SMART BETA FACTOR INVESTING

Is the vast amount of equity return factors economically justifiable with respect to smart beta investing from an investor perspective and to what extent do macroeconomic forces explain Smart Beta performance?

Master's Thesis

M. Sc. Economics and Business Administration
Finance and Investments (FIN)

Authors:
Sara Hassanzadeh (114642)
Felix Gerrit Lüdtke (115059)

Supervisor:
Dr Christian Rix-Nielsen
Department of Economics and Business Economics

Number of characters/pages: 233,841/111
Date of Submission: 13-09-2019

Copenhagen Business School, September 2019

Abstract

This paper guides the reader through the Smart Beta "factor zoo" to unveil the notably *smartest* strategy based on a risk-adjusted analysis. The strategies analysed are Dividends, Earnings-Weighted, Equal-Weighted, Fundamentals, Growth, Low Volatility, Value Momentum, Multi-factor, Non-traditional and Quality. A gap in literature is filled by examining the relationships between the return performance of these eleven SB ETF strategies and a series of eight macroeconomic indicators. An empirical study, consisting of 327 ETFs form the crucial basis of whether US Smart Beta ETFs are able to beat their passive benchmark. Two weighting schemes - equally- and size-weighted portfolios – are constructed to compare SB performance with a raw and risk-adjusted benchmark.

Only two categories, value and multi-factor are found to consistently outperform in both the equally- and size-weighted portfolio. However, it is the equally-weighted portfolio that provides more significant alphas relative to its benchmark suggesting a better performance of smaller SB ETFs. Unexpectedly, the well-established Fama-French factors size (SMB) and value (HML) proved to have no explanatory power and as a result are dropped from the analysis. Within the macroeconomic factors, only four (GDP, IPI, UR, CPI) are statistically significant in explaining SB ETF excess returns. A principle component analysis reduced the dimensionality from eight to only three independent variables without sacrificing much explanatory power and verified the importance of GDP and IPI in explaining SB ETF excess returns.

Nonetheless, one of the most critical points throughout literature is confirmed in the analysis. The promising concept of SB fund managers to tilt their portfolios towards specific (or a combination of) equity factors is attractive to investors only in case of few selected strategies. When taking expenses into account, the superior performance of SB factors seems too unreliable to recommend SB ETFs as an attractive investment vehicle to investors. Lastly, SB outperformance seems to be diminishing since inception for most SB factors – as are Smart Beta ETF fees.

Table of Contents

List of Tables

List of Figures

Acknowledgements

We want to thank our supervisor Dr. Christian Rix-Nielsen for the outstanding support throughout our thesis process and always answering our questions patiently. We appreciate the great input that helped us to steer through the ups and downs of our thesis development.

Thanks to Copenhagen Business School we were provided with access to various resources that allowed a successful completion of our final project.

Lastly, we would like to express our endless gratitude to friends and family who helped us along the way. In particular, we want to thank our parents Touran Nourbakhsh and Abbas Hassanzadeh as well as Ingeborg and Gerhard Lüdtke for the amazing possibilities and spiritual support at all times. You always believe in us and enable us to achieve great things in life.

List of Abbreviations

APT:	Arbitrage Pricing Theory
AUM:	Assets under Management
B/M:	Book-to-Market Equity
BMO:	Bank of Montreal
CAPM:	Capital Asset Pricing Model
CPI:	Consumer Price Index
E/P:	Earnings-Price Ratio
EMH:	Efficient Market Hypothesis
ETF:	Exchange Traded Fund
FDI:	Foreign Direct Investment
FF:	Fama French
GDP:	Gross Domestic Product
GNP:	Gross National Product
HAC:	Heteroscedasticity- and Autocorrelation-Consistent
HC:	Heteroscedasticity-Consistent
HML:	High Minus Low
IMF:	International Monetary Fund
IPI:	Industrial Production Index
IFR:	Information Ratio
IR:	Interest Rate
MS:	Money Supply
OLS:	Ordinary Least Squares
OP:	Oil Price
PC:	Principal Component
PCA:	Principal Component Analysis
SB:	Smart Beta
SEC:	Securities and Exchange Commission
SMB:	Small Minus Big
SoR:	Sortino Ratio
SR:	Sharpe Ratio
TFM:	Three-Factor-Model

