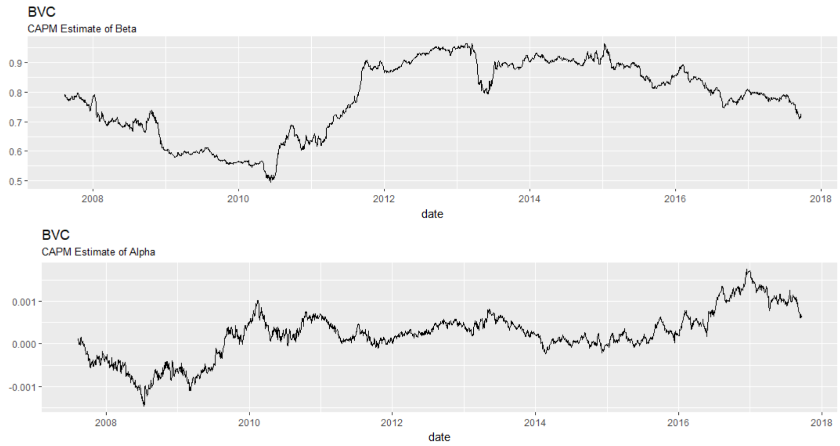
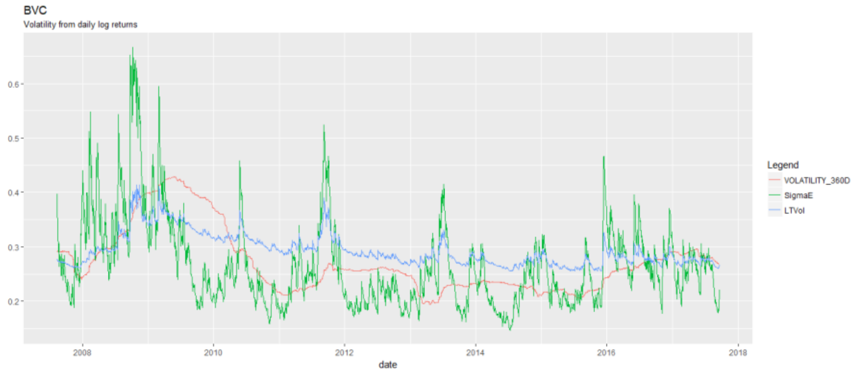


Figure D.3 BVC estimates of equity return volatility



ii. Indicator Results

Table D.2 BVC 1-year predictor performance on stand-alone basis

BVC 1 Year Indicator Performance													
	K	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_360D_CS	1500	0.914	1028	0	88	0	940	0.914	1.000	0.000	0.086	0.000	0.914
ROE_LTVol_CS	1500	0.914	1028	0	88	0	940	0.914	1.000	0.000	0.086	0.000	0.914
ROE_GARCH_CS	1500	0.914	1028	0	88	0	940	0.914	1.000	0.000	0.086	0.000	0.914
CAPM_1YER	1500	0.804	1028	6	82	120	820	0.909	0.872	0.068	0.080	0.117	0.914
Discount_LTVol_BB	1500	0.473	1028	48	40	502	438	0.916	0.466	0.545	0.039	0.488	0.914
ROE_LTVol_BB	1500	0.473	1028	48	40	502	438	0.916	0.466	0.545	0.039	0.488	0.914
ROE_GARCH_BB	1500	0.281	828	55	5	590	178	0.973	0.232	0.917	0.006	0.713	0.928
Discount_GARCH_BB	1500	0.270	1028	83	5	745	195	0.975	0.207	0.943	0.005	0.725	0.914
Discount_360D_BB	1500	0.086	1028	88	0	940	0	NaN	0.000	1.000	0.000	0.914	0.914
ROE_360D_BB	1500	0.086	1028	88	0	940	0	NaN	0.000	1.000	0.000	0.914	0.914
Discount_360D_CS	1500	0.086	1028	88	0	940	0	NaN	0.000	1.000	0.000	0.914	0.914
Discount_LTVol_CS	1500	0.086	1028	88	0	940	0	NaN	0.000	1.000	0.000	0.914	0.914
Discount_GARCH_CS	1500	0.075	976	73	0	903	0	NaN	0.000	1.000	0.000	0.925	0.925

Table D.3 BVC 6-month return prediction performance eon stand-alone basis

BVC 6 Months Indicator Performance													
	K	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_360D_CS	1500	0.711	1028	0	297	0	731	0.711	1.000	0.000	0.289	0.000	0.711
ROE_LTVol_CS	1500	0.711	1028	0	297	0	731	0.711	1.000	0.000	0.289	0.000	0.711
ROE_GARCH_CS	1500	0.711	1028	0	297	0	731	0.711	1.000	0.000	0.289	0.000	0.711
CAPM_6MER	1500	0.641	1028	37	260	109	622	0.705	0.851	0.125	0.253	0.106	0.711
Discount_LTVol_BB	1500	0.476	1028	154	143	396	335	0.701	0.458	0.519	0.139	0.385	0.711
ROE_LTVol_BB	1500	0.476	1028	154	143	396	335	0.701	0.458	0.519	0.139	0.385	0.711
ROE_GARCH_BB	1500	0.459	828	218	21	427	162	0.885	0.275	0.912	0.025	0.516	0.711
Discount_GARCH_BB	1500	0.443	1028	276	21	552	179	0.895	0.245	0.929	0.020	0.537	0.711
Discount_360D_BB	1500	0.289	1028	297	0	731	0	NaN	0.000	1.000	0.000	0.711	0.711
ROE_360D_BB	1500	0.289	1028	297	0	731	0	NaN	0.000	1.000	0.000	0.711	0.711
Discount_360D_CS	1500	0.289	1028	297	0	731	0	NaN	0.000	1.000	0.000	0.711	0.711
Discount_LTVol_CS	1500	0.289	1028	297	0	731	0	NaN	0.000	1.000	0.000	0.711	0.711
Discount_GARCH_CS	1500	0.276	976	269	0	707	0	NaN	0.000	1.000	0.000	0.724	0.724

iii. KNN Prediction Results

Table D.4 BVC 1-year return predictions with KNN (K=5)

BVC 1 Year KNN Prediction Performance 5															
	K	K1	ITD	ITU	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_GARCH_BB	1500	1249	338	911	0.879	828	1	59	41	727	0.925	0.947	0.017	0.071	0.050	0.928
Discount_GARCH_CS	1500	1249	350	899	0.872	976	3	70	55	848	0.924	0.939	0.041	0.072	0.056	0.925
CAPM_1YER	1500	1249	355	894	0.852	1028	3	85	67	873	0.911	0.929	0.034	0.083	0.065	0.914
Discount_GARCH_BB	1500	1249	355	894	0.851	1028	7	81	72	868	0.915	0.923	0.080	0.079	0.070	0.914
ROE_LTVol_CS	1500	1249	355	894	0.786	1028	0	88	132	808	0.902	0.860	0.000	0.086	0.128	0.914
ROE_GARCH_CS	1500	1249	355	894	0.731	1028	9	79	198	742	0.904	0.789	0.102	0.077	0.193	0.914
ROE_360D_CS	1500	1249	355	894	0.720	1028	12	76	212	728	0.905	0.774	0.136	0.074	0.206	0.914
Discount_360D_CS	1500	1249	355	894	0.716	1028	29	59	233	707	0.923	0.752	0.330	0.057	0.227	0.914
Discount_360D_BB	1500	1249	355	894	0.649	1028	42	46	315	625	0.931	0.665	0.477	0.045	0.306	0.914
Discount_LTVol_CS	1500	1249	355	894	0.648	1028	36	52	310	630	0.924	0.670	0.409	0.051	0.302	0.914
ROE_360D_BB	1500	1249	355	894	0.643	1028	38	50	317	623	0.926	0.663	0.432	0.049	0.308	0.914
ROE_LTVol_BB	1500	1249	355	894	0.638	1028	43	45	327	613	0.932	0.652	0.489	0.044	0.318	0.914
Discount_LTVol_BB	1500	1249	355	894	0.627	1028	43	45	338	602	0.930	0.640	0.489	0.044	0.329	0.914

Table D.5 BVC 6-month return predictions with KNN (K=5)

BVC 6 Months KNN Prediction Performance 5															
	K	K1	ITD	ITU	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
Discount_360D_CS	1500	1375	557	818	0.620	1028	113	184	207	524	0.740	0.717	0.380	0.179	0.201	0.711
Discount_GARCH_CS	1500	1375	554	821	0.614	976	58	211	166	541	0.719	0.765	0.216	0.216	0.170	0.724
Discount_GARCH_BB	1500	1375	557	818	0.600	1028	72	225	186	545	0.708	0.746	0.242	0.219	0.181	0.711
Discount_LTVol_BB	1500	1375	557	818	0.581	1028	99	198	233	498	0.716	0.681	0.333	0.193	0.227	0.711
ROE_GARCH_BB	1500	1375	556	819	0.580	828	73	166	182	407	0.710	0.691	0.305	0.200	0.220	0.711
CAPM_6MER	1500	1375	557	818	0.578	1028	69	228	206	525	0.697	0.718	0.232	0.222	0.200	0.711
ROE_LTVol_BB	1500	1375	557	818	0.571	1028	101	196	245	486	0.713	0.665	0.340	0.191	0.238	0.711
ROE_LTVol_CS	1500	1375	557	818	0.562	1028	32	265	185	546	0.673	0.747	0.108	0.258	0.180	0.711
Discount_360D_BB	1500	1375	557	818	0.543	1028	122	175	295	436	0.714	0.596	0.411	0.170	0.287	0.711
ROE_360D_BB	1500	1375	557	818	0.532	1028	128	169	312	419	0.713	0.573	0.431	0.164	0.304	0.711
ROE_GARCH_CS	1500	1375	557	818	0.517	1028	121	176	321	410	0.700	0.561	0.407	0.171	0.312	0.711
ROE_360D_CS	1500	1375	557	818	0.485	1028	68	229	300	431	0.653	0.590	0.229	0.223	0.292	0.711
Discount_LTVol_CS	1500	1375	557	818	0.444	1028	116	181	391	340	0.653	0.465	0.391	0.176	0.380	0.711

iv. Logistic Regression Results

Table D.6 BVC 1-year return predictions with logistic regression

BVC 1 Year Logistic Regression Results																	
	K	K1	ITD	ITU	IT	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_GARCH_BB	1500	1249	338	911	19.967	0.928	828	0	60	0	768	0.928	1.000	0.000	0.072	0.000	0.928
Discount_GARCH_CS	1500	1249	350	899	0.914	0.925	976	0	73	0	903	0.925	1.000	0.000	0.075	0.000	0.925
CAPM_1YER	1500	1249	355	894	-2.098	0.914	1028	0	88	0	940	0.914	1.000	0.000	0.086	0.000	0.914
Discount_360D_BB	1500	1249	355	894	-0.254	0.914	1028	0	88	0	940	0.914	1.000	0.000	0.086	0.000	0.914
Discount_GARCH_BB	1500	1249	355	894	3.227	0.914	1028	0	88	0	940	0.914	1.000	0.000	0.086	0.000	0.914
ROE_360D_BB	1500	1249	355	894	-1.158	0.914	1028	0	88	0	940	0.914	1.000	0.000	0.086	0.000	0.914
Discount_360D_CS	1500	1249	355	894	-0.604	0.914	1028	0	88	0	940	0.914	1.000	0.000	0.086	0.000	0.914
ROE_360D_CS	1500	1249	355	894	0.002	0.914	1028	0	88	0	940	0.914	1.000	0.000	0.086	0.000	0.914
ROE_LTVol_CS	1500	1249	355	894	0.005	0.914	1028	0	88	0	940	0.914	1.000	0.000	0.086	0.000	0.914
ROE_GARCH_CS	1500	1249	355	894	-0.028	0.914	1028	0	88	0	940	0.914	1.000	0.000	0.086	0.000	0.914
ROE_LTVol_BB	1500	1249	355	894	0.000	0.791	1028	37	51	164	776	0.938	0.826	0.420	0.050	0.160	0.914
Discount_LTVol_BB	1500	1249	355	894	0.009	0.747	1028	37	51	209	731	0.935	0.778	0.420	0.050	0.203	0.914
Discount_LTVol_CS	1500	1249	355	894	-0.457	0.695	1028	55	33	281	659	0.952	0.701	0.625	0.032	0.273	0.914

Table D.7 BVC 6-month return predictions with logistic regression

BVC 6 Months Logistic Regression Prediction Results																	
	K	K1	ITD	ITU	IT	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
Discount_GARCH_CS	1500	1375	554	821	-0.289	0.719	976	0	269	5	702	0.723	0.993	0.000	0.276	0.005	0.724
ROE_GARCH_BB	1500	1375	556	819	26.616	0.711	828	0	239	0	589	0.711	1.000	0.000	0.289	0.000	0.711
CAPM_6MER	1500	1375	557	818	-0.114	0.711	1028	0	297	0	731	0.711	1.000	0.000	0.289	0.000	0.711
Discount_360D_BB	1500	1375	557	818	-0.514	0.711	1028	0	297	0	731	0.711	1.000	0.000	0.289	0.000	0.711
ROE_360D_BB	1500	1375	557	818	-0.705	0.711	1028	0	297	0	731	0.711	1.000	0.000	0.289	0.000	0.711
Discount_360D_CS	1500	1375	557	818	-0.611	0.711	1028	0	297	0	731	0.711	1.000	0.000	0.289	0.000	0.711
ROE_360D_CS	1500	1375	557	818	0.042	0.711	1028	0	297	0	731	0.711	1.000	0.000	0.289	0.000	0.711
ROE_LTVol_CS	1500	1375	557	818	0.072	0.711	1028	0	297	0	731	0.711	1.000	0.000	0.289	0.000	0.711
ROE_GARCH_CS	1500	1375	557	818	0.005	0.711	1028	0	297	0	731	0.711	1.000	0.000	0.289	0.000	0.711
Discount_GARCH_BB	1500	1375	557	818	0.304	0.706	1028	0	297	5	726	0.710	0.993	0.000	0.289	0.005	0.711
Discount_LTVol_BB	1500	1375	557	818	-0.003	0.657	1028	55	242	111	620	0.719	0.848	0.185	0.235	0.108	0.711
ROE_LTVol_BB	1500	1375	557	818	-0.011	0.649	1028	62	235	126	605	0.720	0.828	0.209	0.229	0.123	0.711
Discount_LTVol_CS	1500	1375	557	818	-0.447	0.625	1028	70	227	159	572	0.716	0.782	0.236	0.221	0.155	0.711

v. Multiple predictor variables

Variable selection

Table D.8 BVC variable subset selection

BVC Variable Selection												
	1 Year				6 Months				1 Day			
	Adjr2	Cp	Bic	Aic	Adjr2	Cp	Bic	Aic	Adjr2	Cp	Bic	Aic
CAPM	■	■	■	■	■	■	■	■	■			
Discount_360D_BB	■	■	■	■	■	■	■	■				
Discount_LTVol_BB	■	■	■	■	■	■	■	■				
Discount_GARCH_BB	■	■	■	■	■	■	■	■	■			
ROE_360D_BB				■	■	■	■	■				
ROE_LTVol_BB				■	■	■	■	■	■			
ROE_GARCH_BB	■	■		■	■	■	■					
Discount_360D_CS	■	■	■	■	■	■	■	■				
Discount_LTVol_CS	■	■	■		■	■	■	■				
Discount_GARCH_CS	■	■	■	■	■	■	■	■	■			
ROE_360D_CS	■	■	■	■	■	■	■	■				
ROE_LTVol_CS	■	■	■	■	■	■	■	■				
ROE_GARCH_CS	■	■	■		■	■	■	■		■	■	
No. Predictor Variables	11	11	10	11	13	13	13	12	4	1	1	0
1Y Prediction Accuracy	0,31	0,31	0,47	0,27	0,29	0,29	0,29	0,40	0,65	0,91	0,91	
6M Prediction Accuracy	0,35	0,35	0,41	0,35	0,34	0,34	0,34	0,40	0,58	0,71	0,71	
1D Prediction Accuracy	0,50	0,50	0,48	0,50	0,51	0,51	0,51	0,48	0,47	0,49	0,49	

Bivariate predictor performance

Table D.9 BVC 1-year return top performing predictions including pairs of predictors

BVC 1 YearPrediction Results																		
Link	Predictor Variables		K	K1	ITD	ITU	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
LOGIT	ROE_GARCH_BB	Discount_GARCH_CS	1500	1249	337	912	0.933	817	0	55	0	762	0.933	1.000	0.000	0.067	0.000	0.933
LOGIT	CAPM_1YER	ROE_GARCH_BB	1500	1249	338	911	0.928	828	0	60	0	768	0.928	1.000	0.000	0.072	0.000	0.928
LOGIT		ROE_GARCH_BB	1500	1249	338	911	0.928	828	0	60	0	768	0.928	1.000	0.000	0.072	0.000	0.928
LOGIT		Discount_GARCH_CS	1500	1249	350	899	0.925	976	0	73	0	903	0.925	1.000	0.000	0.075	0.000	0.925
LOGIT		CAPM_1YER	1500	1249	355	894	0.914	1028	0	88	0	940	0.914	1.000	0.000	0.086	0.000	0.914
IND		ROE_360D_CS	1500				0.914	1028	0	88	0	940	0.914	1.000	0.000	0.086	0.000	0.914
IND		ROE_LTVol_CS	1500				0.914	1028	0	88	0	940	0.914	1.000	0.000	0.086	0.000	0.914
IND		ROE_GARCH_CS	1500				0.914	1028	0	88	0	940	0.914	1.000	0.000	0.086	0.000	0.914
KNN 5	Discount_GARCH_BB	ROE_GARCH_BB	1500	1249	338	911	0.886	828	2	58	36	732	0.927	0.953	0.033	0.070	0.043	0.928
KNN 5	CAPM_1YER	ROE_GARCH_BB	1500	1249	338	911	0.883	828	4	56	41	727	0.928	0.947	0.067	0.068	0.050	0.928
LOGIT	CAPM_1YER	ROE_LTVol_BB	1500	1249	355	894	0.882	1028	7	81	40	900	0.917	0.957	0.080	0.079	0.039	0.914
KNN 5		ROE_GARCH_BB	1500	1249	338	911	0.879	828	1	59	41	727	0.925	0.947	0.017	0.071	0.050	0.928
KNN 5		Discount_GARCH_CS	1500	1249	350	899	0.872	976	3	70	55	848	0.924	0.939	0.041	0.072	0.056	0.925
KNN 5	ROE_GARCH_BB	ROE_GARCH_CS	1500	1249	338	911	0.856	828	1	59	60	708	0.923	0.922	0.017	0.071	0.072	0.928
KNN 5	ROE_GARCH_BB	ROE_360D_CS	1500	1249	338	911	0.854	828	1	59	62	706	0.923	0.919	0.017	0.071	0.075	0.928
KNN 5	ROE_GARCH_BB	ROE_LTVol_CS	1500	1249	338	911	0.854	828	1	59	62	706	0.923	0.919	0.017	0.071	0.075	0.928
KNN 5		CAPM_1YER	1500	1249	355	894	0.852	1028	3	85	67	873	0.911	0.929	0.034	0.083	0.065	0.914
KNN 5	ROE_GARCH_BB	Discount_360D_CS	1500	1249	338	911	0.841	828	11	49	83	685	0.933	0.892	0.183	0.059	0.100	0.928
KNN 5	CAPM_1YER	Discount_GARCH_CS	1500	1249	350	899	0.826	976	1	72	98	805	0.918	0.891	0.014	0.074	0.100	0.925
KNN 5	CAPM_1YER	Discount_GARCH_BB	1500	1249	355	894	0.816	1028	3	85	104	836	0.908	0.889	0.034	0.083	0.101	0.914
IND		CAPM_1YER	1500				0.804	1028	6	82	120	820	0.909	0.872	0.068	0.080	0.117	0.914
LOGIT		ROE_LTVol_BB	1500	1249	355	894	0.791	1028	37	51	164	776	0.938	0.826	0.420	0.050	0.160	0.914

