**An Investigation into the Non-Linearity of UK Equity Factors**

**ES30029 - Final Year Research Project**

**Barnaby napier (14041)**

2020

**Contents**

[1. Acknowledgements 2](#_Toc32910808)

[2. Abstract 2](#_Toc32910809)

[3. Introduction & Background 2](#_Toc32910810)

[4. Literature Review & Theory 3](#_Toc32910811)

[4.1. Efficient Market Hypothesis 3](#_Toc32910812)

[4.2. Factor Models 3](#_Toc32910813)

[4.2.1. CAPM (the market factor) 3](#_Toc32910814)

[4.2.2. Fama French 3 Factor Model 3](#_Toc32910815)

[4.2.3. Fama-French 5 factor model 4](#_Toc32910816)

[4.3. Market Anomalies 4](#_Toc32910817)

[4.3.1. Momentum 4](#_Toc32910818)

[4.3.2. Volatility 4](#_Toc32910819)

[4.4. Non-Linearity 4](#_Toc32910820)

[4.4.1. Non-Linear Regression Models 4](#_Toc32910821)

[5. Analysis 4](#_Toc32910822)

[5.1. Data 4](#_Toc32910823)

[5.1.1. Collection & Sources 4](#_Toc32910824)

[5.1.2. Exploratory Analysis 5](#_Toc32910825)

[5.2. Models & Methodology 6](#_Toc32910826)

[5.2.1. Linear Regression 6](#_Toc32910827)

[5.2.2. Non-Linear Models 6](#_Toc32910828)

[5.3. Results 7](#_Toc32910829)

[5.3.1. Results 7](#_Toc32910830)

[5.3.2. Limitations 7](#_Toc32910831)

[6. Conclusion 7](#_Toc32910832)

[7. Bibliography 7](#_Toc32910833)

[8. Appendix 7](#_Toc32910834)

1. Acknowledgements

* Patrick
* Andrzej

1. Abstract

* What does the paper investigate and conclude

1. Introduction & Background

For as long as the stock market has existed, its participants have been searching for ways to accurately assess the expected return of a stock. The main motivation is likely to make money, but beyond this an accurate evaluation of a stock’s expected return can be informative for the company’s management in choosing a discount factor for projects it is considering undertaking. Along the same lines, a security’s expected return can be used to estimate the cost of capital of a company and hence determine its discounted fundamental value today.

The theory of equity factors or styles is that the expected return of a stock can be decomposed into a function of various factor exposures.

Where each is a factor exposure for stock .

There are three broad causes of return factors (Pioch, 2018). Firstly, there is compensation for risk, in which an investor earns a higher expected return in exchange for taking mr risk. For example, the illiquidity risk of owning a small company leading to the size effect. Secondly, there are factors caused by behavioural biases of market participants. Our tendency to overweight recent observations, disproportionately fear losses and idolise a compelling story lead to some of the persistent inefficiencies that we observe in the markets today. Thirdly, equity risk premia can be a product of market structure. An example of this is the January effect, where fund managers sell stock in December to lock in their gains, then buy it back in January in order to make returns for the new year, leading to lower returns in December and higher returns in January.

Beyond the drivers mentioned above, some equity factors have their roots in financial economic theory. Fama and French (2015) use the dividend discount model to show that the book to market effect is engrained into traditional and commonly used asset pricing models.

either explain theoretical model or remove it ^^

As will be elaborated on in the following section, equity factors have generally been modelled and tested using a linear specification such as the one shown above. The simplicity of the linear model makes it easy to interpret and hence an attractive choice for the estimation of equity factors, but that simplicity may in turn sacrifice accuracy. We may wish to use more advanced statistical models to estimate the expected return of a stock, allowing for the relaxation of the linearity constraint and also allowing for interaction terms between predictors.

Some analysis of this type has been done over the years, with the idea for this thesis being born out of a recent publication by Gu, Kelly and Xiu (2019). Given the larger dataset in the US, most of the research on this topic has been conducted there. I believe that there should be an attempt made to do the same on UK data, allowing for a comparison of non-linear multi factor models across the two countries. That being the case, UK data has a shorter time scale and less detail, which this thesis acknowledges.