TSS:	Total Sum of Squares
UR:	Unemployment Rate
US:	United States
VAR:	Vector Autoregression
VIF:	Variance Inflation Factor

1 Introduction

1.1 Background

One of the fundamental discussions in the universe of financial economics derives from rather polarized opinions to what extent capital markets are, in fact, efficient. On one hand, prominent researchers such as Malkiel & Fama (1970) have long argued that all information is reflected in current prices. On the other hand, the Efficient Market Hypothesis (EMH) is often challenged. Successive discoveries of anomalies in the following decades erodes the general belief in this theory. Shiller (2003) outlined evidence for excess volatility of returns and investor exuberance, all of which consequently spurred research on behavioural finance topics. The stock market crash 1987, the Dot-com bubble end of the 20th century and the housing bubble several years later portray quite some evidence in regard to EMH inconsistencies (Jones & Netter, 2008). It seems that - at least transitorily - market mispricing exists. Nonetheless, advocates of both sides have presented convincing evidence in favour and against the EMH. As with many debates, the truth is likely to be found somewhere in between the two extremes. Hence, capital markets are not perfectly efficient, neither are they completely inefficient.

The existence of market efficiency is a much-debated topic because its true nature certainly has far reaching implications for participants of the financial markets. Assuming capital markets were efficient, actively managed investment strategies seem unfavourable. Naturally, passive investments that follow the market would be preferable. Vice versa, upon condition that market inconsistencies exist it is smarter to invest in actively managed funds to have arbitrage opportunities identified and exploited by fund managers. This opens up the frequently discussed debate to whether active equity funds are able to outperform and beat the market.

According to Gruber (1996) the performance of actively managed mutual funds has been on average inferior to that of index funds. More recent evidence is provided by Newlands & Marriage (2016) who claim a staggering 99 per cent of all actively managed United States (US) equity funds sold in Europe have underperformed the S&P 500 over the past decade. As a result, active fund managers might find it hard to convince prospects to pay higher fees in exchange. Furthermore, the performance of the few successful managers who do outperform its benchmark quickly vanishes when taking expenses into account. This is the starting point for passively managed funds, which are strongly rising in popularity. Since 1995 these low-cost index funds doubled its net asset value annually and as a result fuel global demand for passive investment strategies (Gastineau, 2001).

1.2 Development of Smart Beta Funds

Exchange Traded Funds (ETFs) constitute the beginning of passive investing through tracking various indices. The well-known S&P 500 index was the first kind of its form and marked the beginning of a new asset class which would grow tremendously over the following decades. By 2018, total net assets of all index-based US ETFs amounted to over $5 trillion. In terms of daily trading volume, ETFs account for almost a quarter of all trades in the US stock market (Evans & Wilson, 2018).

One of the fastest growing segments of ETF market are Smart Beta (SB) ETFs whose popularity in recent years has spiralled upwards. The first SB fund was launched in the US in 2000, seven years after the S&P 500 index. According to Morningstar (2018), the net new inflows into the global SB ETF market in 2018 alone amounted to $87 billion. As of December 2018, the US market offers over 1493 index-based SB ETFs representing collective assets under management (AUM) of approximately $797 billion. After a record high of 257 newly issued funds in 2017, the market seems to be more saturated with only 132 new SB funds brought to market in the following year. Nonetheless, this results in an organic growth rate of nearly 11%. Due to the sophisticated financial infrastructure in the United States, most of the funds are listed here. Europe, in contrast, is still in its early years with an 8 per cent share of global SB ETFs. Nonetheless, the number of new Smart Beta product launches was higher in Europe than all other regions combined for the first time in 2018 so that one can expect a decline of the 88% global market share of US domiciled strategic beta ETFs (Boyadzhiev, Bryan, Choy, Johnson, & Venkataraman, 2017).