Debt variable efficiency

Table D.10 BVC 6-month return predictor variable efficiency

BVC Variable Efficency													
	K	K1	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
CS	1500	1375	0.413	1028	226	71	532	199	0.737	0.272	0.761	0.069	0.518	0.711
BB	1500	1375	0.534	1028	164	133	346	385	0.743	0.527	0.552	0.129	0.337	0.711
CAPM	1500	1375	0.700	1028	0	297	11	720	0.708	0.985	0.000	0.289	0.011	0.711
ALL	1500	1375	0.468	1028	143	154	393	338	0.687	0.462	0.481	0.150	0.382	0.711
BBCS	1500	1375	0.449	1028	185	112	454	277	0.712	0.379	0.623	0.109	0.442	0.711
CAPMCS	1500	1375	0.447	1028	179	118	450	281	0.704	0.384	0.603	0.115	0.438	0.711
CAPMBB	1500	1375	0.493	1028	133	164	357	374	0.695	0.512	0.448	0.160	0.347	0.711

E. CAPITEC RESULTS

i. Estimates of alpha and Beta and equity volatility

Figure E.1 Capitec estimates of alpha and beta

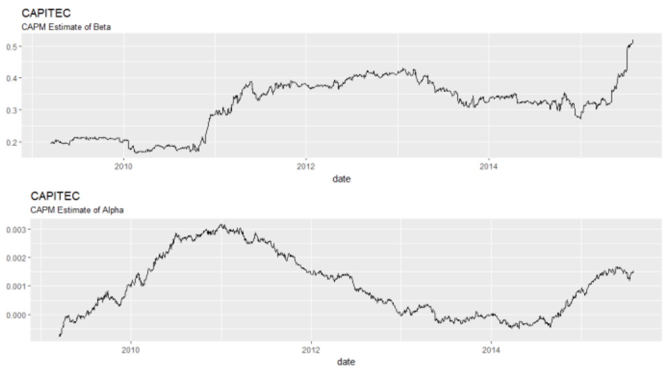
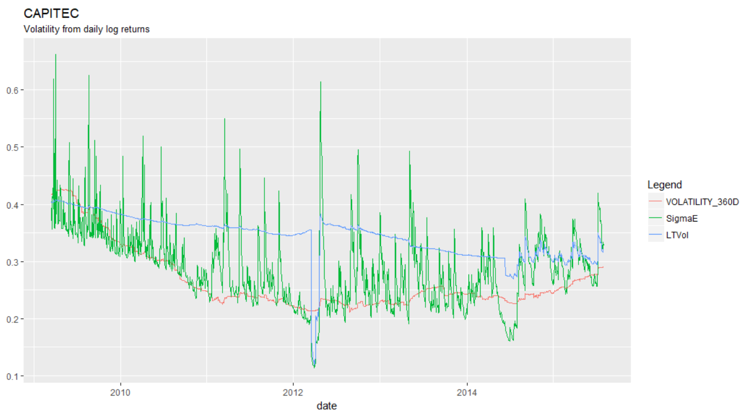


Figure E.2 Capitec estimates of equity return volatility



Structural model predictor variables

Figure E.3 Capitec structural model predictor variables

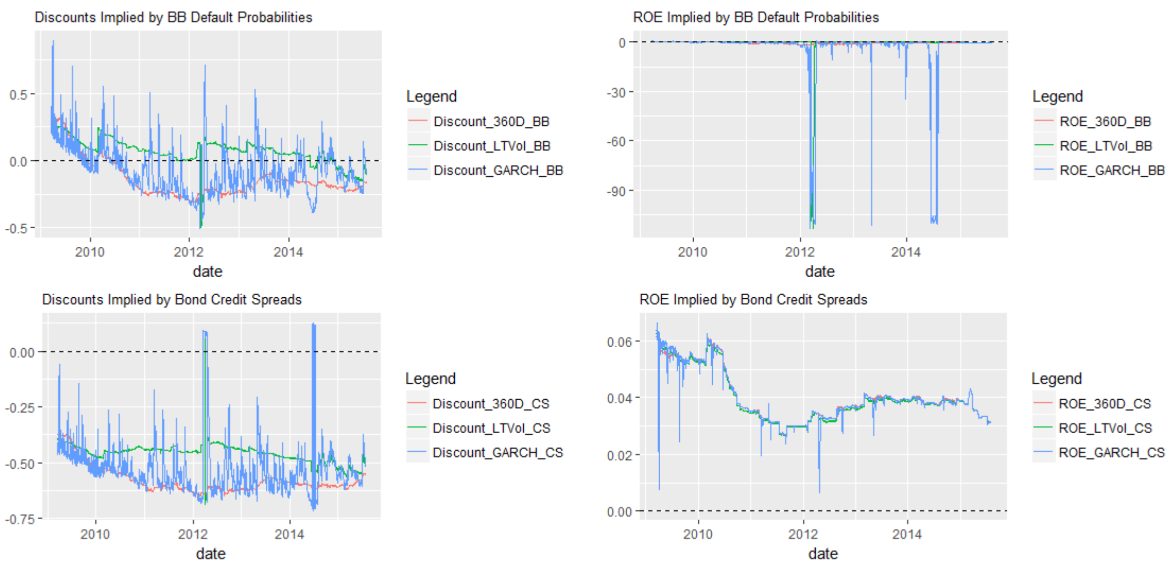


Table E.1 Capitec TermCD check to convergence in simultaneous equations

CAPITEC TermCD Check										
	1	2	3	4	5	6	7	8	9	10
TermCD_M1_360D_BB	1594	0	0	0	0	0	0	0	0	0
TermCD_M1_LTVol_BB	1589	0	5	0	0	0	0	0	0	0
TermCD_M1_GARCH_BB	1586	2	6	0	0	0	0	0	0	0
TermCD_M2_360D_BB	1594	0	0	0	0	0	0	0	0	0
TermCD_M2_LTVol_BB	1582	11	1	0	0	0	0	0	0	0
TermCD_M2_GARCH_BB	1544	47	3	0	0	0	0	0	0	0
TermCD_M1_360D_CS	1594	0	0	0	0	0	0	0	0	0
TermCD_M1_LTVol_CS	1583	0	11	0	0	0	0	0	0	0
TermCD_M1_GARCH_CS	1566	4	24	0	0	0	0	0	0	0
TermCD_M2_360D_CS	1594	0	0	0	0	0	0	0	0	0
TermCD_M2_LTVol_CS	1594	0	0	0	0	0	0	0	0	0
TermCD_M2_GARCH_CS	1594	0	0	0	0	0	0	0	0	0

ii. Indicator Results

Table E.2 Capitec 1-year return prediction performance evaluated on stand-alone prediction

CAPITEC 1 Year Indicator Results													
	K	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_360D_CS	1000	0.909	594	0	54	0	540	0.909	1.000	0.000	0.091	0.000	0.909
ROE_LTVol_CS	1000	0.909	594	0	54	0	540	0.909	1.000	0.000	0.091	0.000	0.909
ROE_GARCH_CS	1000	0.909	594	0	54	0	540	0.909	1.000	0.000	0.091	0.000	0.909
ROE_LTVol_BB	1000	0.603	582	18	24	207	333	0.933	0.617	0.429	0.041	0.356	0.928
Discount_LTVol_BB	1000	0.596	589	18	31	207	333	0.915	0.617	0.367	0.053	0.351	0.917
CAPM_1YER	1000	0.364	594	0	54	324	216	0.800	0.400	0.000	0.091	0.545	0.909
ROE_GARCH_BB	1000	0.290	544	29	4	382	129	0.970	0.252	0.879	0.007	0.702	0.939
Discount_GARCH_BB	1000	0.290	586	41	5	411	129	0.963	0.239	0.891	0.009	0.701	0.922
Discount_360D_BB	1000	0.091	594	54	0	540	0	NaN	0.000	1.000	0.000	0.909	0.909
ROE_360D_BB	1000	0.091	594	54	0	540	0	NaN	0.000	1.000	0.000	0.909	0.909
Discount_360D_CS	1000	0.091	594	54	0	540	0	NaN	0.000	1.000	0.000	0.909	0.909
Discount_LTVol_CS	1000	0.074	583	43	0	540	0	NaN	0.000	1.000	0.000	0.926	0.926
Discount_GARCH_CS	1000	0.062	566	35	0	531	0	NaN	0.000	1.000	0.000	0.938	0.938

Table E.3 Capitec 6-month return prediction performance evaluated on stand-alone basis

CAPITEC 6 Month Indicator Results													
	K	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_360D_CS	1000	0.810	594	0	113	0	481	0.810	1.000	0.000	0.190	0.000	0.810
ROE_LTVol_CS	1000	0.810	594	0	113	0	481	0.810	1.000	0.000	0.190	0.000	0.810
ROE_GARCH_CS	1000	0.810	594	0	113	0	481	0.810	1.000	0.000	0.190	0.000	0.810
ROE_LTVol_BB	1000	0.543	582	30	71	195	286	0.801	0.595	0.297	0.122	0.335	0.826
Discount_LTVol_BB	1000	0.537	589	30	78	195	286	0.786	0.595	0.278	0.132	0.331	0.817
CAPM_6MER	1000	0.502	594	42	71	225	256	0.783	0.532	0.372	0.120	0.379	0.810
ROE_GARCH_BB	1000	0.344	544	73	19	338	114	0.857	0.252	0.793	0.035	0.621	0.831
Discount_GARCH_BB	1000	0.340	586	85	20	367	114	0.851	0.237	0.810	0.034	0.626	0.821
Discount_360D_BB	1000	0.190	594	113	0	481	0	NaN	0.000	1.000	0.000	0.810	0.810
ROE_360D_BB	1000	0.190	594	113	0	481	0	NaN	0.000	1.000	0.000	0.810	0.810
Discount_360D_CS	1000	0.190	594	113	0	481	0	NaN	0.000	1.000	0.000	0.810	0.810
Discount_LTVol_CS	1000	0.175	583	102	0	481	0	NaN	0.000	1.000	0.000	0.825	0.825
Discount_GARCH_CS	1000	0.166	566	94	0	472	0	NaN	0.000	1.000	0.000	0.834	0.834

iii. KNN Prediction Results

Table E.4 Capitec 1-year return predictions with KNN (K=3)

CAPITEC 1 Year KNN 3 Prediction Performance																
	K	K1	ITD	ITU	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
ROE_LTVol_BB	1000	749	24	725	0.885	582	0	42	25	515	0.925	0.954	0.000	0.072	0.043	0.928
Discount_GARCH_BB	1000	749	24	725	0.828	586	1	45	56	484	0.915	0.896	0.022	0.077	0.096	0.922
Discount_LTVol_CS	1000	749	24	725	0.827	583	2	41	60	480	0.921	0.889	0.047	0.070	0.103	0.926
Discount_GARCH_CS	1000	749	24	725	0.822	566	3	32	69	462	0.935	0.870	0.086	0.057	0.122	0.938
ROE_GARCH_BB	1000	749	25	724	0.820	544	1	32	66	445	0.933	0.871	0.030	0.059	0.121	0.939
Discount_LTVol_BB	1000	749	24	725	0.812	589	18	31	80	460	0.937	0.852	0.367	0.053	0.136	0.917
ROE_GARCH_CS	1000	749	24	725	0.805	594	13	41	75	465	0.919	0.861	0.241	0.069	0.126	0.909
ROE_360D_CS	1000	749	24	725	0.803	594	11	43	74	466	0.916	0.863	0.204	0.072	0.125	0.909
CAPM_1YER	1000	749	24	725	0.790	594	0	54	71	469	0.897	0.869	0.000	0.091	0.120	0.909
ROE_360D_BB	1000	749	24	725	0.774	594	0	54	80	460	0.895	0.852	0.000	0.091	0.135	0.909
Discount_360D_CS	1000	749	24	725	0.766	594	3	51	88	452	0.899	0.837	0.056	0.086	0.148	0.909
Discount_360D_BB	1000	749	24	725	0.678	594	17	37	154	386	0.913	0.715	0.315	0.062	0.259	0.909
ROE_LTVol_CS	1000	749	24	725	0.660	594	8	46	156	384	0.893	0.711	0.148	0.077	0.263	0.909

Table E.5 Capitec 6-month return predictions with KNN (K=3)

CAPITEC 6 Months KNN 3 Prediction Performance																
	K	K1	ITD	ITU	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
Discount_LTVol_CS	1000	875	181	694	0.794	583	7	95	25	456	0.828	0.948	0.069	0.163	0.043	0.825
ROE_360D_BB	1000	875	174	701	0.724	594	20	93	71	410	0.815	0.852	0.177	0.157	0.120	0.810
Discount_GARCH_CS	1000	875	185	690	0.710	566	28	66	98	374	0.850	0.792	0.298	0.117	0.173	0.834
ROE_GARCH_BB	1000	875	185	690	0.704	544	22	70	91	361	0.838	0.799	0.239	0.129	0.167	0.831
ROE_LTVol_BB	1000	875	182	693	0.701	582	25	76	98	383	0.834	0.796	0.248	0.131	0.168	0.826
Discount_GARCH_BB	1000	875	180	695	0.698	586	15	90	87	394	0.814	0.819	0.143	0.154	0.148	0.821
CAPM_6MER	1000	875	174	701	0.650	594	3	110	98	383	0.777	0.796	0.027	0.185	0.165	0.810
Discount_360D_CS	1000	875	174	701	0.626	594	22	91	131	350	0.794	0.728	0.195	0.153	0.221	0.810
ROE_360D_CS	1000	875	174	701	0.614	594	19	94	135	346	0.786	0.719	0.168	0.158	0.227	0.810
Discount_LTVol_BB	1000	875	178	697	0.611	589	21	87	142	339	0.796	0.705	0.194	0.148	0.241	0.817
ROE_GARCH_CS	1000	875	174	701	0.572	594	32	81	173	308	0.792	0.640	0.283	0.136	0.291	0.810
Discount_360D_BB	1000	875	174	701	0.522	594	16	97	187	294	0.752	0.611	0.142	0.163	0.315	0.810
ROE_LTVol_CS	1000	875	174	701	0.517	594	32	81	206	275	0.772	0.572	0.283	0.136	0.347	0.810

iv. Logistic Regression Results

Table E.6 Capitec 1-year return predictions with logistic regression

CAPITEC 1 Year Logit Predictions																	
	K	K1	ITD	ITU	IT	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
Discount_GARCH_CS	1000	749	24	725	-0.709	0.938	566	0	35	0	531	0.938	1.000	0	0.062	0.000	0.938
ROE_LTVol_BB	1000	749	24	725	-0.133	0.928	582	0	42	0	540	0.928	1.000	0	0.072	0.000	0.928
Discount_LTVol_CS	1000	749	24	725	-0.658	0.926	583	0	43	0	540	0.926	1.000	0	0.074	0.000	0.926
Discount_GARCH_BB	1000	749	24	725	-0.447	0.922	586	0	46	0	540	0.922	1.000	0	0.078	0.000	0.922
Discount_LTVol_BB	1000	749	24	725	-0.120	0.917	589	0	49	0	540	0.917	1.000	0	0.083	0.000	0.917
ROE_GARCH_BB	1000	749	25	724	-47.809	0.914	544	0	33	14	497	0.938	0.973	0	0.061	0.026	0.939
CAPM_1YER	1000	749	24	725	1.452	0.909	594	0	54	0	540	0.909	1.000	0	0.091	0.000	0.909
Discount_360D_BB	1000	749	24	725	-0.452	0.909	594	0	54	0	540	0.909	1.000	0	0.091	0.000	0.909
ROE_360D_BB	1000	749	24	725	-2.639	0.909	594	0	54	0	540	0.909	1.000	0	0.091	0.000	0.909
Discount_360D_CS	1000	749	24	725	-0.685	0.909	594	0	54	0	540	0.909	1.000	0	0.091	0.000	0.909
ROE_360D_CS	1000	749	24	725	0.010	0.909	594	0	54	0	540	0.909	1.000	0	0.091	0.000	0.909
ROE_LTVol_CS	1000	749	24	725	0.009	0.909	594	0	54	0	540	0.909	1.000	0	0.091	0.000	0.909
ROE_GARCH_CS	1000	749	24	725	0.004	0.909	594	0	54	0	540	0.909	1.000	0	0.091	0.000	0.909

Table E.7 Capitec 6-month return predictions using logistic regression

CAPITEC 6 Months Logit Predictions																	
	K	K1	ITD	ITU	IT	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
Discount_GARCH_CS	1000	875	185	690	-1.080	0.834	566	0	94	0	472	0.834	1.000	0	0.166	0.000	0.834
ROE_LTVol_BB	1000	875	182	693	-2.223	0.826	582	0	101	0	481	0.826	1.000	0	0.174	0.000	0.826
Discount_LTVol_CS	1000	875	181	694	-0.382	0.825	583	0	102	0	481	0.825	1.000	0	0.175	0.000	0.825
Discount_GARCH_BB	1000	875	180	695	-0.553	0.821	586	0	105	0	481	0.821	1.000	0	0.179	0.000	0.821
Discount_LTVol_BB	1000	875	178	697	-0.394	0.817	589	0	108	0	481	0.817	1.000	0	0.183	0.000	0.817
ROE_GARCH_BB	1000	875	185	690	-29.942	0.812	544	0	92	10	442	0.828	0.978	0	0.169	0.018	0.831
CAPM_6MER	1000	875	174	701	1.322	0.810	594	0	113	0	481	0.810	1.000	0	0.190	0.000	0.810
Discount_360D_BB	1000	875	174	701	-0.310	0.810	594	0	113	0	481	0.810	1.000	0	0.190	0.000	0.810
ROE_360D_BB	1000	875	174	701	-1.610	0.810	594	0	113	0	481	0.810	1.000	0	0.190	0.000	0.810
Discount_360D_CS	1000	875	174	701	-0.641	0.810	594	0	113	0	481	0.810	1.000	0	0.190	0.000	0.810
ROE_360D_CS	1000	875	174	701	0.029	0.810	594	0	113	0	481	0.810	1.000	0	0.190	0.000	0.810
ROE_LTVol_CS	1000	875	174	701	0.028	0.810	594	0	113	0	481	0.810	1.000	0	0.190	0.000	0.810
ROE_GARCH_CS	1000	875	174	701	0.028	0.808	594	0	113	1	480	0.809	0.998	0	0.190	0.002	0.810

v. Multiple predictor variables

Variable selection

Table E.8 Capitec subset variable selection results

CAPITEC Variable Selection												
	1 Year				6 Months				1 Day			
	Adjr2	Cp	Bic	Aic	Adjr2	Cp	Bic	Aic	Adjr2	Cp	Bic	Aic
CAPM	■	■	■	■	■	■	■	■				
Discount_360D_BB	■	■	■	■	■	■	■	■			■	
Discount_LTVol_BB	■	■	■	■	■	■	■	■				■
Discount_GARCH_BB	■	■		■	■	■		■	■	■		■
ROE_360D_BB	■	■	■	■	■	■	■	■				
ROE_LTVol_BB	■			■					■	■		
ROE_GARCH_BB	■	■	■	■			■					■
Discount_360D_CS	■	■	■	■	■	■	■	■				■
Discount_LTVol_CS	■	■	■		■	■	■	■	■	■		
Discount_GARCH_CS	■	■							■	■		
ROE_360D_CS	■	■	■	■	■	■	■					
ROE_LTVol_CS	■	■		■	■	■			■			
ROE_GARCH_CS	■	■		■	■	■		■	■			
No. Predictor Variables	13	12	8	11	10	10	8	8	6	4	1	4
1Y Prediction Accuracy	0,54	0,67	0,37	0,58	0,49	0,49	0,37	0,37	0,86	0,94	0,91	0,93
6M Prediction Accuracy	0,50	0,34	0,43	0,67	0,53	0,53	0,43	0,39	0,67	0,83	0,81	0,81
1D Prediction Accuracy	0,46	0,48	0,49	0,49	0,49	0,49	0,49	0,50	0,48	0,50	0,49	0,51