The rest of this paper is broken down into three main sections. Firstly, section 4 provides a summary of the research that has been done on equity factors and their nonlinearity. Following this, section 5 elaborates on the models applied in this thesis, as well as concluding the key findings from their application.

1. Literature Review & Theory
   1. Efficient Market Hypothesis

As a precursor to the literature on equity factors. It is first important to acknowledge the implications for market efficiency. Eugene Fama (1970) proposed the idea of the efficient market hypothesis (EMH) in which he posited that at any given moment, prices will fully reflect all information that is available to the market. Consequently, assuming that the discovery of new information is a random process, stock returns must also follow a random process.

Depending on the driver of excess returns, an equity factor may or may not contradict the EMH. Of the three drivers of risk premia mentioned above, compensation for risk is undoubtedly in line with the EMH given the risk-return trade-off of the markets. The other two drivers however create the potential for market participants to achieve risk adjusted returns that are consistently higher than the market, and therefore contradict EMH.

The implications of the EMH for equity factors are as follows; no investor, statistician or trader can use the set of publicly available information to correctly and consistently forecast asset returns.

* 1. Equity Risk Premia
     1. CAPM (the market factor)

The first model that proposes stock returns are composed of a set of factors is the Capital Asset Pricing Model (CAPM) put forward independently by William Sharpe (1964) and John Lintner (1965). It states that the excess return of a stock above that of the risk-free rate is equal to a stock specific measure of risk, beta, multiplied by the risk premium of the stock above that of the market.

The beta of a stock represents its riskiness relative to the market and can be defined by equation 2 below, or more commonly by the regression coefficient produced when running OLS on equation 1.

graph from financial markets lecture

reason for factor

* + 1. Fama French Three-Factor Model

A few decades after Sharpe and Lintner, Eugene Fama and Kenneth French (1993) introduced the concept of equity factors by suggesting that stock returns are driven by more than just the equity risk premium. In their three-factor model, Fama and French showed that as well as the market risk premium, stocks are also exposed to the size (SMB) factor and value (HML) factor.

*Size* is represented by the difference in average return on the smallest and largest 20% of stocks, resulting in the “Small Minus Big” (SMB) factor.

*Value* factor is calculated as the difference in average return on the 20% of companies with the highest and lowest book-to-market value, giving the “High Minus Low” (HML) factor.

The Fama French (FF) three-factor model is specified as below.

reason for factor

* + 1. Momentum (Jegadeesh, Titman & the Carhart Four-Factor Model)

Momentum is the observation that stocks with strong past performance will continue to outperform in the future, meanwhile those that performed poorly will continue to underperform. The idea that equity returns possess this serial autocorrelation was first put forward by Jegadeesh (1990), who found a negative autocorrelation at short lags such as one month but positive correlation at longer lags, especially twelve months. This paper was consequently built on by two now influential papers.

Firstly, Jegadeesh and Titman (1993) followed up on the 1990 paper, evaluating the consequences of momentum for stock market efficiency.

*The Carhart four-factor model* followed a few years later. Published by Mark Carhart (1997), the model adds a momentum factor to the FF three-factor model. It does so by including an Up Minus Down (UMD) factor which equals the average return of past winners minus the average return of past losers.

reason for factor

* + 1. Quality (Fama French Five-Factor Model)

Many years later, Fama and French (2015) added to their three-factor model by incorporating a Quality factor exposure. Quality implores managers and investors to buy good companies who invest well and make solid profits.

*Low investment* was pass

*Profitability* also was chosen.

Together these additional factors make up the FF five-factor model specified in equation [].

reason for factor

* + 1. Volatility (Andrew Ang)

Finally, Andrew Ang (2006) showed that companies have an exposure to market volatility, and those with a lower sensitivity would have higher expected returns. Specifically, he used the VIX index, a measure of market volatility, to show that companies with less exposure to this index saw statistically significant outperformance in future periods.

In his paper, Ang proposed multiple models for volatility, two of which built off of the CAPM and FF three-factor models respectively, as specified in equations [x] and [x].

reason for factor

* 1. Non-Linearity

So far all of the models put forward use a linear specification to estimate the relationship between the respective factor and equity returns. The following is a summary of the models and respective papers which relax the assumption of linearity, in most cases finding promising results.