1.3 Idea of Smart Beta

In the fast-paced environment of financial markets Smart Beta (SB) are characterized as a rather novel phenomenon where essential characteristics are derived from traditional ETFs. However, instead of a cap-weighted[1] index Smart Beta ETFs apply alternative index construction rules. While most investment strategies aim to generate return while keeping the risk component at reasonable levels, Smart Beta claims to do so in a highly cost-efficient way. Despite having numerous catchy titles, *Smart Beta, Enhanced Beta*, *Alternative Beta* or *Strategic Beta* essentially all imply the following: to manage a relatively passive (low turnover) portfolio to capture high returns without bearing additional risk through tilting the portfolio towards some direction such as value versus growth stocks (Malkiel, 2014). In other words, these funds possess an active trait without requiring a fund manager to constantly adjust the portfolio. Thus, these funds depart from established cap-weighting methods and ultimately claim to reap the benefits of active and passive investing.

Since Smart Beta is the most frequently used name, the paper will stick to this phraseology albeit having received much criticism with respect to the semantic interpretation of "smart" in recent literature (Malkiel, 2014).

1.4 Higher Fees for Higher Returns?

Hence, with regard to expense ratios, Smart Beta products are situated between the costs associated with actively managed funds on one hand and cheaper traditional cap weighted "non-SB" ETFs on the other hand. It is noteworthy, however, that the gap between Smart Beta and traditional ETF fees seems to diminish. Whereas the average cost of SB ETFs as measured by the asset-weighted expense ratio was 70 per cent higher than in 2015 (0.41% vs 0.24%), it is only marginally higher in 2017 (0.27% vs. 0.22%) (Boyadzhiev et al., 2017). Due to increased competition, this trend is expected to continue. Yet, significant differences of expense ratios are found within various Smart Beta products. For example, the Schwab U.S. Large-Cap Growth ETF charges its investors 0.04% according to its prospectus while Barclays ETN+ Select MLP ETN expense ratio is more than 23 times higher (0.95%). This still leaves hundreds of millions of fees available for Smart Beta product providers. Consequently, the question arises to whether these products are capable to outsmart traditional ETFs through factor investing. Are higher fees justified on a risk-adjusted basis?

1 A capitalization-weighted (or cap-weighted) index, also referred to as a market-value-weighted index, weighs its components according to the total market value of their outstanding shares.

Most, if not all SB ETFs, claim to minimize the risk component compared to traditional ETFs by tilting their portfolios towards several rewarding factors. A main objective of the paper is to assess the risk-adjusted performance of SB products relatively to their traditional benchmarks in order to shed light on the justification of Smart Beta fund fees. Furthermore, it is important to identify the actual exposure towards intended and unintended exposures because critics suggest that higher returns of SB ETFs are based on an increased risk exposure. Malkiel (2014) proposes that tilting a portfolio towards one direction results in a less diversified portfolio that exposes investors to a much higher degree of risk compared to the broad market portfolio. This additional risk-taking might be what generates above average returns. Moreover, SB ETFs are argued to not produce alphas when compared to multi-factor risk models such as Fama-French (FF) Three-Factor Model (TFM). Malkiel (2014) argues that timing and the surrounding market valuations of a particular strategy decide whether SB strategies outperform. For example, value strategies were particularly successful during the dot- com bubble when growth stocks were priced extremely richly relative to value stocks. Naturally, as strategies experience higher demand due to its popularity and (short-term) success the corresponding price of the underlying “favoured” stocks will likely increase. Therefore, performance of various factor strategies might prove disappointing. It is unlikely, essentially impossible, to identify a dominant strategy irrespective of these valuation relationships. Likewise, Hsu, Kalesnik & Viswanathan (2015) outline that the time-varying nature of expected returns, volatilities and correlations can quickly result in misleading results when assessing historical samples. Lastly, idiosyncratic risk or unsystematic risk preferences inherent among investors do not allow for a one-size fits all strategy. As a result, plentiful Smart Beta strategies are offered in the marketplace. In order to understand the time-varying performance of various strategies, the paper attempts to identify why and when certain factors tend to outperform.