Bivariate prediction performance

Table E.9 Capitec summary of 1-year return predictions including pairs of predictor variables

CAPITEC 1 Year Prediction Performance																		
Link	Predictor Variables		K	K1	ITD	ITU	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
LOGIT	ROE_GARCH_BB	Discount_360D_CS	1000	749	25	724	0.939	544	0	33	0	511	0.939	1.000	0	0.061	0.000	0.939
LOGIT	Discount_360D_BB	Discount_GARCH_CS	1000	749	24	725	0.938	566	0	35	0	531	0.938	1.000	0	0.062	0.000	0.938
LOGIT	Discount_LTVol_BB	Discount_GARCH_CS	1000	749	24	725	0.938	566	0	35	0	531	0.938	1.000	0	0.062	0.000	0.938
LOGIT		Discount_GARCH_CS	1000	749	24	725	0.938	566	0	35	0	531	0.938	1.000	0	0.062	0.000	0.938
KNN 3	ROE_LTVol_BB	Discount_LTVol_CS	1000	749	24	725	0.928	582	0	42	0	540	0.928	1.000	0.000	0.072	0.000	0.928
LOGIT		ROE_LTVol_BB	1000	749	24	725	0.928	582	0	42	0	540	0.928	1.000	0	0.072	0.000	0.928
LOGIT		Discount_LTVol_CS	1000	749	24	725	0.926	583	0	43	0	540	0.926	1.000	0	0.074	0.000	0.926
IND		ROE_360D_CS	1000				0.909	594	0	54	0	540	0.909	1.000	0.000	0.091	0.000	0.909
KNN 3	ROE_LTVol_BB	Discount_360D_CS	1000	749	24	725	0.904	582	10	32	24	516	0.942	0.956	0.238	0.055	0.041	0.928
KNN 3		ROE_LTVol_BB	1000	749	24	725	0.885	582	0	42	25	515	0.925	0.954	0.000	0.072	0.043	0.928
KNN 3		Discount_GARCH_BB	1000	749	24	725	0.828	586	1	45	56	484	0.915	0.896	0.022	0.077	0.096	0.922
KNN 3		Discount_LTVol_CS	1000	749	24	725	0.827	583	2	41	60	480	0.921	0.889	0.047	0.070	0.103	0.926
IND		ROE_LTVol_BB	1000				0.603	582	18	24	207	333	0.933	0.617	0.429	0.041	0.356	0.928

Table E.10 Capitec summary of 3-monthr return predictions including pairs of predictor variables

CAPITEC 3 Months Prediction Performance																		
Link	Predictor Variables		K	K1	ITD	ITU	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
LOGIT	Discount_GARCH_BB	Discount_GARCH_CS	1000	938	247	691	0.675	566	39	158	26	343	0.685	0.930	0.198	0.279	0.046	0.652
LOGIT	Discount_LTVol_BB	Discount_360D_CS	1000	938	254	684	0.669	589	16	195	0	378	0.660	1.000	0.076	0.331	0.000	0.642
KNN 3	ROE_360D_BB	Discount_LTVol_CS	1000	938	254	684	0.659	583	105	100	99	279	0.736	0.738	0.512	0.172	0.170	0.648
LOGIT	Discount_GARCH_CS	ROE_GARCH_CS	1000	938	247	691	0.654	566	1	196	0	369	0.653	1.000	0.005	0.346	0.000	0.652
LOGIT		Discount_GARCH_CS	1000	938	247	691	0.652	566	0	197	0	369	0.652	1.000	0.000	0.348	0.000	0.652
LOGIT		Discount_LTVol_CS	1000	938	254	684	0.648	583	0	205	0	378	0.648	1.000	0.000	0.352	0.000	0.648
LOGIT		CAPM_3MER	1000	938	254	684	0.636	594	0	216	0	378	0.636	1.000	0.000	0.364	0.000	0.636
IND		ROE_360D_CS	1000				0.636	594	0	216	0	378	0.636	1.000	0.000	0.364	0.000	0.636
IND		ROE_LTVol_CS	1000				0.636	594	0	216	0	378	0.636	1.000	0.000	0.364	0.000	0.636
KNN 3	Discount_360D_CS	ROE_GARCH_CS	1000	938	254	684	0.631	594	114	102	117	261	0.719	0.690	0.528	0.172	0.197	0.636
KNN 3		Discount_GARCH_CS	1000	938	247	691	0.625	566	45	152	60	309	0.670	0.837	0.228	0.269	0.106	0.652
KNN 3		Discount_GARCH_BB	1000	938	254	684	0.608	586	65	143	87	291	0.671	0.770	0.312	0.244	0.148	0.645
KNN 3		ROE_360D_BB	1000	938	254	684	0.593	594	119	97	145	233	0.706	0.616	0.551	0.163	0.244	0.636
IND		ROE_LTVol_BB	1000				0.486	582	65	139	160	218	0.611	0.577	0.319	0.239	0.275	0.649

Debt variable efficiency

Table E.11 Capitec 6-month return predictor variable efficiency

CAPITEC 6M Variable efficiency														
	K	K1	Acc	TSS	TN	FP	FN	TP	Prec	Sens	Spec	FPR	FNR	CB
CS	1000	875	0.285	594	80	33	392	89	0.730	0.185	0.708	0.056	0.660	0.81
BB	1000	875	0.256	594	53	60	382	99	0.623	0.206	0.469	0.101	0.643	0.81
CAPM	1000	875	0.310	594	104	9	401	80	0.899	0.166	0.920	0.015	0.675	0.81
ALL	1000	875	0.200	594	48	65	410	71	0.522	0.148	0.425	0.109	0.690	0.81
BBCS	1000	875	0.251	594	48	65	380	101	0.608	0.210	0.425	0.109	0.640	0.81
CAPMCS	1000	875	0.237	594	80	33	420	61	0.649	0.127	0.708	0.056	0.707	0.81
CAPMBB	1000	875	0.200	594	48	65	410	71	0.522	0.148	0.425	0.109	0.690	0.81

F. R-CODE

DataPrep

Data preparation function

```
DataPrep
function (HAR,Train)
{
# ==INPUTS ==
# HAR:              Input matrix of data
# Train:            The length of time-series observations to train GARCH and LOGIT

# Eventually aim to remove K such that Garch model beigns 1 step update when Credit Spreads h
ave values
# First Bond Issue Date on the Bond Info for particular company
K <- Train

# Declaring Variables as per BLOOMBERG CODES
Share_Price <- HAR[,"PX_LAST"]
JSE_Price <- HAR[,"JSE40_PX_LAST"]
RFR <- HAR[,"SWAP_RATE"]

# creating Daily Log Return Series
ELR_1D <- NULL
LR_1D <- NULL
MR_1D <- NULL
DMER <- NULL

# output vector for garch (1,1) volatility of daily share returns
SigmaE <- NULL
LTVol <- NULL

# Null Vectors for Excess Returns on The Market and Excess Returns Predicted by CAPM
Market_1YER <- NULL
Market_6MER <- NULL
Market_3MER <- NULL
Market_1MER <- NULL
Market_1WER <- NULL

Beta <- NULL                                # Estimated Beta Coefficient
Alpha <- NULL
### daily market and Share returns plus excess log return ###
for (i in 2:nrow(HAR))
         {
         ELR_1D[i] <- log((Share_Price[i])/(Share_Price[i-1]))-(1/252)*RFR[i]
         LR_1D[i] <- log((Share_Price[i])/(Share_Price[i-1]))
         MR_1D[i] <- log((JSE_Price[i])/(JSE_Price[i-1]))
         DMER[i] <- MR_1D[i]-(1/252)*RFR[i]             # Daily Market Excess Return
         }

### HISTORICAL MARKET RETURNS ###
for (i in 252:length(DMER))
         {
         Market_1YER[i] <- sum(DMER[(i-251):i])
         }

for (i in 126:length(DMER))
         {
         Market_6MER[i] <- sum(DMER[(i-125):i])
```

```
         }

for (i in 63:length(DMER))
         {
         Market_3MER[i] <- sum(DMER[(i-62):i])
         }

for (i in 22:length(DMER))
         {
         Market_1MER[i] <- sum(DMER[(i-21):i])
         }

for (i in 5:length(DMER))
         {
         Market_1WER[i] <- sum(DMER[(i-4):i])
         }

### GARCH ###

for (i in 1:(nrow(HAR)-K))
         {
         Vol_Model <- garchFit(formula=~garch(1,1),data=LR_1D[2:(K+i)],trace=FALSE)
         Vol_Pred1 <- predict(Vol_Model,n.ahead=100)[,"standardDeviation"]*sqrt(252)
         Vol_Pred2 <- predict(Vol_Model,n.ahead=1)[,"standardDeviation"]*sqrt(252)
         SigmaE[K+i] <- tail(Vol_Pred2,n=1)
         LTVol[K+i] <- tail(Vol_Pred1,n=1)
         }

### FORWARD REALIZED RETURNS ###

# Realized following day log return
ER_1D <- rep(0,nrow(HAR))
for (i in 1:(nrow(HAR)-1))
         {
         ER_1D[i] <-ELR_1D[i+1]
         }
# Realized following 1W log return
ER_1W <- rep(0,nrow(HAR))
for (i in 1:(nrow(HAR)-5))
         {
         ER_1W[i] <-sum(ELR_1D[(1+i):(5+i)])
         }
# Realized following 1M log return
ER_1M <- rep(0,nrow(HAR))
for (i in 1:(nrow(HAR)-21))
         {
         ER_1M[i] <-sum(ELR_1D[(1+i):(21+i)])
         }
# Realized following 3M log return
ER_3M <- rep(0,nrow(HAR))
for (i in 1:(nrow(HAR)-63))
         {
         ER_3M[i] <-sum(ELR_1D[(1+i):(63+i)])
         }

# Realized following 6M log return
ER_6M <- rep(0,nrow(HAR))
for (i in 1:(nrow(HAR)-126))
         {
         ER_6M[i] <-sum(ELR_1D[(1+i):(126+i)])
         }
# Realized following 1Y log return
```

```
ER_1Y <- rep(0,nrow(HAR))
for (i in 1:(nrow(HAR)-252))
         {
         ER_1Y[i] <-sum(ELR_1D[(1+i):(252+i)])
         }

### CAPM ESTIMATES ###

for (j in K:nrow(HAR))
         {
# linear regression to estimate beta using the daily returns on the market and the stock
         md1 <- lm(ELR_1D[(j-K+1):j]~DMER[(j-K+1):j])
         Beta[j] <- md1$coefficients[[2]]
         Alpha[j] <- md1$coefficients[[1]]
         }
out1 <- cbind(Alpha,Beta,Market_1YER,Market_6MER,Market_3MER,Market_1MER,Market_1WER,DMER)

### RECOMBINING RESULTS AND DATA ###
# Combining the realized future returns and original data
#HAR <- cbind(HAR,ELR_1D,LR_1D,MR_1D,SigmaE,LTVol,ER_1D,ER_1W,ER_1M,ER_3M,ER_6M,ER_1Y)
out2 <- cbind(ELR_1D,LR_1D,MR_1D,SigmaE,LTVol,ER_1D,ER_1W,ER_1M,ER_3M,ER_6M,ER_1Y)
out3 <- cbind(HAR,out2,out1)

# Chopping of the 1st K and last 252 observations for which we do not have realized returns
out4 <- out3[((K+1):(nrow(out3)-252)),]
out5 <- CAPM(out4)
return(out5)
}
```

CAPM

Capital Asset Pricing Model return estimates

```
CAPM
function (DataSet)
{
# prepare the alternative which is to compare against the CAPM model for future returns
# Alpha and Beta determined in Data Prep
# along with Historical Market Excess returns over different horizons
DS <- DataSet

# Null Vectors for Excess Returns on The Market and Excess Returns Predicted by CAPM
Market_1YER <- DS[,"Market_1YER"]
Market_6MER <- DS[,"Market_6MER"]
Market_3MER <- DS[,"Market_3MER"]
Market_1MER <- DS[,"Market_1MER"]
Market_1WER <-DS[,"Market_1WER"]
DMER <- DS[,"DMER"]
Beta <- DS[,"Beta"]          # Estimateed Beta Coefficient
Alpha <- DS[,"Alpha"]

CAPM_1YER <- NULL
CAPM_6MER <- NULL
CAPM_3MER <- NULL
CAPM_1MER <- NULL
CAPM_1WER <- NULL
CAPM_1DER <- NULL

AVM_1YER <- NULL
AVM_6MER <- NULL
```

```
AVM_3MER <- NULL
AVM_1MER <- NULL
AVM_1WER <- NULL
AVM_1DER <- NULL

# now create average Market Excess return over previous lag
for (i in 1:nrow(DS))
{
if (i > 252){AVM_1YER[i] <- mean(Market_1YER[i-252:i])
          AVM_6MER[i] <- mean(Market_6MER[i-252:i])
          AVM_3MER[i] <- mean(Market_3MER[i-252:i])}
else{AVM_1YER[i] <- mean(Market_1YER[1:i])
AVM_6MER[i] <- mean(Market_6MER[1:i])
AVM_3MER[i] <- mean(Market_3MER[1:i])}

if (i > 126){AVM_1MER[i] <- mean(Market_1MER[i-126:i])}
else{AVM_1MER[i] <- mean(Market_1MER[1:i])}

if (i > 22){AVM_1WER[i] <- mean(Market_1WER[i-22:i])}
else{AVM_1WER[i] <- mean(Market_1WER[1:i])}
if (i > 5){AVM_1DER[i] <- mean(DMER[i-5:i])}
else{AVM_1DER[i] <- mean(DMER[1:i])}
}

CAPM_1YER <- Beta*AVM_1YER+Alpha*252
CAPM_6MER <- Beta*AVM_6MER+Alpha*126
CAPM_3MER <- Beta*AVM_3MER+Alpha*63
CAPM_1MER <- Beta*AVM_1MER+Alpha*21
CAPM_1WER <- Beta*AVM_1WER+Alpha*5
CAPM_1DER <- Beta*AVM_1DER+Alpha

out1 <- cbind(CAPM_1YER,CAPM_6MER,CAPM_3MER,CAPM_1MER,CAPM_1WER,CAPM_1DER)
out2 <- cbind(DataSet,out1)
out2
}
```

Structural model functions

Black-Scholes Call option price

```
BSCall
function (V,F,Sigma,r,T)
{
# Function that prices BS price of a Call option
# ==INPUTS==
# V:      Current asset price
# F:      Strike price
# Sigma: Volatility of asset returns
# r:      risk-free rate
# T:      Time-to-maturity on option

# Outputs
# The value of a call C under the BS formula
d1 <-(1/(Sigma*sqrt(T)))*(log(V/F)+(r+0.5*(Sigma^2))*T)
d2 <- d1-Sigma*sqrt(T)
C <- V*pnorm(d1)-F*exp(-r*T)*pnorm(d2)
return(C)
}
```

Discount from PD

```
Merton.v1
function (E,SigmaE,r,T,F,PD)
{
# Function to solve for Merton parameters {V,Sigma} using PD as an input
# This version uses Risk-neutral inputs

# ==INPUTS==
# E:                        Equity price
# r:                        Risk-free rate
# T:                        Time to expiration
# F:                        Facevalue of debt outstanding
# SigmaE:                   Volatility of equity
# PD:                       Risk-neutral Probability of Default

# ==OUTPUTS==
# V:                        Implied Asset value
# SigmaV:                   Implied volatility of asset value
# L:                        Degree of leverage
# Et:                       Implied Equity value
# DER:                      Discount Excess Return

F <- F*exp(r*T)
# we now adapt the set of simultaneous equations to solve from a given PD
SimEqMerton <- function(x)
         {
         d1 <-(1/(x[2]*sqrt(T)))*(log(x[1]/F)+(r+0.5*(x[2]^2))*T)
         d2 <- d1-x[2]*sqrt(T)

         y <- numeric(2)
         y[1] <- (PD-pnorm(-d2))*10000 # to uses risk-neutral pd
         y[2] <- (SigmaE*E-(x[1]*x[2]*pnorm(d1))) # to solve for V and sigmaV
         y
         }

# Starting guesses for V and SigmaV as per Crosbie & Bohn
GuessV <- E+F
GuessSigmaV <- SigmaE*(E/(E+F))
xstart <- c(GuessV,GuessSigmaV)
solutions <- nleqslv(xstart,SimEqMerton,method="Newton",control=list(allowSingular=TRUE,maxit
=10000,ftol=0.001))
V_M1 <-solutions$x[1] # estimate of V
SigmaV_M1 <- solutions$x[2] # estimate of SigmaV
L_M1 <-  (exp(-r*T)*F)/V_M1
d1 <- (1/(SigmaV_M1*sqrt(T)))*(log(V_M1/F)+(r+0.5*(SigmaV_M1^2))*T)
d2 <- d1-SigmaV_M1*sqrt(T)
RR_M1 <- exp(-r*T)*(V_M1/F)*(pnorm(-d1)/pnorm(-d2))
Et_M1 <- BSCall(V_M1,F,SigmaV_M1,r,T)
Discount <- (Et_M1-E)/E
DER <- Discount-r                    # Discount Excess Return over risk-free rate
TermCD_M1 <- solutions$termcd[1]
PD_M1 <-pnorm(-d2)
output <- cbind(V_M1,SigmaV_M1,L_M1,RR_M1,Et_M1,Discount,DER,TermCD_M1,PD_M1)
output
}
```