* + 1. Generalised Linear Models

where have polynomial regressions been used in equity factors

* + 1. Machine Learning

Although not much of the literature around equity factors has implemented machine learning models, those that did have shown promising results. Gu, Kelly and Xiu (2018) investigate the performance of various machine learning models in forecasting expected returns in the US market. They find that machine learning models have stronger predictive power than linear regression, with neural networks and regression trees being the most performant. The authors identify the advantage of relaxing the nonlinearity assumption, as well as the capacity for a larger predictor set as potential reasons for the strong performance of machine learning models.

Further to this,

* + 1. Deep Learning

Beyond traditional machine learning, research is also being conducted in the area of deep learning, which involves using neural networks with multiple hidden layers for prediction.

The first to investigate how artificial neural networks (ANN) can be applied to equity factors was Levin (1996), meanwhile more recently Nakagawa, Uchida and Aoshima (2018) and Nakagawa (2019) estimate deep factor models of equity returns.

As this thesis falls under the Department of Economics, little will be said of these models, which lie closer to the study of computer science. Furthermore, Gu, Kelly and Xiu (2018) find some evidence that deep learning models do not outperform shallow learning or traditional machine learning on their dataset. The reader is encouraged to survey the aforementioned literature for an idea of the state of research in this area.

1. Data
   1. Collection & Sources

The data used in this thesis is collected from Thomson Reuters DataStream over the period 31/12/1995 to 31/12/2018. It consists of monthly, stock level data for all companies in the FSTE All-Share, conditional on sufficient data availability. Investment trusts, unit trusts and other investment vehicles are filtered out at this stage of the process due to their anomalous characteristics, which are both undesirable and unintuitive when estimating equity risk premia. All predictors are either downloaded directly from DataStream or calculated using data that is. The full list of downloaded and calculated predictors can be found in table A of the appendix.

* 1. Exploratory Analysis

descriptive stats and graphs showing nonlinearity

compare factor returns to market return (graph, sharpe ratio)

A brief graphical investigation into the relationship between various predictors and returns suggests that statistical models which allow for non-linear relationships between independent and dependent variables may be better suited to the problem of asset pricing. The figures below show the relationship between various predictors of asset return and the asset return itself. In most cases the argument can be made that a linear model would fail to effectively fit the respective relationship.

A close up of a map

Description automatically generated

A close up of a map

Description automatically generated

A screenshot of a cell phone

Description automatically generated

Secondly, quadratic and cubic polynomials of each factor are statistically significant when used as independent variables to predict excess returns. In fact, in the case of the SMB factor these higher order polynomials are more significant than the linear form of the variable.

A close up of text on a white background

Description automatically generated

Finally, correlation matrices over various characteristics of stocks suggest that there may be interaction effects between them. The main results from these matrices can be rationalised from what is known about equity factors: Small cap has more momentum

A screenshot of a cell phone

Description automatically generated

redo this correlation matrix with new variables and data (eg. factors)

1. Models
   1. Linear Regression

Before investigating the performance of non-linear models, it is necessary to first estimate the linear model in order to provide a benchmark for their performance.

The Ordinary Least Squares (OLS) model aims to minimise the sum of squared residuals (SSR) as defined in equation [x].

Given that all residuals are weighted the same, large outliers can have a disproportionate effect on the performance of the model and its predictions.

* 1. Polynomial Regression

The first nonlinear model implemented is the polynomial regression, which builds on the linear regression by allowing for quadratic and cubic polynomials of variables to be used as predictors.

Firstly, a cubic polynomial predictor set is created from the existing data simply by including all variables up to the third power. Each model specified above is expanded to include all predictors up to a cubic power, hence permitting a nonlinear relationship.

why is cubic polynomial used instead of quadratic or fourth power?

The goal of the cubic polynomial regression is the same as that of the linear regression, namely, to minimise the sum of squared residuals. Hence outliers are just as important in this scenario, however the idea is that by permitting a nonlinear relationship between predictors and returns the model will be a better fit.