Table I: United States Ranking of Strategic-Beta ETPs by Secondary Attribute

Secondary Attribute	# of ETPs	Assets ($bn)	% of Assets
Value	43	150.1	24.1
Dividend Screened/Weighted	111	150.0	24.1
Growth	34	140.0	22.5
Multi-Factor	215	46.0	7.4
Low/Minimum Volatility/ Variance	28	38.1	6.1

Equal Weighted	55	34.9	5.6
Fundamentals Weighted	12	18.7	3.0
Non-Traditional Fixed Income	23	10.2	1.6
Momentum	28	9.8	1.6
Non-Traditional Commodity	49	7.9	1.3
Quality	13	6.4	1.0
Earnings Weighted	6	3.2	0.5
Multi-Asset	7	1.7	0.3
Revenue Weighted	6	1.6	0.3
Buyback/Shareholder Yield	5	1.5	0.2
Risk-Weighted	11	1.3	0.2
Low/High Beta	3	0.2	0.0
Expected Returns	1	0.2	0.0

Source: Morningstar, 2018

1.5 Universe of Smart Beta Strategies

Currently, the universe of SB ETFs can be divided into eighteen different strategies (Table I) with a clear dominance of multi-factor models and dividend weighted strategies according to Morningstar (2018). In theory, there does not seem to be a clear boundary of how many new factors can be proposed to (supposedly) outsmart the market assuming capital markets are not perfectly efficient. Indeed, empirical evidence suggests many allegedly premium-bearing factors attempt to find patterns of mispriced securities. Technological progress has allowed to identify an enormous number of new factors which Cochrane (2011) refers to as "the factor zoo". According to Hsu et al. (2015) approximately 250 factors can be found in reputable journals with recent experience suggesting this number to increase by 40 factors each year. This is a rather counterintuitive trend with respect to the EMH which suggests the number of equity return factors to decline over time. However, 20 years ago, there were only five equity factors (the market, value, small-cap, momentum, and low beta factors). Hence, scepticism of recent developments seems adequate. The vast amount of factor publications is often criticized as the result of intensive data mining (Hsu et al., 2015). The key challenge lies in identifying an actual return factor that is able to continuously outperform a traditional capitalization-weighted market beta.

Similarly, Harvey, Liu & Zhu (2016) assessed the factor zoo and came to the conclusion that any factor should exceed a t-statistic of 3.0. Thus, most factors are deemed significant when they in fact are simply the result of excessive data mining in recent decades. Traditional cut-off levels (>2.0) tend to be too low nowadays. Possible reasons include a substantial decrease in costs associated with data mining methods, a lack of new observations (since historical data is commonly used to identify new factors) next to correlation among factors which are largely driven by macroeconomic and market-wide variables. Whereas a handful of rather prominent factors performed well (High Minus Low (Fama & French, 1992), Momentum (Carhart, 1997), Durable Consumption Growth (Yogo, 2006), Short-Run Volatility (Adrian & Rosenberg, 2008), Market Beta (Fama & MacBeth, 1973) no empirical evidence could be found for most of the remaining tilts. Instead, Harvey et al. (2016) suggest macroeconomic variables explain much of the stock movements. In the rather young field of SB ETFs several previous studies have already tested many existing factors both on the US and European market. Often times, SB ETFs did not result in significant outperformance over the entire sample holding period as suggested by Glushkov (2015) and De Meyer (2016).