ROE from PD

```
Merton.v2
function (E,SigmaE,r,T,F,PD)
{
# Function to solve for Merton parameters {V,Sigma,Mu} using PD,Et,Vol as an input
```

```
# ==INPUTS==
# E:                            Equity price
# r:                            Risk-free rate
# T:                            Time to expiration
# F:                            Facevalue of debt outstanding
# SigmaE:                       Volatility of equity
# PD:                           Risk-neutral Probability of Default

# ==OUTPUTS==
# V:                            Implied Asset value
# SigmaV:                       Implied volatility of asset value
# L:                            Degree of leverage
# Et:                           Implied Equity value
# Mu:                           Real-World Asset drift rate
# EAD:                          Excess Asset Drift Rate
# ROE:                          Excess Return on Equity implied by EAD

F <- F*exp(r*T)
# we now adapt the set of simultaneous equations to solve from a given PD
# x[3] is the Excess Asset Drift over risk-free rate
SimEqMerton2 <- function(x)
         {
         d1 <- (1/(x[2]*sqrt(T)))*(log(x[1]/F)+((x[3]+r)+0.5*(x[2]^2))*T)
         d2 <- d1-x[2]*sqrt(T)
         y <- numeric(3)
         y[1] <- (E-BSCall(x[1],F,x[2],x[3]+r,T))
         y[2] <- (SigmaE*E-x[1]*x[2]*pnorm(d1)) # to solve for V and sigmaV
         y[3] <- (PD-pnorm(-d2))*1000 # to uses risk-neutral pd
         y
         }
# Starting guesses for V and SigmaV as per Crosbie & Bohn
GuessV <- E+F
GuessSigmaV <- SigmaE*(E/(E+F))
GuessMu <- 0
xstart <- c(GuessV,GuessSigmaV,GuessMu)
solutions <- nleqslv(xstart,SimEqMerton2,method="Newton",control=list(allowSingular=TRUE,maxi
t=10000,ftol=0.001))
V_M2 <-solutions$x[1] # estimate of V
SigmaV_M2 <- solutions$x[2] # estimate of SigmaV
EAD <- solutions$x[3]
Mu <- EAD+r
L_M2 <-(exp(-r*T)*F)/V_M2
d1 <- (1/(SigmaV_M2*sqrt(T)))*(log(V_M2/F)+(Mu+0.5*(SigmaV_M2^2))*T)
d2 <- d1-SigmaV_M2*sqrt(T)
RR_M2 <- exp(-r*T)*(V_M2/F)*(pnorm(-d1)/pnorm(-d2))
Et_M2 <- BSCall(V_M2,F,SigmaV_M2,Mu,T)
TermCD_M2 <- solutions$termcd[1]
PD_M2 <-pnorm(-d2)
ROE <- EAD*(V_M2/E)
output <- cbind(V_M2,SigmaV_M2,L_M2,RR_M2,Et_M2,Mu,EAD,ROE,TermCD_M2,PD_M2)
output
}
```

PD models estimates applied to dataset

```
Merton.v4
function (DataMatrix,Maturity)
{
# ==INPUTS ==
# DataMatrix:                         Input matrix of data
```

```
# Maturity:                         Time to Maturity on debt/PD
T <- Maturity
output4 <- DataMatrix
#DataSet <- DataPrep(DataMatrix)

# Merton Results on Test Set
Share_Price <- DataMatrix[,"PX_LAST"]
risk_free_rate <- DataMatrix[,"SWAP_RATE"]
Vol_Class <- c("VOLATILITY_360D","LTVol","SigmaE")
Vol_Names <- c("360D","LTVol","GARCH")
CAPM_Returns <- c("CAPM_1YER","CAPM_6MER","CAPM_3MER","CAPM_1MER","CAPM_1WER","CAPM_1DER")
CAPM_names <- c("1Y","6M","3M","1M","1W","1D")
#dps <- HAR[,"TOT_DEBT_TO_COM_EQY"]
dps <- DataMatrix[,"TOTAL_DEBT_PER_SHARE"]

#PD_1YR <- DataMatrix[,"BB_1YR_DEFAULT_PROB"]
#PD_4YR <- DataMatrix[ ,"BB_4Y_DEFAULT_PROB"]
PD_5YR <- DataMatrix[,"BB_5Y_DEFAULT_PROB"]

# if statements to control PD Matching based on T inputed by user
if(T==1){
PD <- PD_1YR}
else if (T==4){
PD <- PD_4YR}
else{
PD <- PD_5YR}
for (j in 1:length(Vol_Class))
{
Equity_Vol <- DataMatrix[,Vol_Class[j]]
output1 <- Merton.v1(Share_Price[1],Equity_Vol[1],risk_free_rate[1],T,dps[1],PD[1])
output2 <- Merton.v2(Share_Price[1],Equity_Vol[1],risk_free_rate[1],T,dps[1],PD[1])

         for (i in 2:(nrow(DataMatrix)))
                  {
                  output1 <- rbind(output1,Merton.v1(Share_Price[i],Equity_Vol[i],risk_free_
rate[i],T,dps[i],PD[i]))
                  output2 <- rbind(output2,Merton.v2(Share_Price[i],Equity_Vol[i],risk_free_
rate[i],T,dps[i],PD[i]))
                  }
varnames1 <- NULL
varnames2 <- NULL
aaa <- colnames(output1)
bbb <- colnames(output2)

         for(z in 1:length(aaa))
                  {
                  varnames1[z] <- paste(aaa[z],Vol_Names[j],"BB",sep="_")
                  }

         for(u in 1:length(bbb))
                  {
                  varnames2[u] <- paste(bbb[u],Vol_Names[j],"BB",sep="_")
                  }

colnames(output1) <- varnames1
colnames(output2) <- varnames2
output3 <- cbind(output1,output2)
output4 <- cbind(output4,output3)
}
zz <- na.locf(output4,na.rm=TRUE)
return(zz)
}
```

Discount from Credit Spread (CS)

MV.1

```
function (E,SigmaE,r,T,F,CS)
{
# Function to solve for Merton parameters {V,Sigma} using Credit Spread as an input
# This version uses Risk-neutral inputs

# ==INPUTS==
# E:                    Equity price
# r:                    Risk-free rate
# T:                    Time to expiration
# F:                    Facevalue of debt outstanding
# SigmaE:               Volatility of equity
# CS:                   Credit Spread on Bond

# ==OUTPUTS==
# V:                    Implied Asset value
# SigmaV:               Implied volatility of asset value
# L:                    Degree of leverage
# Et:                   Implied Equity value
# DER:                  Discount Excess Return

#F <- F*exp(r*T)
# we now adapt the set of simultaneous equations to solve from a given PD
SimEqMerton <- function(x)
         {
         d1 <-(1/(x[2]*sqrt(T)))*(log(x[1]/F)+(r+0.5*(x[2]^2))*T)
         d2 <- d1-x[2]*sqrt(T)
         Dt <- x[1]-BSCall(x[1],F,x[2],r,T)
         S <- -((log(Dt/F))/T)-r
         y <- numeric(2)
         y[1] <- (CS-S)*1000 # to uses risk-neutral pd
         y[2] <- (SigmaE*E-(x[1]*x[2]*pnorm(d1))) # to solve for V and sigmaV
         y
         }

# Starting guesses for V and SigmaV as per Crosbie & Bohn
GuessV <- E+F
GuessSigmaV <- SigmaE*(E/(E+F))
xstart <- c(GuessV,GuessSigmaV)
solutions <- nleqslv(xstart,SimEqMerton,method="Newton",control=list(allowSingular=TRUE,maxit
=10000,ftol=0.001))
V_M1 <-solutions$x[1] # estimate of V
SigmaV_M1 <- solutions$x[2] # estimate of SigmaV
L_M1 <-  (exp(-r*T)*F)/V_M1
d1 <- (1/(SigmaV_M1*sqrt(T)))*(log(V_M1/F)+(r+0.5*(SigmaV_M1^2))*T)
d2 <- d1-SigmaV_M1*sqrt(T)
RR_M1 <- exp(-r*T)*(V_M1/F)*(pnorm(-d1)/pnorm(-d2))
Et_M1 <- BSCall(V_M1,F,SigmaV_M1,r,T)
Discount <- (Et_M1-E)/E
DER <- Discount-r                    # Discount Excess Return over risk-free rate
TermCD_M1 <- solutions$termcd[1]
PD_M1 <-pnorm(-d2)
Dt <- V_M1-BSCall(V_M1,F,SigmaV_M1,r,T)
S_M1 <- -((log(Dt/F))/T)-r
output <- cbind(V_M1,SigmaV_M1,L_M1,RR_M1,Et_M1,Discount,DER,TermCD_M1,PD_M1,Dt,S_M1)
output
```

```
}
```

ROE from credit Spreads

```
MV.2
function (E,SigmaE,r,T,F,CS)
{
# Function to solve for Merton parameters {V,Sigma,Mu} using PD,Et,Vol as an input

# ==INPUTS==
# E:                            Equity price
# r:                            Risk-free rate
# T:                            Time to expiration
# F:                            Facevalue of debt outstanding
# SigmaE:                       Volatility of equity
# CS:                           Credit Spread
# ==OUTPUTS==
# V:                            Implied Asset value
# SigmaV:                       Implied volatility of asset value
# L:                            Degree of leverage
# Et:                           Implied Equity value
# Mu:                           Real-World Asset drift rate
# EAD:                          Excess Asset Drift Rate
# ROE:                          Excess Return on Equity implied by EAD
F <- F*exp(r*T)
# we now adapt the set of simultaneous equations to solve from a given PD
# x[3] is the Excess Asset Drift over risk-free rate
SimEqMerton2 <- function(x)
          {
          d1 <- (1/(x[2]*sqrt(T)))*(log(x[1]/F)+((x[3]+r)+0.5*(x[2]^2))*T)
          d2 <- d1-x[2]*sqrt(T)
          Dt <- x[1]-BSCall(x[1],F,x[2],x[3]+r,T)
          S <- -((log(Dt/F))/T)-r
          y <- numeric(3)
          y[1] <- (E-BSCall(x[1],F,x[2],x[3]+r,T))
          y[2] <- (SigmaE*E-x[1]*x[2]*pnorm(d1)) # to solve for V and sigmaV
          y[3] <- (CS-S)*1000 # to uses risk-neutral pd
          y
          }
# Starting guesses for V and SigmaV as per Crosbie & Bohn
GuessV <- E+F
GuessSigmaV <- SigmaE*(E/(E+F))
GuessMu <- 0
xstart <- c(GuessV,GuessSigmaV,GuessMu)
solutions <- nleqslv(xstart,SimEqMerton2,method="Newton",control=list(allowSingular=TRUE,maxi
t=10000,ftol=0.001))
V_M2 <-solutions$x[1] # estimate of V
SigmaV_M2 <- solutions$x[2] # estimate of SigmaV
EAD <- solutions$x[3]
Mu <- EAD+r
L_M2 <-(exp(-r*T)*F)/V_M2
d1 <- (1/(SigmaV_M2*sqrt(T)))*(log(V_M2/F)+(Mu+0.5*(SigmaV_M2^2))*T)
d2 <- d1-SigmaV_M2*sqrt(T)
RR_M2 <- exp(-r*T)*(V_M2/F)*(pnorm(-d1)/pnorm(-d2))
Et_M2 <- BSCall(V_M2,F,SigmaV_M2,Mu,T)
TermCD_M2 <- solutions$termcd[1]
PD_M2 <-pnorm(-d2)
ROE <- EAD*(V_M2/E)
Dt_M2 <- V_M2-Et_M2
S_M2 <- -((log(Dt_M2/F))/T)-r
```

```
output <- cbind(V_M2,SigmaV_M2,L_M2,RR_M2,Et_M2,Mu,EAD,ROE,TermCD_M2,PD_M2,Dt_M2,S_M2)
output
}
```

Results from structural models with credit spread input

```
MV.4
function (DataMatrix,Maturity)
{
# ==INPUTS ==
# DataMatrix:                Input matrix of data
# Maturity:                  Time to Maturity on debt/PD
T <- Maturity
output4 <- DataMatrix
#DataSet <- DataPrep(DataMatrix)
# Merton Results on Test Set
Share_Price <- DataMatrix[,"PX_LAST"]
risk_free_rate <- DataMatrix[,"SWAP_RATE"]
Vol_Class <- c("VOLATILITY_360D","LTVol","SigmaE")
Vol_Names <- c("360D","LTVol","GARCH")
CAPM_Returns <- c("CAPM_1YER","CAPM_6MER","CAPM_3MER","CAPM_1MER","CAPM_1WER","CAPM_1DER")
CAPM_names <- c("1Y","6M","3M","1M","1W","1D")
#dps <- HAR[,"TOT_DEBT_TO_COM_EQY"]
dps <- DataMatrix[,"TOTAL_DEBT_PER_SHARE"]
CS <- DataMatrix[,"MC_SPREAD"]

for (j in 1:length(Vol_Class))
{
Equity_Vol <- DataMatrix[,Vol_Class[j]]
output1 <- MV.1(Share_Price[1],Equity_Vol[1],risk_free_rate[1],T,dps[1],CS[1])
output2 <- MV.2(Share_Price[1],Equity_Vol[1],risk_free_rate[1],T,dps[1],CS[1])

          for (i in 2:(nrow(DataMatrix)))
                    {
output1 <- rbind(output1,MV.1(Share_Price[i],Equity_Vol[i],risk_free_rate[i],T,dps[i],CS[i]))
output2 <- rbind(output2,MV.2(Share_Price[i],Equity_Vol[i],risk_free_rate[i],T,dps[i],CS[i]))
                    }
varnames1 <- NULL
varnames2 <- NULL
aaa <- colnames(output1)
bbb <- colnames(output2)

          for(z in 1:length(aaa))
                    {
                    varnames1[z] <- paste(aaa[z],Vol_Names[j],"CS",sep="_")
                    }

          for(u in 1:length(bbb))
                    {
                    varnames2[u] <- paste(bbb[u],Vol_Names[j],"CS",sep="_")
                    }

colnames(output1) <- varnames1
colnames(output2) <- varnames2
output3 <- cbind(output1,output2)
output4 <- cbind(output4,output3)
}
zz <- na.locf(output4,na.rm=TRUE)
return(zz)
}
```

Utility functions

Confustion matrix to results vector

```
Table_Results
function (PredictionTable)
{
PR <- PredictionTable
# how to deal with table when results contain predictions of only 1 Direction
TD <- dim(PR)
# 1 Row of only "Down" predictions + 2 columns
if(TD[1]==1 & TD[2]==2 & rownames(PR)[1]=="Down")
          {
          FP <-0
          TP <-0
          TN <-PR[1,1]
          FN <-PR[1,2]
          }
# 1 Row of only "Up" predictions+ 2 columns
else if(TD[1]==1 & TD[2]==2 & rownames(PR)[1]=="Up")
          {
          TN <- 0
          FN <- 0
          FP <-PR[1,1]
          TP <-PR[1,2]
          }
# 2 rows + 1 column of only "Down" realizations
else if(TD[1]==2 & TD[2]==1 & colnames(PR)[1]=="Down")
          {
          TN <- PR[1,1]
          FP <-PR[2,1]
          FN <- 0
          TP <- 0
          }
# 2 rows + 1 column of only "Up" realizations
else if(TD[1]==2 & TD[2]==1 & colnames(PR)[1]=="Up")
          {
          TN <- 0
          FP <- 0
          FN <- PR[1,1]
          TP <- PR[2,1]
          }
# 1 Row + 1 Column both "Up"
else if(TD[1]==1 & TD[2]==1 & colnames(PR)[1]=="Up" & rownames(PR)[1]=="Up")
          {
          TN <- 0
          FP <- 0
          FN <- 0
          TP <- PR[1,1]
          }
# 1 Row + 1 Column both "Down"
else if(TD[1]==1 & TD[2]==1 & colnames(PR)[1]=="Down" & rownames(PR)[1]=="Down")
          {
          TN <- PR[1,1]
          FP <- 0
          FN <- 0
          TP <- 0
          }
# 1 Row prediction only "Down" + 1 Column only "Up" realized
else if(TD[1]==1 & TD[2]==1 & colnames(PR)[1]=="Up" & rownames(PR)[1]=="Down")
          {
```

```
          TN <- 0
          FP <- 0
          FN <- PR[1,1]
          TP <- 0
          }
# 1 Row prediction only "Up" + 1 Column only "Down" realized
else if(TD[1]==1 & TD[2]==1 & rownames(PR)[1]=="Up" & colnames(PR)[1]=="Down")
          {
          TN <- 0
          FP <- PR[1,1]
          FN <- 0
          TP <- 0
          }
# Assuming 2 rows + 2 columns
else{
          FP <-PR[2,1]
          TP <-PR[2,2]
          TN <-PR[1,1]
          FN <-PR[1,2]
          }
# end of for loop and extracting TP, FP,TN,FN & Accuracy from prediction table Merton
# Prediction Robustness Metrics
Prec <- TP/(TP+FP)
Sens <- TP/(TP+FN)
Spec <- TN/(TN+FP)
FPR <- FP/(TP+TN+FN+FP)
FNR <- FN/(TP+TN+FN+FP)
# Class Balance: Proportion of upwards movements actually realized in test prediction set
CB <- (TP+FN)/(TP+TN+FN+FP)
TSS <- TP+FN+FP+TN
output <- cbind(TSS,TN,FP,FN,TP,Prec,Sens,Spec,FPR,FNR,CB)
return(output)
}
```

