**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 .

When all is said and done, the stock market is the coming together of countless individuals and machines. Equity factors are to some extent a result of the human element of financial markets, 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.

Fama and French had shown that the much of a fund manager’s performance can be explained through their models

Beyond human biases, 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 paper 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 paper acknowledges.

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. This idea has significant implications. It means that there exists no trading or forecasting system which can achieve a risk-adjusted expected return in excess of the equilibrium return, and hence no investor can beat the market over the long run without taking sufficient risk.

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.

does the idea of equity risk premia go against the emh?

* 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.

* + 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.

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

* + - 1. Size Effect

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.

* + - 1. Value Effect

The 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.

* + 1. Momentum (Carhart, Jagadeesh, Titman)
* The idea that high past returns will lead to high future returns and vise versa (12-month time scale)
* Momentum implies that market returns have some memory (i.e there is autocorrelation in stock returns)
* Carhart four factor model 🡪 Up minus down (UMD)
* Jagadeesh + Titman

who actually came up with momentum?

* + 1. Fama French Five-Factor model

Building on their three-factor model, Fama and French (2015) added to this the factors of profitability and low investment, together incorporating a Quality factor exposure to their model.

* + 1. Volatility
  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. Non-Linear Multi-Factor Model

As early as the mid-ninety’s researchers were considering a non-linear form of the recently published FF three-factor model. Levin

* + 1. Machine Learning
* Not much of the literature has used machine learning in relation to equity factors
* This paper mentions how ML is used because it allows for non-linearity and more predictors
* The few papers which have used it have found promising results
* Main models that I have found being used in relation to equity factors
  + Random forests & Regression trees (AQR)
  + Generalised linear regression (AQR)
  + Penalised linear regression (AQR)
  + Partial Least Squares (AQR)

Although not much of the literature around equity factors has implemented machine learning models, those that did have seen shown promising results.

* + 1. Deep Learning
* Neural networks (Nakagawa)

1. Analysis
   1. Data
      1. Collection & Sources

This paper uses data collected from Thomson Reuters DataStream. It consists of monthly, stock level data for all companies in the FSTE All-Share with all data available in the respective month, beginning on 31/12/1995. A full list of variables used and calculated can be found in the appendix. Note that investment 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 factors as they are not, in the traditional sense, equities.

* + 1. Exploratory Analysis
  + Stationarity
  + Normality
  + Autocorrelation

A brief investigation into the data 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 below scatter charts show the relationship between various predictors of asset return and the asset return itself. In most cases the argument can easily be made that a linear model would fail to effectively model 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, correlation matrices over different time frames 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

* 1. Models & Methodology
     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.

The Ordinary Least Squares (OLS) model aims to minimise the sum of squared residuals from the regression line of best fit. Consequently, outliers can have a disproportionate effect on the performance of the model and its predictions.

* + 1. Non-Linear Models
* Polynomial regressions
* ML models tested
  + Which ones performed well?
  + Why do you think they performed well?
  + Stats behind the models
  1. Results
     1. Results
* R-squared (expect values to be low as this is financial data)
* Mean squared error (MSE) – Used in Patrick’s lectures so he will like it being in here
  + 1. Limitations
* Data availability
* Economic significance (investors most likely cannot use this to their advantage given then existence of trading costs)

1. Conclusion
2. 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.

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\_3m* | 3-month price return |
| *ret\_6m* | 6-month price return |
|  |  |