1.6 Research Gap

Nevertheless, recent literature has not yet taken a macroeconomic perspective in explaining SB ETF returns. To that extent, this study assesses the various investment strategies in terms of their empirical evidence for risk-adjusted performance and additionally links various SB strategies towards eight different macroeconomic metrics, including Gross Domestic Product (GDP), Industrial Production Index (IPI), Money Supply (MS), Interest Rate (IR), Oil Prices (OP), Unemployment Rate (UR), Consumer Price Index (CPI) and lastly Foreign Direct Investment (FDI). Moreover, both a stepwise regression and a Principal Component (PC) regression attempt to identify to what extent macroeconomic variables explain and affect SB ETF excess returns compared to established factors such as HML[2] and SMB[3]. For that purpose, a sample of 327 US domiciled domestic equity SB ETFs between June 2000 and December 2018 is constructed and sorted into categories based on its common factor strategy (low volatility, quality, fundamentals weighted, etc.) The relative performance is evaluated against

[2] HML (High Minus Low) is the average return on two value portfolios minus the average return on two growth portfolios (Fama & French, 1993).

[3] SMB (Small Minus Big) is the average return on three small portfolios minus the average return on three big portfolios (Fama & French, 1993).

traditional ETF benchmarks on a risk-adjusted basis. In the next step, different models are constructed for both an equal- and size-weighting portfolio of excess returns to identify the macroeconomic main drivers behind SB ETF performance and answer the research question:

"Is the vast amount of equity return factors economically justifiable with respect to Smart Beta investing from an investor perspective and to what extent do macroeconomic forces explain Smart Beta performance?"

In order to implement the risk-adjusted performance analysis while extending it to a macroeconomic factor exposure analysis in a later stage, it must be outlined that the study exclusively focuses on SB ETFs listed on the US equity market. A total of eight macroeconomic factors are investigated with respect to SB ETF excess returns due to their relativity in academia. Lastly, the study only includes US SB ETFs because of the rather short existence of SB ETFs outside of the US. Thus, local and global macroeconomic factors cannot be compared across countries at this point in time whereas it should be noted that data availability likely allows such an analysis in the near future.

1.7 Outline

The paper is organized as follows: Section 3 outlines the theoretical foundation of factor investing, after which well-known multi-factor models such as the Fama-French three factor model and Carhart four factor model will be touched upon. An in-depth comparison of ETFs versus SB ETFs is necessary to discuss the universe of SB strategies, both traditional and alternative SB factors before introducing the aforementioned macroeconomic variables. Next, section 4 presents the empirical framework used throughout the analysis after which section 5 describes sample collection and outlines the specific research method. Section 6 presents main results including the descriptive statistics, risk-adjusted performance of various SB strategies and macroeconomic factor exposure of SB strategies. Furthermore, section 7 discusses previously found results and attempts to link our findings to previous studies and provide certain recommendations for investors. Next, limitations (section 8) as well as implications for future research (section 9) are presented. Section 10 concludes.