Predictor variable efficiency (PVE)

```
PVE
function (Results,Target_Variable,Train)
{
# Test function to see if transformation of variables to option is information productive
# Predictor Variable Efficiency (PVE)

K <- Train
nn <- nrow(Results)
Direction <- rep("Down",nn)
TV <- Results[,Target_Variable]
# Predictors from BB created Variables
P1 <- Results[,"PX_LAST"]
P2 <- Results[,"TOTAL_DEBT_PER_SHARE"]
P3 <- Results[,"SWAP_RATE"]
P4 <- Results[,"MC_SPREAD"]
P5 <- Results[,"BB_5Y_DEFAULT_PROB"]
P6 <- Results[,"JSE40_PX_LAST"]
Direction[TV>0] <-"Up"
if(Target_Variable=="ER_1Y"){KA <- 251}
else if(Target_Variable=="ER_6M"){KA <- 125}
else if(Target_Variable=="ER_3M"){KA <- 62}
else if(Target_Variable=="ER_1M"){KA <- 20}
else if(Target_Variable=="ER_1W"){KA <- 4}
else if(Target_Variable=="ER_1D"){KA <- 0}
```

```
K1 <- K-KA
#Credit Spread Variables
DataMat <- cbind(Results,Direction,P1,P2,P3,P4)
glm.pred <- rep("Down",(nn-K))
for (i in K:(nn-1))
         {
         df_train <-DataMat[1:(i-KA),]
         df_test <- DataMat[(i+1),]
         glm.fit <- glm(Direction~P1+P2+P3+P4,family=binomial,data=df_train)
         glm.probs <-predict(glm.fit,df_test,type="response")
         if(glm.probs>.5){glm.pred[i-K+1]<-"Up"}
         }
Direct <- DataMat[((K+1):nn),"Direction"]
t1<-table(glm.pred,Direct)
Tc <- mean(glm.pred==Direct)
Acc <- Tc
TR <- Table_Results(t1)
output <- cbind(K,K1,Acc,TR)
rownames(output)<-"CS"
#BB Variables
DataMat <- cbind(Results,Direction,P1,P2,P3,P5)
glm.pred <- rep("Down",(nn-K))
for (i in K:(nn-1))
         {
         df_train <-DataMat[1:(i-KA),]
         df_test <- DataMat[(i+1),]
         glm.fit <- glm(Direction~P1+P2+P3+P5,family=binomial,data=df_train)
         glm.probs <-predict(glm.fit,df_test,type="response")
         if(glm.probs>.5){glm.pred[i-K+1]<-"Up"}
         }
Direct <- DataMat[((K+1):nn),"Direction"]
t1<-table(glm.pred,Direct)
Tc <- mean(glm.pred==Direct)
Acc <- Tc
TR <- Table_Results(t1)
output1 <- cbind(K,K1,Acc,TR)
rownames(output1)<-"BB"
#CAPM
DataMat <- cbind(Results,Direction,P1,P6)
glm.pred <- rep("Down",(nn-K))
for (i in K:(nn-1))
         {
         df_train <-DataMat[1:(i-KA),]
         df_test <- DataMat[(i+1),]
         glm.fit <- glm(Direction~P1+P6,family=binomial,data=df_train)
         glm.probs <-predict(glm.fit,df_test,type="response")
         if(glm.probs>.5){glm.pred[i-K+1]<-"Up"}
         }
Direct <- DataMat[((K+1):nn),"Direction"]
t1<-table(glm.pred,Direct)
Tc <- mean(glm.pred==Direct)
Acc <- Tc
TR <- Table_Results(t1)
output2 <- cbind(K,K1,Acc,TR)
rownames(output2)<-"CAPM"
#ALL
DataMat <- cbind(Results,Direction,P1,P2,P3,P4,P5,P6)
glm.pred <- rep("Down",(nn-K))

for (i in K:(nn-1))
         {
         df_train <-DataMat[1:(i-KA),]
```

```
          df_test <- DataMat[(i+1),]
          glm.fit <- glm(Direction~P1+P2+P3+P4+P5+P6,family=binomial,data=df_train)
          glm.probs <-predict(glm.fit,df_test,type="response")
          if(glm.probs>.5){glm.pred[i-K+1]<-"Up"}
          }
Direct <- DataMat[((K+1):nn),"Direction"]
t1<-table(glm.pred,Direct)
Tc <- mean(glm.pred==Direct)
Acc <- Tc
TR <- Table_Results(t1)
output3 <- cbind(K,K1,Acc,TR)
rownames(output3)<-"ALL"

# BB+CS
DataMat <- cbind(Results,Direction,P1,P2,P3,P4,P5)
glm.pred <- rep("Down",(nn-K))

for (i in K:(nn-1))
          {
          df_train <-DataMat[1:(i-KA),]
          df_test <- DataMat[(i+1),]
          glm.fit <- glm(Direction~P1+P2+P3+P4+P5,family=binomial,data=df_train)
          glm.probs <-predict(glm.fit,df_test,type="response")
          if(glm.probs>.5){glm.pred[i-K+1]<-"Up"}
          }

Direct <- DataMat[((K+1):nn),"Direction"]
t1<-table(glm.pred,Direct)
Tc <- mean(glm.pred==Direct)
Acc <- Tc
TR <- Table_Results(t1)
output4 <- cbind(K,K1,Acc,TR)
rownames(output4)<-"BBCS"
# CAPM+CS
DataMat <- cbind(Results,Direction,P1,P2,P3,P4,P6)
glm.pred <- rep("Down",(nn-K))

for (i in K:(nn-1))
          {
          df_train <-DataMat[1:(i-KA),]
          df_test <- DataMat[(i+1),]
          glm.fit <- glm(Direction~P1+P2+P3+P4+P6,family=binomial,data=df_train)
          glm.probs <-predict(glm.fit,df_test,type="response")
          if(glm.probs>.5){glm.pred[i-K+1]<-"Up"}
          }

Direct <- DataMat[((K+1):nn),"Direction"]
t1<-table(glm.pred,Direct)
Tc <- mean(glm.pred==Direct)
Acc <- Tc
TR <- Table_Results(t1)
output5 <- cbind(K,K1,Acc,TR)
rownames(output5)<-"CAPMCS"
# CAPM+BB
DataMat <- cbind(Results,Direction,P1,P2,P3,P5,P6)
glm.pred <- rep("Down",(nn-K))

for (i in K:(nn-1))
          {
          df_train <-DataMat[1:(i-KA),]
          df_test <- DataMat[(i+1),]
          glm.fit <- glm(Direction~P1+P2+P3+P5+P6,family=binomial,data=df_train)
```

```
          glm.probs <-predict(glm.fit,df_test,type="response")
          if(glm.probs>.5){glm.pred[i-K+1]<-"Up"}
          }
Direct <- DataMat[((K+1):nn),"Direction"]
t1<-table(glm.pred,Direct)
Tc <- mean(glm.pred==Direct)
Acc <- Tc
TR <- Table_Results(t1)
output6 <- cbind(K,K1,Acc,TR)
rownames(output6)<-"CAPMBB"
output7 <- rbind(output,output1,output2,output3,output4,output5,output6)
output8 <- round(output7,digits=3)
return(output8)
}
```

TermCD Check for convergence to solutions in solving for structural model parameters

```
TermCD_Check
function (DataSet)
{
# function follows Merton.v9 currently
# First goal is to check TermCD on all relevant ROE and Discount From all PDs
DS <- DataSet
Models <- c("TermCD_M1_360D_BB","TermCD_M1_LTVol_BB","TermCD_M1_GARCH_BB",
          "TermCD_M2_360D_BB","TermCD_M2_LTVol_BB","TermCD_M2_GARCH_BB",
          "TermCD_M1_360D_CS","TermCD_M1_LTVol_CS","TermCD_M1_GARCH_CS",
          "TermCD_M2_360D_CS","TermCD_M2_LTVol_CS","TermCD_M2_GARCH_CS")

# create 1st row vector for Model 1
A <- table(DS[,Models[1]])
AA <- dimnames(A)
output1 <- rep(0,10)
for (k in 1:length(AA[[1]]))
          {
          output1[eval(parse(text=AA[[1]][k]))] <-A[k]
          }
# Reiterate over all models binding the row output each time
for (j in 2:length(Models))
{
C <- table(DS[,Models[j]])
CC <- dimnames(C)
output3 <- rep(0,10)
for (i in 1:length(CC[[1]]))
          {
          output3[eval(parse(text=CC[[1]][i]))] <-C[i]
          }
output1 <- rbind(output1,output3)
}
rownames(output1) <- Models
colnames(output1) <- c(1:10)
return(output1)
}
```

Stand-alone performance evaluation

Indicator function

```
Indicator
```

```
function (DS,x1,x2,K)
{
# == INPUTS == #
# DS:     Data Set
# x1:     Predictions
# x2:     Actual
Prefix <- substr(x1,start=1,stop=3)
NC <- nchar(x1)
if(Prefix=="ROE")
{Label <- "TermCD_M2"
Suffix <- substr(x1,start=5,stop=NC)}
else if(Prefix=="Dis")
{Label <-"TermCD_M1"
Suffix <- substr(x1,start=10,stop=NC)}
if(Prefix=="CAP"){DM <- DS}
else{TermCD <- paste(Label,Suffix,sep="_")
DM <- DS[DS[,TermCD]==1,]}
NT <- nrow(DM)
DataMat <- DM[((K+1):NT),]
x <- DataMat[,x1]
y <- DataMat[,x2]
# Indicator of how well predictions split the positive and negative returns
N <- length(x) # Total number of predictions made

Actual_Indicator <- rep("Down",N)
Actual_Indicator[y>0] <- "Up"
Pred_Indicator <- rep("Down",N)
Pred_Indicator[x>0] <- "Up"

PR <- table(Pred_Indicator,Actual_Indicator)
Acc <- mean(Pred_Indicator==Actual_Indicator)
TR <- Table_Results(PR)

AA <- cbind(K,Acc,TR)
AA
}
```

Indicator Results

```
Indicator_List
function (Results,K)
{
# function to list the indicator around fixed value of zero for all models against different
return horizons
Target_Variable <- c("ER_1Y","ER_6M","ER_3M","ER_1M","ER_1W","ER_1D")
CAPM <- c("CAPM_1YER","CAPM_6MER","CAPM_3MER","CAPM_1MER","CAPM_1WER","CAPM_1DER")
Models <- c("CAPM",
         "Discount_360D_BB","Discount_LTVol_BB","Discount_GARCH_BB",
         "ROE_360D_BB","ROE_LTVol_BB","ROE_GARCH_BB",
         "Discount_360D_CS","Discount_LTVol_CS","Discount_GARCH_CS",
         "ROE_360D_CS","ROE_LTVol_CS","ROE_GARCH_CS")

ExRet_1Y <- NULL
ExRet_6M <- NULL
ExRet_3M <- NULL
ExRet_1M <- NULL
ExRet_1W <- NULL
ExRet_1D <- NULL
ListNames <- NULL

## For loop over each Target Variable
```

```
for (m in 1:length(Target_Variable))
{
# initial run over CAPM
TVar <- Target_Variable[m]
Pr <- CAPM[m]
output <- Indicator(Results,Pr,TVar,K)
rownames(output) <- CAPM[m]
for (h in 2:length(Models))
{
Pr <- Models[h]
output1 <- Indicator(Results,Pr,TVar,K)
rownames(output1) <- Models[h]
output <- rbind(output,output1)
}

output2<- output[sort.list(output[,2],decreasing=TRUE),]
output3<- round(output2,digits=3)
if(m==1){ExRet_1Y <-output3}
else if(m==2){ExRet_6M <- output3}
else if(m==3){ExRet_3M <- output3}
else if(m==4){ExRet_1M <- output3}
else if(m==5){ExRet_1W <- output3}
else{ExRet_1D <- output3}
ListNames[m] <- paste(Target_Variable[m],"Indicator")
}
outlist <- list(ExRet_1Y,ExRet_6M,ExRet_3M,ExRet_1M,ExRet_1W,ExRet_1D)
names(outlist) <- ListNames
return(outlist)
}
```

Logistic regression functions

Logistic regression predictions for single predictor variable

```
Logit
function (DS,Target_Variable,Predictor,Train)
{
NC <- nchar(Predictor)
Prefix <- substr(Predictor,start=1,stop=3)
if(Prefix=="ROE")
{Label <- "TermCD_M2"
Suffix <- substr(Predictor,start=5,stop=NC)}
else if(Prefix=="Dis")
{Label <-"TermCD_M1"
Suffix <- substr(Predictor,start=10,stop=NC)}
if(Prefix=="CAP"){Results <- DS}
else{
TermCD <- paste(Label,Suffix,sep="_")
Results <- DS[DS[,TermCD]==1,]}

K <- Train
nn <- nrow(Results)
Direction <- rep("Down",nn)
TV <- Results[,Target_Variable]
Pred <- Results[,Predictor]
Direction[TV>0] <-"Up"

DataMat <- cbind(Results,Direction,Pred)
glm.pred <- rep("Down",(nn-K))
B0 <- NULL
```

```
B1 <- NULL
Threshold <- NULL

if(Target_Variable=="ER_1Y"){KA <- 251}
else if(Target_Variable == "ER_6M"){KA <- 125}
else if(Target_Variable == "ER_3M"){KA <- 62}
else if(Target_Variable == "ER_1M"){KA <- 20}
else if(Target_Variable == "ER_1W"){KA <- 4}
else if(Target_Variable == "ER_1D"){KA <- 0}

# in 1step forward update fashion
Signal <- rep(0,nn)
Signal[Results[,Target_Variable]>0] <- 1
ss <- length(Signal)
TD <- rep(0,ss)
TU <- rep(0,ss)

for (i in 1:ss)
         {
         TU[i] <- sum(Signal[1:i])
         TD[i] <- i-TU[i]
         }

for (i in K:(nn-1))
         {
         df_train <-DataMat[(1:(i-KA)),]
         df_test <- DataMat[(i+1),]
         glm.fit <- glm(Direction~Pred,family=binomial,data=df_train)
         B0[i] <- glm.fit$coefficients[[1]]
         B1[i] <- glm.fit$coefficients[[2]]
         Threshold[i] <- (log(1)-B0[i])/B1[i]
         glm.probs <-predict(glm.fit,df_test,type="response")
         if(glm.probs>.5){glm.pred[i-K+1]<-"Up"}
         }

Direct <- DataMat[((K+1):nn),"Direction"]
t1<-table(glm.pred,Direct)
Tc <- mean(glm.pred==Direct)
outlist <- list("Prediction Table"=t1,"Total Correct"=Tc,"Threshold"=Threshold,"TU"=TU,"TD"=T
D)
return(outlist)
}
```

Logistic Regression results for single predictor variable

```
Logit_Results
function (Results,MinTrain)
{
# function recursivley draws 1-step update predictions from SINGLE starting points
# See V13 for varying start points

# Results:                    Full set Returns and Return Predictors
# MinTrain:                   The Minimum observations in the initial training set

K <- MinTrain
Target_Variable <- c("ER_1Y","ER_6M","ER_3M","ER_1M","ER_1W","ER_1D")
CAPM <- c("CAPM_1YER","CAPM_6MER","CAPM_3MER","CAPM_1MER","CAPM_1WER","CAPM_1DER")
Models <- c("CAPM",
         "Discount_360D_BB","Discount_LTVol_BB","Discount_GARCH_BB",
         "ROE_360D_BB","ROE_LTVol_BB","ROE_GARCH_BB",
```

```
          "Discount_360D_CS","Discount_LTVol_CS","Discount_GARCH_CS",
          "ROE_360D_CS","ROE_LTVol_CS","ROE_GARCH_CS")

ExRet_1Y <- NULL
ExRet_6M <- NULL
ExRet_3M <- NULL
ExRet_1M <- NULL
ExRet_1W <- NULL
ExRet_1D <- NULL
ListNames <- NULL

## For loop over each Target Variable
for (m in 1:length(Target_Variable))
{
# initial run over CAPM
TVar <- Target_Variable[m]
if(TVar=="ER_1Y"){KA <- 251}
else if(TVar=="ER_6M"){KA <- 125}
else if(TVar=="ER_3M"){KA <- 62}
else if(TVar=="ER_1M"){KA <- 20}
else if(TVar=="ER_1W"){KA <- 4}
else if(TVar=="ER_1D"){KA <- 0}
K1 <- K-KA
Pr <- CAPM[m]

ff <- Logit(Results,TVar,Pr,K)
PR <- ff[["Prediction Table"]]
Acc <-ff[["Total Correct"]]
Threshold <- ff[["Threshold"]]
TrainD <- ff[["TD"]]
TrainU <- ff[["TU"]]
# how to deal with table when results contain predictions of only 1 Direction
TR <- Table_Results(PR)
IT <- Threshold[K]
ITU <- TrainU[K1]
ITD <- TrainD[K1]

output <- cbind(K,K1,ITD,ITU,IT,Acc,TR)
rownames(output) <- CAPM[m]

for (h in 2:length(Models))
          {
          Pr <- Models[h]
          ff <- Logit(Results,TVar,Pr,K)
          PR <- ff[["Prediction Table"]]
          Acc <-ff[["Total Correct"]]
          Threshold <- ff[["Threshold"]]
          TrainD <- ff[["TD"]]
          TrainU <- ff[["TU"]]

          IT <- Threshold[K] # Initial Threshold
          ITU <- TrainU[K1]  # initial upward obs in train set
          ITD <- TrainD[K1]  # initial D obs in train set
          TR <- Table_Results(PR)
          output1 <- cbind(K,K1,ITD,ITU,IT,Acc,TR)
          rownames(output1) <- Models[h]
          output <- rbind(output,output1)
          }

output2 <- output[sort.list(output[,6],decreasing=TRUE),]
output3 <- round(output2,digits=3)
```

```
if(m==1){ExRet_1Y <-output3}
else if(m==2){ExRet_6M <- output3}
else if(m==3){ExRet_3M <- output3}
else if(m==4){ExRet_1M <- output3}
else if(m==5){ExRet_1W <- output3}
else{ExRet_1D <- output3}
ListNames[m] <- paste(Target_Variable[m],"Logit Prediction")

}
outlist <- list(ExRet_1Y,ExRet_6M,ExRet_3M,ExRet_1M,ExRet_1W,ExRet_1D)
names(outlist) <- ListNames
return(outlist)
}
```

Logistic regression predictions for pairs of explanatory variables

```
Multiple_Logit
function (DS,Target_Variable,P1,P2,Train)
{
NC1 <- nchar(P1)
Prefix1 <- substr(P1,start=1,stop=3)
NC2 <- nchar(P2)
Prefix2 <- substr(P2,start=1,stop=3)
if(Prefix1=="ROE")
{Label1 <- "TermCD_M2"
Suffix1 <- substr(P1,start=5,stop=NC1)}
else if(Prefix1=="Dis")
{Label1 <-"TermCD_M1"
Suffix1 <- substr(P1,start=10,stop=NC1)}
if(Prefix1=="CAP"){R1 <- DS}
else{
TermCD1 <- paste(Label1,Suffix1,sep="_")
R1 <- DS[DS[,TermCD1]==1,]}
if(Prefix2=="ROE")
{Label2 <- "TermCD_M2"
Suffix2 <- substr(P2,start=5,stop=NC2)}
else if(Prefix2=="Dis")
{Label2 <-"TermCD_M1"
Suffix2 <- substr(P2,start=10,stop=NC2)}
if(Prefix2=="CAP"){Results <- R1}
else{
TermCD2 <- paste(Label2,Suffix2,sep="_")
Results <- R1[R1[,TermCD2]==1,]}

if(Target_Variable=="ER_1Y"){KA <- 251}
else if(Target_Variable=="ER_6M"){KA <- 125}
else if(Target_Variable=="ER_3M"){KA <- 62}
else if(Target_Variable=="ER_1M"){KA <- 20}
else if(Target_Variable=="ER_1W"){KA <- 4}
else if(Target_Variable=="ER_1D"){KA <- 0}

K <- Train
nn <- nrow(Results)

Direction <- rep("Down",nn)
TV <- Results[,Target_Variable]
Pred1 <- Results[,P1]
Pred2 <- Results[,P2]
Direction[TV>0] <-"Up"
DataMat <- cbind(Results,Direction,Pred1,Pred2)
```

```
glm.pred <- rep("Down",(nn-K))

# in 1step forward update fashion
Signal <- rep(0,nn)
Signal[Results[,Target_Variable]>0] <- 1
ss <- length(Signal)
TD <- rep(0,ss)
TU <- rep(0,ss)

for (i in 1:ss)
         {
         TU[i] <- sum(Signal[1:i])
         TD[i] <- i-TU[i]
         }

for (i in K:(nn-1))
         {
         df_train <-DataMat[(1:(i-KA)),]
         df_test <- DataMat[(i+1),]
         glm.fit <- glm(Direction~Pred1+Pred2,family=binomial,data=df_train)
         glm.probs <-predict(glm.fit,df_test,type="response")
         if(glm.probs>.5){glm.pred[i-K+1]<-"Up"}
         }

Direct <- DataMat[((K+1):nn),"Direction"]
t1<-table(glm.pred,Direct)
Tc <- mean(glm.pred==Direct)
outlist <- list("Prediction Table"=t1,"Total Correct"=Tc,"TU"=TU,"TD"=TD)
return(outlist)

}
```