* 1. Lasso Regression

The lasso regression is a form of penalised regression in which a shrinkage factor attempts to limit the number of predictors incorporated into the model. The same predictor set is used as the polynomial regression, ensuring that this model can incorporate nonlinearities as well. The lasso regression implements L1 regularisation, which means that it includes a penalisation term which applies to the sum of the absolute values of predictor coefficients. Mathematically, instead of minimising the SSR, the optimisation aims to minimise equation [x]

The shrinkage term, , determines the severity with which additional predictors are penalised. A higher value of will limit the size of the predictor set more strongly than lower values, and so the choice of is important.

* 1. Ridge Regression

The ridge regression also falls into the category of penalised regressions. It is similar to the lasso, but instead of using L1 regularisation to constrain the number of predictors it implements L2 regularisation. L2 regularisation means that instead of considering the sum of the absolute values of coefficients the model uses the sum of the squared coefficients. This results in an optimisation problem which aims to minimise equation [x].

* 1. Regression Tree

The final model used is a regression tree.

1. Methodology

This section will discuss how each model is constructed, as well as the predictor sets used and the rationale for doing so. The models summarised in the literature review of section [y] are used as the foundation upon which this thesis builds. For the purposes of this thesis, each model is specified as below, with the addition of nonlinearities from section [y] onward.

FF Three-Factor Model

Carhart Four-Factor Model

FF Five-Factor Model

FF Three-Factor Model with Volatility

Traditionally, a process of k-fold cross validation would be used to train the model and evaluate its true performance. K-fold cross validation involves dividing the dataset into segments, or “folds”. One by one each fold is temporarily removed from the dataset. The model is then trained on the other folds and tested on the fold that was left out. Once each fold has been tested, an average performance can be taken over all folds and a fair evaluation of the model’s performance made.

This approach is commonly used in situations where the ordering of data does not matter. However, it is not applicable to time series data given that it does not account for the intertemporal relation between different folds.

rolling-origin evaluation (Tashman 2000)

When evaluating each model, is used as a measure of performance. is defined as the percentage variation in the dependent variable that is explained by the model and is specified by equation [x].

1. Results
   1. Model Performance

mean squared error (MSE)

r-squared

results from k-fold cross validation

results from other fitting methods

comparison of linear to non-linear

discussion of regression coefficients

* 1. Limitations of Results

Data availability

might not be economically significant (trading costs)

1. Conclusion

Bibliography

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

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

Fama, E. and French, K., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), pp.3-56.

Gu, S., Kelly, B. and Xiu, D., 2018. Empirical Asset Pricing Via Machine Learning. *SSRN Electronic Journal*.

Fama, E. and French, K., 2015. Dissecting Anomalies with a Five-Factor Model. *Review of Financial Studies*, 29(1), pp.69-103.

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

Ang, A., Hodrick, R., Xing, Y. and Zhang, X., 2006. The Cross-Section of Volatility and Expected Returns. *The Journal of Finance*, 61(1), pp.259-299.

Fama, E., 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance, 25(2), p.383.

Pioch, A., 2018. *How market frictions live in harmony with persistent factor premia*. [online] Macromatters.lgim.com. Available from: https://macromatters.lgim.com/categories/portfolio-thinking/how-market-frictions-live-in-harmony-with-persistent-factor-premia [Accessed 25 Feb. 2020].

Jegadeesh, N., 1990. Evidence of Predictable Behavior of Security Returns. The Journal of Finance, 45(3), pp.881-898.

JEGADEESH, N. and TITMAN, S., 1993. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance*, 48(1), pp.65-91.

Gu, S., Kelly, B. and Xiu, D., 2018. Empirical Asset Pricing Via Machine Learning. *SSRN Electronic Journal*.

1. Appendix
   1. Python Code

The code used to run the regression models and machine learning models. Programmed in Jupyter notebook.

* 1. Variables

The following variables are sourced directly from Thomson Reuters DataStream.

|  |  |
| --- | --- |
| **Name** | **Description** |
| *ret* | 1-month price return in excess of risk free rate |
| *mv* | Market value |
| *allshare* | Monthly price return of FTSE All Share |

The following variables were calculated using the data in the above table.

|  |  |
| --- | --- |
| **Name** | **Description** |
| *ret\_12m* | 3-month price return in excess of risk free rate (gilt3m) |
| *gilt3m* | The annualised interest rate on 3-month UK government bonds. Used as risk free rate |
|  |  |