2 Theory

2.1 Efficient Market Hypothesis

Fama, Fisher, Jensen, & Roll (1969) first introduced the concept of efficient markets and theorized around how prices of securities and other assets adjust when new information is released to markets. An efficient market was originally defined as one that *"[…] adjusts rapidly to new information"*. Revising this definition two decades later as data transmission had become visibly seamless, it allowed Fama (1991) to redefine the hypothesis of prices to *"[…] fully reflect all available information"*. The contemporary cornerstone of asset pricing disregards the case of asymmetric information, making it impossible to achieve excess returns through stock market expertise rather than success, since all available information is reflected in the price of securities. The only factor that can influence the price of a security is tomorrow's news. News is unpredictable by definition, and, thus, resulting price changes must be unpredictable and random. In other words, even if uninformed investors bought a diversified portfolio at current prices given by the market, they would obtain a rate of return as benevolent as that achieved by financial experts (Malkiel, 2003). Summarized the idea of an efficient market is built on two pillars: 1) in efficient markets, available information is already incorporated in stock prices; 2) in efficient markets, investors cannot earn a risk-weighted excess return. Determined by the theory of random walks in stock market prices, the market follows the idea of a random walk, meaning that subsequent price changes of securities will be a random step away from the last registered price (Fama, 1965). It is linked to the definition of an "efficient" market including a large number of rational, profit-maximisers actively competing, while predicting future market movements of individual securities. Stating that the stock price fluctuations are independent of each other whilst having the same probability distribution, the theory is connected to the belief that markets are efficient. Malkiel & Fama (1970) mentions that the following two hypotheses constituting the Random Walk Model are implicitly assumed in the EMH; 1) successive price changes are independent and, 2) successive changes are identically distributed. The latter author classifies three degrees of informational efficiency, including the weak form, the semi-strong form and the strong form. While the weak-form efficient market stipulates that the price of a stock fully reflects all information contained in past prices, the semi-strong form of EMH extends the weak form by not only reflecting past history of prices but all information that is publicly[4] available at time t. The strong form of EMH assumes that share prices are reflective of all information known at time t, public as well as private (Jensen, 1978). Again, public and private information is incorporated in share prices by the time it is released. Malkiel & Fama (1970) defines the strong form of EMH as one where investors can earn comparable returns to professionally managed funds by building their own portfolio without any expert guidance. Note that throughout this paper, the semi-strong EMH is assumed to hold.

2.2 Capital Asset Pricing Model

Jack Treynor (1961), William Sharpe (1964), John Lintner (1965) and Jan Mossin (1966) marked the birth of the Capital Asset Pricing Model (CAPM) introducing the problem of optimal portfolio selection in the 1960's. Building on the earlier work on the model of portfolio choice developed by Harry Markowitz (1952), it assumes risk-averse investors that are choosing "mean-variance-efficient" portfolios aiming to 1) minimize the variance of portfolio return, given expected return, and 2) maximize expected return, given variance. Arising from Markowitz's (1952) portfolio theory - often called a "mean-variance model"- the CAPM provides a simple one-factor asset pricing model which attempts to capture excess market return. Predicting the relation between risk and expected return means that the return of any given security depends on its exposure to the systematic risk factor (Fama & French, 2004). Sharpe (1964) and Lintner (1965) extend Markowitz's model by two key assumptions. The first one implies a borrowing and lending rate at a risk-free rate, which is the same for all investors and independent of the amount borrowed or lent. The second assumption indicates that all investors have homogeneous expectations leading to identical probability distributions for future return; i.e. total agreement on the distribution of asset returns from $t-1$ to t.

Putting the CAPM into numbers, the formula for calculating the expected return of an asset given its risk is defined as follows:

$$\bar{r}_a = r_f + \beta_a(\bar{r}_m - r_f) \tag{1}$$

[4] Public information covers capital market information as known from the weak-form EMH as well as non-market information such as earnings, dividend announcements, price earnings ratio, information about the economy and political news.

The variables are defined as followed; $\bar{r}_a$ is the expected return of asset a, r_f the risk-free rate of return, β_a stands for the systematic risk factor beta and $\bar{r}_m$ equals the expected return of the market. Individual investment contains two types of risk, which are systematic and unsystematic risk. Systematic risk, also known as "undiversifiable risk" or "market risk" is the risk inherent to the entire market. It is not only unpredictable but also impossible to mitigate through diversification. Unsystematic risk, "specific risk", or idiosyncratic risk is exposed to individual stocks and can be reduced through diversification. In other words, it represents the proportion of a stock's return that is uncorrelated with general market movements. Modern portfolio theory shows that systematic risk cannot be eliminated; even not with a portfolio containing all the shares in the stock market. CAPM measures the exposure of systematic risk, represented by β in (1). That is to say, that CAPM captures the amount of risk premium $(\bar{r}_m - r_f)$ investors demand in order to accept a riskier asset. Allowing risk-averse investors to invest in the risk-free rate, enables them to secure any component of its portfolio with a fixed return and thus, ensures a minimum level of return for other riskier assets. Only securities exceeding risk-free rate are classified as an attractive investment. The above equation is interpreted as a regression in the following:

$$R_{it} - R_{ft} = \alpha_i + \beta_i \gamma_{it} + \varepsilon_{it} \tag{2}$$

Assuming the CAPM to hold, the regression coefficient α_i, in the above time series model, must equal zero. Jensen (1968, 1969) introduced alpha to analyze performance and is until today a widely-used performance measure, to test whether portfolios did or did not beat the market.

2.2.1 Beta: Does it keep its promise?

Measuring a stock's relative volatility, beta is the only relevant measure of a stock's risk according to the CAPM. The market beta is defined as followed:

$$\beta_{iM} = \frac{COV(R_i R_M)}{\sigma^2 R_M} \tag{3}$$

As the formula indicates, beta is measured by the covariance of the asset return with the market return divided by the market's return variance (Elbannan, 2014). Hence, beta measures the sensitivity of the asset's return to variation in the market return. It shows how much the price of a particular stock deviates according to the movements of the stock market. Black, Jensen, &, Scholes (1972) analysed the price movements of the stocks on the New York Stock Exchange between 1931 and 1965 and could confirm a linear relationship between the financial returns of stock portfolios and their betas. By definition, the market beta equals one, indicating that the share price moves exactly in line with the market. A stock swinging more than the market is above one, whereas a stock moving less than the market is less than one. As already explained in the first half of this section, high-beta stocks provide a potential for higher returns; low-beta stocks pose less systematic risk with lower returns. From equation (2) it becomes evident, that systematic risk, attributable to its sensitivity to macroeconomic factors, is reflected in β_i. Non-systematic risk, the unexpected component due to unexpected events that are relevant only to security, is reflected in ε_{it}. In other words, the expected return on an asset depends only on its systematic risk, regardless of how much total risk an asset contains.

Mirza and Shabbir (2005) published a critical review on "*The Death of CAPM*" debating whether beta is an appropriate measure of systematic risk. Defining beta as the "problem child" (Mirza & Shabbir, 2005, p. 38), the stability of this coefficient has been a controversial issue in literature. Several studies exist, amongst which Blume (1971), Baesel (1971), Roenfeldt, Griepentrog, & Pflaum (1978) used different sets of data over various time periods with the aim to observe the change in beta estimates through time. Their outcomes led them to the common conclusion, that betas are not stable. According to Jegadeesh (1992), as well as Fama & French (1991, 1996) betas are not statistically related to returns. Consequently, they declared beta to be "dead" and just continued to encourage literature suspecting the validity of beta in measuring risk.

Not the entire literature is degrading the CAPM and its exposure to market risk. For the last 30 years academics have been debating the merits of CAPM with the common goal to identify whether beta is an appropriate measure of risk. Research supporting systematic risk is provided by Fama & MacBeth (1973), who tested the validity of CAPM by evaluating the relationship between beta and returns. Providing evidence for a significant average excess return of 1.30% per month, they furthermore found an existing positive relationship between beta and monthly returns. Their results are reinforcing the CAPM in the US stock market and made beta a valid measure of systematic risk as a logic consequence.