Results from logistic regression predictions with pairs of explanatory variables

```
Multiple_Logit_Results
function (Results,MinTrain)
{
# Function that calculates KNN Prediction Accuracy for combinations of 2 predictor Variables
K <- MinTrain
Target_Variable <- c("ER_1Y","ER_6M","ER_3M","ER_1M","ER_1W","ER_1D")
CAPM <- c("CAPM_1YER","CAPM_6MER","CAPM_3MER","CAPM_1MER","CAPM_1WER","CAPM_1DER")
M1 <- c("Discount_360D_BB","Discount_LTVol_BB","Discount_GARCH_BB")
M2 <- c("ROE_360D_BB","ROE_LTVol_BB","ROE_GARCH_BB")
M3 <- c("Discount_360D_CS","Discount_LTVol_CS","Discount_GARCH_CS")
M4 <- c("ROE_360D_CS","ROE_LTVol_CS","ROE_GARCH_CS")

Models <- c(M1,M2,M3,M4)
Models1 <- c(M2,M3,M4)
Models2 <- c(M3,M4)
ExRet_1Y <- NULL
ExRet_6M <- NULL
ExRet_3M <- NULL
ExRet_1M <- NULL
ExRet_1W <- NULL
ExRet_1D <- NULL
ListNames <- NULL

## For loop over each Target Variable
for (m in 1:length(Target_Variable))
{
```

```
# initial run over CAPM
TVar <- Target_Variable[m]

if(TVar=="ER_1Y"){KA <- 251}
else if(TVar=="ER_6M"){KA <- 125}
else if(TVar=="ER_3M"){KA <- 62}
else if(TVar=="ER_1M"){KA <- 20}
else if(TVar=="ER_1W"){KA <- 4}
else if(TVar=="ER_1D"){KA <- 0}

K1 <- K-KA
Pr1 <- CAPM[m]
Pr2 <- Models[1]
ff <- Multiple_Logit(Results,TVar,Pr1,Pr2,K)
PR <- ff[["Prediction Table"]]
Acc <-ff[["Total Correct"]]
TrainD <- ff[["TD"]]
TrainU <- ff[["TU"]]
ITU <- TrainU[K1]
ITD <- TrainD[K1]
TR <- Table_Results(PR)
output <- cbind(K,K1,ITD,ITU,Acc,TR)
rownames(output) <- paste(CAPM[m],Models[1])

for (h in 2:length(Models))
        {
        Pr2 <- Models[h]
        ff <- Multiple_Logit(Results,TVar,Pr1,Pr2,K)
        PR <- ff[["Prediction Table"]]
        Acc <-ff[["Total Correct"]]
        TrainD <- ff[["TD"]]
        TrainU <- ff[["TU"]]
        ITU <- TrainU[K1]  # initial upward obs in train set
        ITD <- TrainD[K1]  # initial D obs in train set
        TR <- Table_Results(PR)
        output1 <- cbind(K,K1,ITD,ITU,Acc,TR)
        rownames(output1) <- paste(CAPM[m],Models[h])
        output <- rbind(output,output1)
        }
###############################################################
# COMBINATIONS WITH M1 ##
###############################################
for (Z in 1:length(M1))
        {
        Pr1 <- M1[Z]
        for(X in 1:length(Models1))
                {
                Pr2 <- Models1[X]
                ff <- Multiple_Logit(Results,TVar,Pr1,Pr2,K)
                PR <- ff[["Prediction Table"]]
                Acc <-ff[["Total Correct"]]
                TrainD <- ff[["TD"]]
                TrainU <- ff[["TU"]]
                ITU <- TrainU[K1]  # initial upward obs in train set
                ITD <- TrainD[K1]  # initial D obs in train set
                TR <- Table_Results(PR)
                output1 <- cbind(K,K1,ITD,ITU,Acc,TR)
                rownames(output1) <- paste(M1[Z],Models1[X])
                output <- rbind(output,output1)
                }
        }
###############################################################
```

```
# COMBINATIONS WITH M2 ##
###############################################
for (Z in 1:length(M2))
        {
        Pr1 <- M2[Z]
        for(X in 1:length(Models2))
                {
                Pr2 <- Models2[X]
                ff <- Multiple_Logit(Results,TVar,Pr1,Pr2,K)
                PR <- ff[["Prediction Table"]]
                Acc <-ff[["Total Correct"]]
                TrainD <- ff[["TD"]]
                TrainU <- ff[["TU"]]
                ITU <- TrainU[K1]  # initial upward obs in train set
                ITD <- TrainD[K1]  # initial D obs in train set
                TR <- Table_Results(PR)
                output1 <- cbind(K,K1,ITD,ITU,Acc,TR)
                rownames(output1) <- paste(M2[Z],Models2[X])
                output <- rbind(output,output1)
                }
        }
##############################################################
# COMBINATIONS WITH M3 ##
###############################################
for (Z in 1:length(M3))
        {
        Pr1 <- M3[Z]
        for(X in 1:length(M4))
                {
                Pr2 <- M4[X]
                ff <- Multiple_Logit(Results,TVar,Pr1,Pr2,K)
                PR <- ff[["Prediction Table"]]
                Acc <-ff[["Total Correct"]]
                TrainD <- ff[["TD"]]
                TrainU <- ff[["TU"]]
                ITU <- TrainU[K1]  # initial upward obs in train set
                ITD <- TrainD[K1]  # initial D obs in train set
                TR <- Table_Results(PR)
                output1 <- cbind(K,K1,ITD,ITU,Acc,TR)
                rownames(output1) <- paste(M3[Z],M4[X])
                output <- rbind(output,output1)
                }
        }

output2 <- output[sort.list(output[,5],decreasing=TRUE),]
output3 <- round(output2,digits=3)
if(m==1){ExRet_1Y <-output3}
else if(m==2){ExRet_6M <- output3}
else if(m==3){ExRet_3M <- output3}
else if(m==4){ExRet_1M <- output3}
else if(m==5){ExRet_1W <- output3}
else{ExRet_1D <- output3}
ListNames[m] <- paste(Target_Variable[m],"Multiple Logit Prediction")
}

outlist <- list(ExRet_1Y,ExRet_6M,ExRet_3M,ExRet_1M,ExRet_1W,ExRet_1D)
names(outlist) <- ListNames
return(outlist)
}
```

Graphics from univariate logistic regression predictions

```
Logit_Graphics
function (DS,Target_Variable,Predictor,Train)
{
# Function to graph the predictions of chosen 1 variable model
# Making sure only TermCD=1 used in predictor
NC <- nchar(Predictor)
Prefix <- substr(Predictor,start=1,stop=3)
if(Prefix=="ROE")
{Label <- "TermCD_M2"
Suffix <- substr(Predictor,start=5,stop=NC)}
else if(Prefix=="Dis")
{Label <-"TermCD_M1"
Suffix <- substr(Predictor,start=10,stop=NC)}

if(Prefix=="CAP"){Results <- DS}
else{
TermCD <- paste(Label,Suffix,sep="_")
Results <- DS[DS[,TermCD]==1,]}
# we now have clean result set

K <- Train
nn <- nrow(Results)
Direction <- rep("Down",nn)
TV <- Results[,Target_Variable]
Pred <- Results[,Predictor]
Direction[TV>0] <-"Up"
DataMat <- cbind(Results,Direction,Pred)
glm.pred <- rep("Down",(nn-K))
B0 <- NULL
B1 <- NULL
Threshold <- rep(0,nn)
if(Target_Variable=="ER_1Y"){KA <- 251}
else if(Target_Variable =="ER_6M"){KA <- 125}
else if(Target_Variable =="ER_3M"){KA <- 62}
else if(Target_Variable =="ER_1M"){KA <- 20}
else if(Target_Variable =="ER_1W"){KA <- 4}
else if(Target_Variable =="ER_1D"){KA <- 0}

# in 1step forward update fashion
Signal <- rep(0,nn)
Signal[Results[,Target_Variable]>0] <- 1
ss <- length(Signal)
TD <- rep(0,ss)
TU <- rep(0,ss)

for (i in 1:ss)
         {
         TU[i] <- sum(Signal[1:i])
         TD[i] <- i-TU[i]
         }
for (i in K:(nn-1))
         {
         df_train <-DataMat[(1:(i-KA)),]
         df_test <- DataMat[(i+1),]
         glm.fit <- glm(Direction~Pred,family=binomial,data=df_train)
         B0[i] <- glm.fit$coefficients[[1]]
         B1[i] <- glm.fit$coefficients[[2]]
         Threshold[i] <- (log(1)-B0[i])/B1[i]
```

```
          glm.probs <-predict(glm.fit,df_test,type="response")
          if(glm.probs>.5){glm.pred[i-K+1]<-"Up"}
          }

Direct <- DataMat[((K+1):nn),"Direction"]
t1<-table(glm.pred,Direct)
Tc <- mean(glm.pred==Direct)
fam1 <- rep("Init Train",(K-KA))
fam2 <- rep("Pred Dist",KA)
fam <- rep("True",(nn-K))
fam[Direct=="Down" & glm.pred=="Down"]="True"
fam[Direct=="Down" & glm.pred=="Up"]="False"
fam[Direct=="Up" & glm.pred=="Down"]="False"
fam3 <- c(fam1,fam2,fam)
Date <- as.Date(rownames(DataMat),"%Y-%m-%d")
Mat <- data.frame(Date=Date,TargVar=TV,Prediction=fam3,Thresh=Threshold,PVar=Pred)

x <- substitute(DS)
if(x=="ABSA_R2"){Company <-"ABSA"}
else if(x=="BVC_R2"){Company<-"BVC"}
else if(x=="CAPITEC_R2"){Company<-"CAPITEC"}
else if(x=="INL_R2"){Company<-"INL"}
else if(x=="GRF_R2"){Company<-"GRF"}

gg <- ggplot(Mat,aes(x=Date,y=TargVar))
p <- gg+geom_point(aes(color=Prediction))+geom_hline(yintercept=0,linetype="dashed",color="bl
ack")+
geom_line(aes(x=Date,y=Thresh),linetype="dashed")+
geom_line(aes(x=Date,y=PVar))+
scale_color_discrete(name="Legend")+
labs(x="Date",y=paste(Target_Variable,""),title=paste(Company, "Logistic Regression Predictio
ns"),
subtitle=paste(Predictor," As predictor Variable"))
multiplot(p)



outlist <- list("Prediction Table"=t1,"Total Correct"=Tc)#,"Threshold"=Threshold,"TU"=TU,"TD"
=TD)
return(outlist)
}
```

Graphics from logistic regression predictions with pair of predictor variables

```
MLG
function (DS,Target_Variable,P1,P2,Train)
{
NC1 <- nchar(P1)
Prefix1 <- substr(P1,start=1,stop=3)
NC2 <- nchar(P2)
Prefix2 <- substr(P2,start=1,stop=3)
if(Prefix1=="ROE")
{Label1 <- "TermCD_M2"
Suffix1 <- substr(P1,start=5,stop=NC1)}
else if(Prefix1=="Dis")
{Label1 <-"TermCD_M1"
Suffix1 <- substr(P1,start=10,stop=NC1)}
if(Prefix1=="CAP"){R1 <- DS}
else{
TermCD1 <- paste(Label1,Suffix1,sep="_")
```

```
R1 <- DS[DS[,TermCD1]==1,]}
if(Prefix2=="ROE")
{Label2 <- "TermCD_M2"
Suffix2 <- substr(P2,start=5,stop=NC2)}
else if(Prefix2=="Dis")
{Label2 <-"TermCD_M1"
Suffix2 <- substr(P2,start=10,stop=NC2)}
if(Prefix2=="CAP"){Results <- R1}
else{
TermCD2 <- paste(Label2,Suffix2,sep="_")
Results <- R1[R1[,TermCD2]==1,]}

if(Target_Variable=="ER_1Y"){KA <- 251}
else if(Target_Variable=="ER_6M"){KA <- 125}
else if(Target_Variable=="ER_3M"){KA <- 62}
else if(Target_Variable=="ER_1M"){KA <- 20}
else if(Target_Variable=="ER_1W"){KA <- 4}
else if(Target_Variable=="ER_1D"){KA <- 0}

K <- Train
nn <- nrow(Results)

Direction <- rep("Down",nn)
TV <- Results[,Target_Variable]
Pred1 <- Results[,P1]
Pred2 <- Results[,P2]
Direction[TV>0] <-"Up"
DataMat <- cbind(Results,Direction,Pred1,Pred2)
glm.pred <- rep("Down",(nn-K))

# Creating vector of running indication of class balance of U vs D in the training set
# in 1step forward update fashion
Signal <- rep(0,nn)
Signal[Results[,Target_Variable]>0] <- 1
ss <- length(Signal)
TD <- rep(0,ss)
TU <- rep(0,ss)

for (i in 1:ss)
         {
         TU[i] <- sum(Signal[1:i])
         TD[i] <- i-TU[i]
         }

for (i in K:(nn-1))
         {
         df_train <-DataMat[(1:(i-KA)),]
         df_test <- DataMat[(i+1),]
         glm.fit <- glm(Direction~Pred1+Pred2,family=binomial,data=df_train)
         glm.probs <-predict(glm.fit,df_test,type="response")
         if(glm.probs>.5){glm.pred[i-K+1]<-"Up"}
         }

Direct <- DataMat[((K+1):nn),"Direction"]
t1<-table(glm.pred,Direct)
Tc <- mean(glm.pred==Direct)
fam1 <- rep("Init Train",(K-KA))
fam2 <- rep("Pred Dist",KA)
fam <- rep("True",(nn-K))
fam[Direct=="Down" & glm.pred=="Down"]="True"
fam[Direct=="Down" & glm.pred=="Up"]="False"
fam[Direct=="Up" & glm.pred=="Down"]="False"
```

```
fam3 <- c(fam1,fam2,fam)

Date <- as.Date(rownames(DataMat),"%Y-%m-%d")
Mat <- data.frame(Date=Date,TargVar=TV,PredAccuracy=fam3)
x <- substitute(DS)
if(x=="ABSA_R2"){Company <-"ABSA"}
else if(x=="BVC_R2"){Company<-"BVC"}
else if(x=="CAPITEC_R2"){Company<-"CAPITEC"}
else if(x=="INL_R2"){Company<-"INL"}
else if(x=="GRF_R2"){Company<-"GRF"}
gg <- ggplot(Mat,aes(x=Date,y=TargVar))
p <- gg+geom_point(aes(color=PredAccuracy))+geom_hline(yintercept=0,linetype="dashed",color="
black")+
abs(x="Date",y=paste(Target_Variable,""),title=paste(Company, "Multiple Logistic Regression P
rediction"),subtitle=paste(P1,"&", P2))
multiplot(p)

outlist <- list("Prediction Table"=t1,"Total Correct"=Tc)#,"Threshold"=Threshold,"TU"=TU,"TD"
=TD)
return(outlist)

}
```

Varying training & test sets for logistic regression with pairs of predictors

```
VLR.v2
function (DS,Target_Variable,P1,P2,MinTrain,MinPred,Jump)
{
# function recursivley draws 1-step update predictions from Multiple starting points
# Results:          Full set Returns and Return Predictors
# MinTrain:         The Minimum observations in the initial training set
# MinPred:          The minimum number of observations to be predicted in the 1-step update
NC1 <- nchar(P1)
Prefix1 <- substr(P1,start=1,stop=3)
NC2 <- nchar(P2)
Prefix2 <- substr(P2,start=1,stop=3)
if(Prefix1=="ROE")
{Label1 <- "TermCD_M2"
Suffix1 <- substr(P1,start=5,stop=NC1)}
else if(Prefix1=="Dis")
{Label1 <-"TermCD_M1"
Suffix1 <- substr(P1,start=10,stop=NC1)}
if(Prefix1=="CAP"){R1 <- DS}
else{
TermCD1 <- paste(Label1,Suffix1,sep="_")
R1 <- DS[DS[,TermCD1]==1,]}
if(Prefix2=="ROE")
{Label2 <- "TermCD_M2"
Suffix2 <- substr(P2,start=5,stop=NC2)}
else if(Prefix2=="Dis")
{Label2 <-"TermCD_M1"
Suffix2 <- substr(P2,start=10,stop=NC2)}
if(Prefix2=="CAP"){Results <- R1}
else{
TermCD2 <- paste(Label2,Suffix2,sep="_")
Results <- R1[R1[,TermCD2]==1,]}
index <- seq(from=MinTrain,to=(nrow(Results)-MinPred),by=Jump)
NN <- length(index)

TVar <- Target_Variable
```

```
if(TVar=="ER_1Y"){KA <- 251}
else if(TVar=="ER_6M"){KA <- 125}
else if(TVar=="ER_3M"){KA <- 62}
else if(TVar=="ER_1M"){KA <- 20}
else if(TVar=="ER_1W"){KA <- 4}
else if(TVar=="ER_1D"){KA <- 0}

K <- index[1]
K1 <- K-KA
ff <- Multiple_Logit(Results,TVar,P1,P2,K)
PR <- ff[["Prediction Table"]]
Acc <-ff[["Total Correct"]]
TrainD <- ff[["TD"]]
TrainU <- ff[["TU"]]
TR <- Table_Results(PR)
ITU <- TrainU[K1]
ITD <- TrainD[K1]
output <- cbind(K,K1,ITD,ITU,Acc,TR)

for (h in 2:NN)
         {
         K <- index[h]
         K1 <- K-KA
         ff <- Multiple_Logit(Results,TVar,P1,P2,K)
         PR <- ff[["Prediction Table"]]
         Acc <-ff[["Total Correct"]]
         TrainD <- ff[["TD"]]
         TrainU <- ff[["TU"]]

         ITU <- TrainU[K1]  # initial upward obs in train set
         ITD <- TrainD[K1]  # initial D obs in train set
         TR <- Table_Results(PR)
         output1 <- cbind(K,K1,ITD,ITU,Acc,TR)
         output <- rbind(output,output1)
         }

output3 <- round(output,digits=3)
return(output3)
}
```

Varying training and test sets for univariate logistic regression

```
VLR
function (DS,Target_Variable,P1,MinTrain,MinPred,Jump)
{
# function recursivley draws 1-step update predictions from Multiple starting points

# Results:        Full set Returns and Return Predictors
# MinTrain:       The Minimum observations in the initial training set
# MinPred:        The minimum number of obser to be predicted in the 1-step update
NC <- nchar(P1)
Prefix <- substr(P1,start=1,stop=3)
if(Prefix=="ROE")
{Label <- "TermCD_M2"
Suffix <- substr(P1,start=5,stop=NC)}
else if(Prefix=="Dis")
{Label <-"TermCD_M1"
Suffix <- substr(P1,start=10,stop=NC)}
else if(Prefix=="CMD"){
Label <- "TermCD_M3"
Suffix <- substr(P1,start=5,stop=NC)}
```

```
if(Prefix=="CAP"){Results <- DS}
else{
TermCD <- paste(Label,Suffix,sep="_")
Results <- DS[DS[,TermCD]==1,]}
index <- seq(from=MinTrain,to=(nrow(Results)-MinPred),by=Jump)
NN <- length(index)

TVar <- Target_Variable
if(TVar=="ER_1Y"){KA <- 251}
else if(TVar=="ER_6M"){KA <- 125}
else if(TVar=="ER_3M"){KA <- 62}
else if(TVar=="ER_1M"){KA <- 20}
else if(TVar=="ER_1W"){KA <- 4}
else if(TVar=="ER_1D"){KA <- 0}

K <- index[1]
K1 <- K-KA

ff <- Logit(Results,TVar,P1,K)
PR <- ff[["Prediction Table"]]
Acc <-ff[["Total Correct"]]
Threshold <- ff[["Threshold"]]
TrainD <- ff[["TD"]]
TrainU <- ff[["TU"]]

TR <- Table_Results(PR)
IT <- Threshold[K]
ITU <- TrainU[K1]
ITD <- TrainD[K1]
output <- cbind(K,K1,ITD,ITU,IT,Acc,TR)

for (h in 2:NN)
         {
         K <- index[h]
         K1 <- K-KA
         ff <- Logit(Results,TVar,P1,K)
         PR <- ff[["Prediction Table"]]
         Acc <-ff[["Total Correct"]]
         Threshold <- ff[["Threshold"]]
         TrainD <- ff[["TD"]]
         TrainU <- ff[["TU"]]

         IT <- Threshold[K] # Initial Threshold
         ITU <- TrainU[K1]  # initial upward obs in train set
         ITD <- TrainD[K1]  # initial D obs in train set
         TR <- Table_Results(PR)
         output1 <- cbind(K,K1,ITD,ITU,IT,Acc,TR)
         output <- rbind(output,output1)
         }

output3 <- round(output,digits=3)
return(output3)
}
```

KNN functions

KNN-Classifier predictions for single explanatory variable

```
KNN_Class
function (DS,Target_Variable,Predictor,Train,Neighbors)
{
J <- Neighbors
NC <- nchar(Predictor)
Prefix <- substr(Predictor,start=1,stop=3)
if(Prefix=="ROE")
{Label <- "TermCD_M2"
Suffix <- substr(Predictor,start=5,stop=NC)}
else if(Prefix=="Dis")
{Label <-"TermCD_M1"
Suffix <- substr(Predictor,start=10,stop=NC)}
else if(Prefix=="CMD"){
Label <- "TermCD_M3"
Suffix <- substr(Predictor,start=5,stop=NC)}
if(Prefix=="CAP"){Results <- DS}
else{
TermCD <- paste(Label,Suffix,sep="_")
Results <- DS[DS[,TermCD]==1,]}

if(Target_Variable=="ER_1Y"){KA <- 251}
else if(Target_Variable=="ER_6M"){KA <- 125}
else if(Target_Variable=="ER_3M"){KA <- 62}
else if(Target_Variable=="ER_1M"){KA <- 20}
else if(Target_Variable=="ER_1W"){KA <- 4}
else if(Target_Variable=="ER_1D"){KA <- 0}

K <- Train
nn <- nrow(Results)
Direction <- rep("Down",nn)
TV <- Results[,Target_Variable]
Direction[TV>0] <-"Up"

DataMat <- cbind(Results,Direction)
glm.pred <- rep("Down",(nn-K))
# Creating vector of running indication of class balance of U vs D in the training set
forward update fashion
Signal <- rep(0,nn)
Signal[Results[,Target_Variable]>0] <- 1
ss <- length(Signal)
TD <- rep(0,ss)
TU <- rep(0,ss)

for (i in 1:ss)
         {
         TU[i] <- sum(Signal[1:i])
         TD[i] <- i-TU[i]
         }
for (i in K:(nn-1))
         {
         train.x <- as.matrix(cbind(DataMat[(1:(i-KA)),Predictor]))
         test.x <- as.matrix(cbind(DataMat[(i+1),Predictor]))
         train.direction <- DataMat[(1:(i-KA)),"Direction"]
         knn.prediction <- knn(train.x,test.x,train.direction,k=J)
         AB <- summary(knn.prediction)
         if(AB[2]==1){glm.pred[i-K+1]<-"Up"}
         }
```

```
Direct <- DataMat[((K+1):nn),"Direction"]
t1<-table(glm.pred,Direct)
Tc <- mean(glm.pred==Direct)
outlist <- list("Prediction Table"=t1,"Total Correct"=Tc,"TU"=TU,"TD"=TD)
return(outlist)
}
```

KNN classifier results for single explanatory variable

```
KNN_Results
function (Results,MinTrain,Neighbors)
{
# Results:                            Full set of ROE VS ER_3M
# MinTrain:                 The Minimum observations in the initial training set
K <- MinTrain
J <- Neighbors
Target_Variable <- c("ER_1Y","ER_6M","ER_3M","ER_1M","ER_1W","ER_1D")
CAPM <- c("CAPM_1YER","CAPM_6MER","CAPM_3MER","CAPM_1MER","CAPM_1WER","CAPM_1DER")
Models <- c("CAPM",
         "Discount_360D_BB","Discount_LTVol_BB","Discount_GARCH_BB",
         "ROE_360D_BB","ROE_LTVol_BB","ROE_GARCH_BB",
         "Discount_360D_CS","Discount_LTVol_CS","Discount_GARCH_CS",
         "ROE_360D_CS","ROE_LTVol_CS","ROE_GARCH_CS")

ExRet_1Y <- NULL
ExRet_6M <- NULL
ExRet_3M <- NULL
ExRet_1M <- NULL
ExRet_1W <- NULL
ExRet_1D <- NULL
ListNames <- NULL

## For loop over each Target Variable
for (m in 1:length(Target_Variable))
{
# initial run over CAPM
TVar <- Target_Variable[m]
if(TVar=="ER_1Y"){KA <- 251}
else if(TVar=="ER_6M"){KA <- 125}
else if(TVar=="ER_3M"){KA <- 62}
else if(TVar=="ER_1M"){KA <- 20}
else if(TVar=="ER_1W"){KA <- 4}
else if(TVar=="ER_1D"){KA <- 0}

K1 <- K-KA
Pr <- CAPM[m]
ff <- KNN_Class(Results,TVar,Pr,K,J)
PR <- ff[["Prediction Table"]]
Acc <-ff[["Total Correct"]]
TrainD <- ff[["TD"]]
TrainU <- ff[["TU"]]
ITU <- TrainU[K1]
ITD <- TrainD[K1]
TR <- Table_Results(PR
(K,K1,ITD,ITU,Acc,TR)
rownames(output) <- CAPM[m]

for (h in 2:length(Models))
         {
         Pr <- Models[h]
```

```
        ff <- KNN_Class(Results,TVar,Pr,K,J)
        PR <- ff[["Prediction Table"]]
        Acc <-ff[["Total Correct"]]
        TrainD <- ff[["TD"]]
        TrainU <- ff[["TU"]]
        ITU <- TrainU[K1]  # initial upward obs in train set
        ITD <- TrainD[K1]  # initial D obs in train set
        TR <- Table_Results(PR) # how to deal with table when results contain predictions of
 only 1 Direction
        output1 <- cbind(K,K1,ITD,ITU,Acc,TR)
        rownames(output1) <- Models[h]
        output <- rbind(output,output1)
        }

output2<- output[sort.list(output[,5],decreasing=TRUE),]
output3<- round(output2,digits=3)

if(m==1){ExRet_1Y <-output3}
else if(m==2){ExRet_6M <- output3}
else if(m==3){ExRet_3M <- output3}
else if(m==4){ExRet_1M <- output3}
else if(m==5){ExRet_1W <- output3}
else{ExRet_1D <- output3}
ListNames[m] <- paste(Target_Variable[m],"KNN Prediction",Neighbors)
}

outlist <- list(ExRet_1Y,ExRet_6M,ExRet_3M,ExRet_1M,ExRet_1W,ExRet_1D)
names(outlist) <- ListNames
return(outlist)
}
```

KNN classifier predictions for predictor pairs

```
Multiple_KNN_Class
function (DS,Target_Variable,P1,P2,Train,Neighbors)
{
# KNN Function to predict Negative and Positive Returns using 2 Predictions
# Dichotomous stratification
J <- Neighbors
NC1 <- nchar(P1)
Prefix1 <- substr(P1,start=1,stop=3)
NC2 <- nchar(P2)
Prefix2 <- substr(P2,start=1,stop=3)
if(Prefix1=="ROE")
{Label1 <- "TermCD_M2"
Suffix1 <- substr(P1,start=5,stop=NC1)}
else if(Prefix1=="Dis")
{Label1 <-"TermCD_M1"
Suffix1 <- substr(P1,start=10,stop=NC1)}

if(Prefix1=="CAP"){R1 <- DS}
else{
TermCD1 <- paste(Label1,Suffix1,sep="_")
R1 <- DS[DS[,TermCD1]==1,]}
if(Prefix2=="ROE")
{Label2 <- "TermCD_M2"
Suffix2 <- substr(P2,start=5,stop=NC2)}
else if(Prefix2=="Dis")
{Label2 <-"TermCD_M1"
Suffix2 <- substr(P2,start=10,stop=NC2)}
```

```
if(Prefix2=="CAP"){Results <- R1}
else{
TermCD2 <- paste(Label2,Suffix2,sep="_")
Results <- R1[R1[,TermCD2]==1,]}

if(Target_Variable=="ER_1Y"){KA <- 251}
else if(Target_Variable=="ER_6M"){KA <- 125}
else if(Target_Variable=="ER_3M"){KA <- 62}
else if(Target_Variable=="ER_1M"){KA <- 20}
else if(Target_Variable=="ER_1W"){KA <- 4}
else if(Target_Variable=="ER_1D"){KA <- 0}

K <- Train
nn <- nrow(Results)
Direction <- rep("Down",nn)
TV <- Results[,Target_Variable]
Direction[TV>0] <-"Up"
DataMat <- cbind(Results,Direction)
glm.pred <- rep("Down",(nn-K))
Signal <- rep(0,nn)
Signal[Results[,Target_Variable]>0] <- 1
ss <- length(Signal)
TD <- rep(0,ss)
TU <- rep(0,ss)

for (i in 1:ss)
         {
         TU[i] <- sum(Signal[1:i])
         TD[i] <- i-TU[i]
         }

for (i in K:(nn-1))
         {
         train.x <- as.matrix(cbind(DataMat[(1:(i-KA)),P1],DataMat[(1:(i-KA)),P2]))
         test.x <- as.matrix(cbind(DataMat[(i+1),P1],DataMat[(i+1),P2]))
         train.direction <- DataMat[(1:(i-KA)),"Direction"]
         knn.prediction <- knn(train.x,test.x,train.direction,k=J)
         AB <- summary(knn.prediction)
         if(AB[2]==1){glm.pred[i-K+1]<-"Up"}
         }

Direct <- DataMat[((K+1):nn),"Direction"]
t1<-table(glm.pred,Direct)
Tc <- mean(glm.pred==Direct)
outlist <- list("Prediction Table"=t1,"Total Correct"=Tc,"TU"=TU,"TD"=TD)
return(outlist)
}
```

Full results for KNN classifier with pairs of explanatory variables

```
Multiple_KNN_Results
function (Results,MinTrain,Neighbors)
{
# Function that calculates KNN Prediction Accuracy for combinations of 2 predictor Variables
K <- MinTrain
J <- Neighbors
Target_Variable <- c("ER_1Y","ER_6M","ER_3M","ER_1M","ER_1W","ER_1D")
CAPM <- c("CAPM_1YER","CAPM_6MER","CAPM_3MER","CAPM_1MER","CAPM_1WER","CAPM_1DER")
M1 <- c("Discount_360D_BB","Discount_LTVol_BB","Discount_GARCH_BB")
M2 <- c("ROE_360D_BB","ROE_LTVol_BB","ROE_GARCH_BB")
```

```
M3 <- c("Discount_360D_CS","Discount_LTVol_CS","Discount_GARCH_CS")
M4 <- c("ROE_360D_CS","ROE_LTVol_CS","ROE_GARCH_CS")
Models <- c(M1,M2,M3,M4)
Models1 <- c(M2,M3,M4)
Models2 <- c(M3,M4)
ExRet_1Y <- NULL
ExRet_6M <- NULL
ExRet_3M <- NULL
ExRet_1M <- NULL
ExRet_1W <- NULL
ExRet_1D <- NULL
ListNames <- NULL

## For loop over each Target Variable
for (m in 1:length(Target_Variable))
{
# initial run over CAPM
TVar <- Target_Variable[m]
if(TVar=="ER_1Y"){KA <- 251}
else if(TVar=="ER_6M"){KA <- 125}
else if(TVar=="ER_3M"){KA <- 62}
else if(TVar=="ER_1M"){KA <- 20}
else if(TVar=="ER_1W"){KA <- 4}
else if(TVar=="ER_1D"){KA <- 0}

K1 <- K-KA
Pr1 <- CAPM[m]
Pr2 <- Models[1]
ff <- Multiple_KNN_Class(Results,TVar,Pr1,Pr2,K,J)
PR <- ff[["Prediction Table"]]
Acc <-ff[["Total Correct"]]
TrainD <- ff[["TD"]]
TrainU <- ff[["TU"]]
ITU <- TrainU[K1]
ITD <- TrainD[K1]
TR <- Table_Results(PR)# how to deal with table when results contain predictions of only 1 Di
rection
output <- cbind(K,K1,ITD,ITU,Acc,TR)
rownames(output) <- paste(CAPM[m],Models[1])

for (h in 2:length(Models))
          {
          Pr2 <- Models[h]
          ff <- Multiple_KNN_Class(Results,TVar,Pr1,Pr2,K,J)
          PR <- ff[["Prediction Table"]]
          Acc <-ff[["Total Correct"]]
          TrainD <- ff[["TD"]]
          TrainU <- ff[["TU"]]
          ITU <- TrainU[K1]  # initial upward obs in train set
          ITD <- TrainD[K1]  # initial D obs in train set
          TR <- Table_Results(PR)
          output1 <- cbind(K,K1,ITD,ITU,Acc,TR)
          rownames(output1) <- paste(CAPM[m],Models[h])
          output <- rbind(output,output1)
          }
##############################################################
# COMBINATIONS WITH M1 ##
###############################################
for (Z in 1:length(M1))
          {
          Pr1 <- M1[Z]
          for(X in 1:length(Models1))
```

```
                    {
                    Pr2 <- Models1[X]
                    ff <- Multiple_KNN_Class(Results,TVar,Pr1,Pr2,K,J)
                    PR <- ff[["Prediction Table"]]
                    Acc <-ff[["Total Correct"]]
                    TrainD <- ff[["TD"]]
                    TrainU <- ff[["TU"]]
                    ITU <- TrainU[K1]  # initial upward obs in train set
                    ITD <- TrainD[K1]  # initial D obs in train set
                    TR <- Table_Results(PR) # how to deal with table when results contain pred
ictions of only 1 Direction
                    output1 <- cbind(K,K1,ITD,ITU,Acc,TR)
                    rownames(output1) <- paste(M1[Z],Models1[X])
                    output <- rbind(output,output1)
                    }
          }
##############################################################
# COMBINATIONS WITH M2 ##
###############################################
for (Z in 1:length(M2))
          {
          Pr1 <- M2[Z]
          for(X in 1:length(Models2))
                    {
                    Pr2 <- Models2[X]
                    ff <- Multiple_KNN_Class(Results,TVar,Pr1,Pr2,K,J)
                    PR <- ff[["Prediction Table"]]
                    Acc <-ff[["Total Correct"]]
                    TrainD <- ff[["TD"]]
                    TrainU <- ff[["TU"]]
                    ITU <- TrainU[K1]  # initial upward obs in train set
                    ITD <- TrainD[K1]  # initial D obs in train set
                    TR <- Table_Results(PR)
                    output1 <- cbind(K,K1,ITD,ITU,Acc,TR)
                    rownames(output1) <- paste(M2[Z],Models2[X])
                    output <- rbind(output,output1)
                    }
          }
##############################################################
# COMBINATIONS WITH M3 ##
###############################################
for (Z in 1:length(M3))
          {
          Pr1 <- M3[Z]
          for(X in 1:length(M4))
                    {
                    Pr2 <- M4[X]
                    ff <- Multiple_KNN_Class(Results,TVar,Pr1,Pr2,K,J)
                    PR <- ff[["Prediction Table"]]
                    Acc <-ff[["Total Correct"]]
                    TrainD <- ff[["TD"]]
                    TrainU <- ff[["TU"]]
                    ITU <- TrainU[K1]  # initial upward obs in train set
                    ITD <- TrainD[K1]  # initial D obs in train set
                    TR <- Table_Results(PR)
                    output1 <- cbind(K,K1,ITD,ITU,Acc,TR)
                    rownames(output1) <- paste(M3[Z],M4[X])
                    output <- rbind(output,output1)
                    }
          }

output2 <- output[sort.list(output[,5],decreasing=TRUE),]
```

```
output3 <- round(output2,digits=3)
if(m==1){ExRet_1Y <-output3}
else if(m==2){ExRet_6M <- output3}
else if(m==3){ExRet_3M <- output3}
else if(m==4){ExRet_1M <- output3}
else if(m==5){ExRet_1W <- output3}
else{ExRet_1D <- output3}
ListNames[m] <- paste(Target_Variable[m],"Multiple KNN Prediction",Neighbors)
}

outlist <- list(ExRet_1Y,ExRet_6M,ExRet_3M,ExRet_1M,ExRet_1W,ExRet_1D)
names(outlist) <- ListNames
return(outlist)
}
```

Graphics of KNN predictions with single predictor variable

```
KNN_Graphics
function (DS,Target_Variable,Predictor,Train,Neighbors)
{
# Function to Graph Test prediction over sample for KNN prediction
# for single chosen variable

J <- Neighbors
NC <- nchar(Predictor)
Prefix <- substr(Predictor,start=1,stop=3)
if(Prefix=="ROE")
{Label <- "TermCD_M2"
Suffix <- substr(Predictor,start=5,stop=NC)}
else if(Prefix=="Dis")
{Label <-"TermCD_M1"
Suffix <- substr(Predictor,start=10,stop=NC)}
else if(Prefix=="CMD"){
Label <- "TermCD_M3"
Suffix <- substr(Predictor,start=5,stop=NC)}

if(Prefix=="CAP"){Results <- DS}
else{
TermCD <- paste(Label,Suffix,sep="_")
Results <- DS[DS[,TermCD]==1,]}

if(Target_Variable=="ER_1Y"){KA <- 251}
else if(Target_Variable=="ER_6M"){KA <- 125}
else if(Target_Variable=="ER_3M"){KA <- 62}
else if(Target_Variable=="ER_1M"){KA <- 20}
else if(Target_Variable=="ER_1W"){KA <- 4}
else if(Target_Variable=="ER_1D"){KA <- 0}

K <- Train
nn <- nrow(Results)
Direction <- rep("Down",nn)
TV <- Results[,Target_Variable]
Direction[TV>0] <-"Up"
DataMat <- cbind(Results,Direction)
glm.pred <- rep("Down",(nn-K))

# Creating vector of running indication of class balance of U vs D in the training set
# in 1step forward update fashion
Signal <- rep(0,nn)
Signal[Results[,Target_Variable]>0] <- 1
```

```
ss <- length(Signal)
TD <- rep(0,ss)
TU <- rep(0,ss)

for (i in 1:ss)
          {
          TU[i] <- sum(Signal[1:i])
          TD[i] <- i-TU[i]
          }

for (i in K:(nn-1))
          {
          train.x <- as.matrix(cbind(DataMat[(1:(i-KA)),Predictor]))
          test.x <- as.matrix(cbind(DataMat[(i+1),Predictor]))
          train.direction <- DataMat[(1:(i-KA)),"Direction"]
          knn.prediction <- knn(train.x,test.x,train.direction,k=J)
          AB <- summary(knn.prediction)
          if(AB[2]==1){glm.pred[i-K+1]<-"Up"}
          }

Direct <- DataMat[((K+1):nn),"Direction"]
t1<-table(glm.pred,Direct)
Tc <- mean(glm.pred==Direct)
fam1 <- rep("Init Train",(K-KA))
fam2 <- rep("Pred Dist",KA)
fam <- rep("Correct",(nn-K))
fam[Direct=="Down" & glm.pred=="Down"]="Correct"
fam[Direct=="Down" & glm.pred=="Up"]="False"
fam[Direct=="Up" & glm.pred=="Down"]="False"
fam3 <- c(fam1,fam2,fam)
Pred <- DataMat[,Predictor]

Date <- as.Date(rownames(DataMat),"%Y-%m-%d")
Mat <- data.frame(Date=Date,TargVar=TV,PredAccuracy=fam3,PVar=Pred)

x <- substitute(DS)
if(x=="ABSA_R2"){Company <-"ABSA"}
else if(x=="BVC_R2"){Company<-"BVC"}
else if(x=="CAPITEC_R2"){Company<-"CAPITEC"}
else if(x=="INL_R2"){Company<-"INL"}
else if(x=="GRF_R2"){Company<-"GRF"}

gg <- ggplot(Mat,aes(x=Date,y=TargVar))
p <- gg+geom_point(aes(color=PredAccuracy))+geom_hline(yintercept=0,linetype="dashed",color="
black")+
geom_line(aes(x=Date,y=PVar))+
labs(x="Date",y=paste(Target_Variable,""),title=paste(Company, "KNN Predictions"),
                   subtitle=paste(Predictor,""))
multiplot(p)
outlist <- list("Prediction Table"=t1,"Total Correct"=Tc)#,"TU"=TU,"TD"=TD)
return(outlist)
}
```

Graphics from KNN predictions with pair of predictor variables

MKG

```
function (DS,Target_Variable,P1,P2,Train,Neighbors)
{
# KNN Function to predict Negative and Positive Returns using 2 Predictions
# Dichotomous stratification
```

```
J <- Neighbors
NC1 <- nchar(P1)
Prefix1 <- substr(P1,start=1,stop=3)
NC2 <- nchar(P2)
Prefix2 <- substr(P2,start=1,stop=3)
if(Prefix1=="ROE")
{Label1 <- "TermCD_M2"
Suffix1 <- substr(P1,start=5,stop=NC1)}
else if(Prefix1=="Dis")
{Label1 <-"TermCD_M1"
Suffix1 <- substr(P1,start=10,stop=NC1)}
if(Prefix1=="CAP"){R1 <- DS}
else{
TermCD1 <- paste(Label1,Suffix1,sep="_")
R1 <- DS[DS[,TermCD1]==1,]}
if(Prefix2=="ROE")
{Label2 <- "TermCD_M2"
Suffix2 <- substr(P2,start=5,stop=NC2)}
else if(Prefix2=="Dis")
{Label2 <-"TermCD_M1"
Suffix2 <- substr(P2,start=10,stop=NC2)}
if(Prefix2=="CAP"){Results <- R1}
else{
TermCD2 <- paste(Label2,Suffix2,sep="_")
Results <- R1[R1[,TermCD2]==1,]}


if(Target_Variable=="ER_1Y"){KA <- 251}
else if(Target_Variable=="ER_6M"){KA <- 125}
else if(Target_Variable=="ER_3M"){KA <- 62}
else if(Target_Variable=="ER_1M"){KA <- 20}
else if(Target_Variable=="ER_1W"){KA <- 4}
else if(Target_Variable=="ER_1D"){KA <- 0}

K <- Train
nn <- nrow(Results)
Direction <- rep("Down",nn)
TV <- Results[,Target_Variable]
Direction[TV>0] <-"Up"
DataMat <- cbind(Results,Direction)
glm.pred <- rep("Down",(nn-K))
Signal <- rep(0,nn)
Signal[Results[,Target_Variable]>0] <- 1
ss <- length(Signal)
TD <- rep(0,ss)
TU <- rep(0,ss)

for (i in 1:ss)
         {
         TU[i] <- sum(Signal[1:i])
         TD[i] <- i-TU[i]
         }

for (i in K:(nn-1))
         {
         train.x <- as.matrix(cbind(DataMat[(1:(i-KA)),P1],DataMat[(1:(i-KA)),P2]))
         test.x <- as.matrix(cbind(DataMat[(i+1),P1],DataMat[(i+1),P2]))
         train.direction <- DataMat[(1:(i-KA)),"Direction"]
         knn.prediction <- knn(train.x,test.x,train.direction,k=J)
         AB <- summary(knn.prediction)
         if(AB[2]==1){glm.pred[i-K+1]<-"Up"}
         }
```

```
Direct <- DataMat[((K+1):nn),"Direction"]
t1<-table(glm.pred,Direct)
Tc <- mean(glm.pred==Direct)
fam1 <- rep("Init Train",(K-KA))
fam2 <- rep("Pred Dist",KA)
fam <- rep("True",(nn-K))
fam[Direct=="Down" & glm.pred=="Down"]="True"
fam[Direct=="Down" & glm.pred=="Up"]="False"
fam[Direct=="Up" & glm.pred=="Down"]="False"
fam3 <- c(fam1,fam2,fam)

Date <- as.Date(rownames(DataMat),"%Y-%m-%d")
Mat <- data.frame(Date=Date,TargVar=TV,PredAccuracy=fam3)

x <- substitute(DS)
if(x=="ABSA_R2"){Company <-"ABSA"}
else if(x=="BVC_R2"){Company<-"BVC"}
else if(x=="CAPITEC_R2"){Company<-"CAPITEC"}
else if(x=="INL_R2"){Company<-"INL"}
else if(x=="GRF_R2"){Company<-"GRF"}

gg <- ggplot(Mat,aes(x=Date,y=TargVar))
p <- gg+geom_point(aes(color=PredAccuracy))+geom_hline(yintercept=0,linetype="dashed",color="
black")+
labs(x="Date",y=paste(Target_Variable,""),title=paste(Company, "Multiple KNN Prediction"),
         subtitle=paste(P1,"&", P2))
multiplot(p)
outlist <- list("Prediction Table"=t1,"Total Correct"=Tc)#,"TU"=TU,"TD"=TD)
return(outlist)

}
```

Varying training and test sets for KNN Classifier with single predictor variable

```
VKR
function (DS,Target_Variable,P1,Neighbors,MinTrain,MinPred,Jump)
{
# Results:                             Full set of ROE VS ER_3M
# MinTrain:                            The Minimum observations in the initial training set
J <- Neighbors
NC <- nchar(P1)
Prefix <- substr(P1,start=1,stop=3)
if(Prefix=="ROE")
{Label <- "TermCD_M2"
Suffix <- substr(P1,start=5,stop=NC)}
else if(Prefix=="Dis")
{Label <-"TermCD_M1"
Suffix <- substr(P1,start=10,stop=NC)}
else if(Prefix=="CMD"){
Label <- "TermCD_M3"
Suffix <- substr(P1,start=5,stop=NC)}
if(Prefix=="CAP"){Results <- DS}
else{
TermCD <- paste(Label,Suffix,sep="_")
Results <- DS[DS[,TermCD]==1,]}
index <- seq(from=MinTrain,to=(nrow(Results)-MinPred),by=Jump)
NN <- length(index)

TVar <- Target_Variable
if(TVar=="ER_1Y"){KA <- 251}
```

```
else if(TVar=="ER_6M"){KA <- 125}
else if(TVar=="ER_3M"){KA <- 62}
else if(TVar=="ER_1M"){KA <- 20}
else if(TVar=="ER_1W"){KA <- 4}
else if(TVar=="ER_1D"){KA <- 0}

K <- index[1]
K1 <- K-KA

ff <- KNN_Class(Results,TVar,P1,K,J)
PR <- ff[["Prediction Table"]]
Acc <-ff[["Total Correct"]]
TrainD <- ff[["TD"]]
TrainU <- ff[["TU"]]
ITU <- TrainU[K1]
ITD <- TrainD[K1]
TR <- Table_Results(PR)# how to deal with table when results contain predictions of only 1 Di
rection
output <- cbind(K,K1,ITD,ITU,Acc,TR)

for (h in 2:NN)
         {
         K <- index[h]
         K1 <- K-KA
         ff <- KNN_Class(Results,TVar,P1,K,J)
         PR <- ff[["Prediction Table"]]
         Acc <-ff[["Total Correct"]]
         TrainD <- ff[["TD"]]
         TrainU <- ff[["TU"]]
         ITU <- TrainU[K1]  # initial upward obs in train set
         ITD <- TrainD[K1]  # initial D obs in train set
         TR <- Table_Results(PR) # how to deal with table when results contain predictions of
 only 1 Direction
         output1 <- cbind(K,K1,ITD,ITU,Acc,TR)
         output <- rbind(output,output1)
         }

output3<- round(output,digits=3)
return(output3)
}
```

Varying training and test sets for KNN Classifier with pairs of predictor variables

```
VKR.v2
function (DS,Target_Variable,P1,P2,Neighbors,MinTrain,MinPred,Jump)
{
J <- Neighbors
NC1 <- nchar(P1)
Prefix1 <- substr(P1,start=1,stop=3)
NC2 <- nchar(P2)
Prefix2 <- substr(P2,start=1,stop=3)
if(Prefix1=="ROE")
{Label1 <- "TermCD_M2"
Suffix1 <- substr(P1,start=5,stop=NC1)}
else if(Prefix1=="Dis")
{Label1 <-"TermCD_M1"
Suffix1 <- substr(P1,start=10,stop=NC1)}
if(Prefix1=="CAP"){R1 <- DS}
else{
TermCD1 <- paste(Label1,Suffix1,sep="_")
R1 <- DS[DS[,TermCD1]==1,]}
```

```
if(Prefix2=="ROE")
{Label2 <- "TermCD_M2"
Suffix2 <- substr(P2,start=5,stop=NC2)}
else if(Prefix2=="Dis")
{Label2 <-"TermCD_M1"
Suffix2 <- substr(P2,start=10,stop=NC2)}
if(Prefix2=="CAP"){Results <- R1}
else{
TermCD2 <- paste(Label2,Suffix2,sep="_")
Results <- R1[R1[,TermCD2]==1,]}
index <- seq(from=MinTrain,to=(nrow(Results)-MinPred),by=Jump)
NN <- length(index)

TVar <- Target_Variable
if(TVar=="ER_1Y"){KA <- 251}
else if(TVar=="ER_6M"){KA <- 125}
else if(TVar=="ER_3M"){KA <- 62}
else if(TVar=="ER_1M"){KA <- 20}
else if(TVar=="ER_1W"){KA <- 4}
else if(TVar=="ER_1D"){KA <- 0}
K <- index[1]
K1 <- K-KA

ff <- Multiple_KNN_Class(Results,TVar,P1,P2,K,J)
PR <- ff[["Prediction Table"]]
Acc <-ff[["Total Correct"]]
TrainD <- ff[["TD"]]
TrainU <- ff[["TU"]]
ITU <- TrainU[K1]
ITD <- TrainD[K1]
TR <- Table_Results(PR)
output <- cbind(K,K1,ITD,ITU,Acc,TR)

for (h in 2:NN)
         {
         K <- index[h]
         K1 <- K-KA
         ff <- Multiple_KNN_Class(Results,TVar,P1,P2,K,J)
         PR <- ff[["Prediction Table"]]
         Acc <-ff[["Total Correct"]]
         TrainD <- ff[["TD"]]
         TrainU <- ff[["TU"]]
         ITU <- TrainU[K1]  # initial upward obs in train set
         ITD <- TrainD[K1]  # initial D obs in train set
         TR <- Table_Results(PR)
         output1 <- cbind(K,K1,ITD,ITU,Acc,TR)
         output <- rbind(output,output1)
         }

output3<- round(output,digits=3)
return(output3)
}
```

Variable selection

Subset selection using regitfull

```
Var_Select
function (DS,TVar,K)
{
# Function to perform Regression variable subset selection

#Target_Variable <- c("ER_1Y","ER_6M","ER_3M","ER_1M","ER_1W","ER_1D")
#CAPM <- c("CAPM_1YER","CAPM_6MER","CAPM_3MER","CAPM_1MER","CAPM_1WER","CAPM_1DER")

Models <- c("CAPM_1YER","Discount_360D_BB","Discount_LTVol_BB","Discount_GARCH_BB",
                            "ROE_360D_BB","ROE_LTVol_BB","ROE_GARCH_BB",
                            "Discount_360D_CS","Discount_LTVol_CS","Discount_GARCH_CS",
                            "ROE_360D_CS","ROE_LTVol_CS","ROE_GARCH_CS")

# first TermCd Check
PX <- Models[1]
NC <- nchar(PX)
Prefix <- substr(PX,start=1,stop=3)
if(Prefix=="ROE")
{Label <- "TermCD_M2"
Suffix <- substr(PX,start=5,stop=NC)}
else if(Prefix=="Dis")
{Label <-"TermCD_M1"
Suffix <- substr(PX,start=10,stop=NC)}
if(Prefix=="CAP"){R1 <- DS}
else{
TermCD <- paste(Label,Suffix,sep="_")
R1 <- DS[DS[,TermCD]==1,]}

for(i in 2:length(Models))
          {
          PX1 <- Models[i]
          NC1 <- nchar(PX1)
          Prefix1 <- substr(PX1,start=1,stop=3)
          if(Prefix1=="ROE")
          {Label1 <- "TermCD_M2"
          Suffix1 <- substr(PX1,start=5,stop=NC1)}
          else if(Prefix1=="Dis")
          {Label1 <-"TermCD_M1"
          Suffix1 <- substr(PX1,start=10,stop=NC1)}
          if(Prefix1=="CAP"){R1 <- R1}
          else{
          TermCD1 <- paste(Label1,Suffix1,sep="_")
          R1 <- R1[R1[,TermCD1]==1,]}
          }
# R1 is now Data Set with all variables TermCD 1

if(TVar=="ER_1Y"){P1 <- R1[,"CAPM_1YER"]}
else if(TVar=="ER_6M"){P1 <- R1[,"CAPM_6MER"]}
else if(TVar=="ER_3M"){P1 <- R1[,"CAPM_3MER"]}
else if(TVar=="ER_1M"){P1 <- R1[,"CAPM_1MER"]}
else if(TVar=="ER_1W"){P1 <- R1[,"CAPM_1WER"]}
else if(TVar=="ER_1D"){P1 <- R1[,"CAPM_1DER"]}

P2 <- R1[,"Discount_360D_BB"]
P3 <- R1[,"Discount_LTVol_BB"]
P4 <- R1[,"Discount_GARCH_BB"]
```

```
P5 <- R1[,"ROE_360D_BB"]
P6 <- R1[,"ROE_LTVol_BB"]
P7 <- R1[,"ROE_GARCH_BB"]
P8 <- R1[,"Discount_360D_CS"]
P9 <- R1[,"Discount_LTVol_CS"]
P10 <- R1[,"Discount_GARCH_CS"]
P11 <- R1[,"ROE_360D_CS"]
P12 <- R1[,"ROE_LTVol_CS"]
P13 <- R1[,"ROE_GARCH_CS"]

y <- R1[,TVar]

DF <- data.frame(CAPM=P1,
                   Discount_360D_BB=P2,Discount_LTVol_BB=P3,Discount_GARCH_BB=P4,
                   ROE_360D_BB=P5,ROE_LTVol_BB=P6,ROE_GARCH_BB=P7,
                   Discount_360D_CS=P8,Discount_LTVol_CS=P9,Discount_GARCH_CS=P10,
                   ROE_360D_CS=P11,ROE_LTVol_CS=P12,ROE_GARCH_CS=P13,Y=y)

DF1 <- DF[1:K,]
regit.full <- regsubsets(Y~.,data=DF1,nvmax=NULL,method="seqrep",nbest=1)
sub.graph(regit.full)
#return(reg.sum)
}
```

Subset selection using logistic regression AIC

```
Var_Select2
function (DS,TVar,K,Label)
{
# Function to perform Regression variable subset selection
Models <- c("CAPM_1YER","Discount_360D_BB","Discount_LTVol_BB","Discount_GARCH_BB",
                            "ROE_360D_BB","ROE_LTVol_BB","ROE_GARCH_BB",
                            "Discount_360D_CS","Discount_LTVol_CS","Discount_GARCH_CS",
                            "ROE_360D_CS","ROE_LTVol_CS","ROE_GARCH_CS")

# first TermCd Check
PX <- Models[1]
NC <- nchar(PX)
Prefix <- substr(PX,start=1,stop=3)
if(Prefix=="ROE")
{Label <- "TermCD_M2"
Suffix <- substr(PX,start=5,stop=NC)}
else if(Prefix=="Dis")
{Label <-"TermCD_M1"
Suffix <- substr(PX,start=10,stop=NC)}
if(Prefix=="CAP"){R1 <- DS}
else{
TermCD <- paste(Label,Suffix,sep="_")
R1 <- DS[DS[,TermCD]==1,]}

for(i in 2:length(Models))
         {
         PX1 <- Models[i]
         NC1 <- nchar(PX1)
         Prefix1 <- substr(PX1,start=1,stop=3)
         if(Prefix1=="ROE")
         {Label1 <- "TermCD_M2"
         Suffix1 <- substr(PX1,start=5,stop=NC1)}
         else if(Prefix1=="Dis")
         {Label1 <-"TermCD_M1"
         Suffix1 <- substr(PX1,start=10,stop=NC1)}
```

```
          if(Prefix1=="CAP"){R1 <- R1}
          else{
          TermCD1 <- paste(Label1,Suffix1,sep="_")
          R1 <- R1[R1[,TermCD1]==1,]}
          }
# R1 is now Data Set with all variables TermCD 1

if(TVar=="ER_1Y"){P1 <- R1[,"CAPM_1YER"]}
else if(TVar=="ER_6M"){P1 <- R1[,"CAPM_6MER"]}
else if(TVar=="ER_3M"){P1 <- R1[,"CAPM_3MER"]}
else if(TVar=="ER_1M"){P1 <- R1[,"CAPM_1MER"]}
else if(TVar=="ER_1W"){P1 <- R1[,"CAPM_1WER"]}
else if(TVar=="ER_1D"){P1 <- R1[,"CAPM_1DER"]}

P2 <- R1[,"Discount_360D_BB"]
P3 <- R1[,"Discount_LTVol_BB"]
P4 <- R1[,"Discount_GARCH_BB"]
P5 <- R1[,"ROE_360D_BB"]
P6 <- R1[,"ROE_LTVol_BB"]
P7 <- R1[,"ROE_GARCH_BB"]
P8 <- R1[,"Discount_360D_CS"]
P9 <- R1[,"Discount_LTVol_CS"]
P10 <- R1[,"Discount_GARCH_CS"]
P11 <- R1[,"ROE_360D_CS"]
P12 <- R1[,"ROE_LTVol_CS"]
P13 <- R1[,"ROE_GARCH_CS"]

y <- R1[,TVar]
Direction <- rep("Down",K)
Direction[y>0] <-"Up"

DF <- data.frame(CAPM=P1,
                   Discount_360D_BB=P2,Discount_LTVol_BB=P3,Discount_GARCH_BB=P4,
                   ROE_360D_BB=P5,ROE_LTVol_BB=P6,ROE_GARCH_BB=P7,
                   Discount_360D_CS=P8,Discount_LTVol_CS=P9,Discount_GARCH_CS=P10,
                   ROE_360D_CS=P11,ROE_LTVol_CS=P12,ROE_GARCH_CS=P13,Y=Direction)

DF1 <- DF[1:K,]
fullmod <- glm(Y~.,family=binomial,data=DF1)
nothing <- glm(Y~1,family=binomial,data=DF1)
backwards <- step(fullmod,trace=0)
#summary(backwards)
forwards <- step(nothing,scope=list(lower=formula(nothing),upper=formula(fullmod)),direction=
"forward",trace=0)
bothways <- step(nothing,scope=list(lower=formula(nothing),upper=formula(fullmod)),direction=
"forward",trace=0)
#summary(bothways)
stargazer(bothways,type="text",out="regression.txt",dep.var.labels=Label)
}
```