Howbeit, academics revealed the findings of Fama & MacBeth (1973) to only provide very weak support for a positive risk return trade off due to the fact that the positive risk return relationship found is not significant across sub periods. Contrary to the verifying findings of the latter authors, Reinganum (1981) spreads doubts about beta in his study, as he found that the cross-sectional differences in portfolio betas, and the differences in average portfolio returns are not reliably related. Hence, the returns on high beta portfolio are not significantly higher than the returns on low beta portfolios. Lakonishok & Shapiro (1986) conducted a study of monthly returns of all stocks traded on the NYSE with the outcome that the return on an individual security is not specifically related to its degree of systematic risk but is significantly related to the market capitalization value. Thus, instead of beta only size can significantly explain the cross-sectional variation in return. Similarly, Haugen & Baker (1993) scanned the risk and return characteristics of 1000 US stocks that have large capitalization over all US stock exchanges and markets between 1972 and 1989. Low risk stocks tended to have abnormally high returns, contradicting the relationship between beta and returns as prescribed by CAPM. The anti-CAPM club was joined and led by Fama, who used to be a great supporter of CAPM until the early 90’s. This was a huge setback itself, when Fama & French (1991) found an insignificant relationship between beta and average returns based on their analysis of the monthly average returns of NYSE stocks. In their conclusion, they state that CAPM has no explanatory power in describing the last 50 years of average stock returns, but rather market capitalization and the ratio of book value to market value.

3 Literature Review

3.1 ETF vs. Smart Beta ETFs

Following their main goal to obtain efficient portfolios that maximize return whilst minimizing risk, investors seek to eliminate idiosyncratic risk. For decades now, investors believe or are convinced to realize the aforementioned strategy by replicating indices. In other words, they follow the performance of an index passively, by buying all of its components, or, at least, a representative sample of them. Exchange-Traded Funds are a marketable security who aim to track a stock index, a commodity, bonds, or a basket of assets, with the intention to mimic an index's return. The demand for ETFs has grown markedly. Also referred to as index-based ETFs, they are designed to track the performance of their designated indices. By the end of 2017, the U.S. ETF market reached $3.4 trillion assets with more than 1,832 funds offered, and it remained the largest in the world, accounting for 72 percent of $4.7 trillion in ETF assets around the world (2018 Investment Company Fact Book, 2018, p. 86). They have become one of the most popular investment vehicles for both institutional and individual investors. Their roots can be followed back for more than two decades now, but were first approved by the US Securities and Exchange Commission (SEC) in 1993, when the State Street Global Investors released the S&P 500 Trust ETF[5], known as the "spider", which is still one of the most actively-traded ETFs today.

The evolution to active ETFs accommodates active investing, improving the risk management of passive ETFs when striving for better risk-adjusted returns at the same time. Until 2008, the SEC had only approved ETFs that tracked specific indices. However, in early 2008, the SEC granted approval to several fund sponsors to offer fully transparent, actively managed ETFs meeting certain requirements. Instead of tracking the return of a particular index, actively managed ETFs seek to create a unique mix of investments in order to meet a particular investment objective and strategy. At year-end 2017, 194 actively managed ETFs – with $45 billion in assets – were registered with the SEC as investment companies. "Active" means having a portfolio manager or team who makes investment decisions in the portfolio, also trading on exchanges. It can either be based on an underlying index or consist of individual

[5] A broad-based domestic equity fund tracking the S&P 500 index.

securities. However, during the first decade of ETFs existence, cap-weighted equity indexes have been the main players in the market for equity index products.

In recent years, there has been an increasing number for alternatives to cap-weighted equity indexes launched. The main differentiating factor is that they are constructed using other weighting schemes in order to improve on capitalization weighting leading to the ability to provide investors with “improved beta”. Already more than two decades ago, several academics, such as Ferson, Kandel &, Stambaugh (1987), Haugen & Baker (1991) and Grinold (1992), concluded in their analyses that cap-weighted indexes provide an inefficient risk/return trade-off. Based on this outcome, investors nowadays have a wide array of indexes to choose from.

A recent breakthrough in capital markets is the evolution of Smart Beta investing. However, they are not especially new, since we can trace them conceptually to a 1976 paper on Arbitrage Pricing Theory (APT) by Stephen Ross. In his work, he posited return premiums associated with various risk factors. Since then, researchers have continued to document positive risk-adjusted returns associated with factors beyond (systematic) market risk, including behavioural anomalies. Seeking to fulfil the mission of generating return whilst maintaining a healthy proportion of risk, SB ETFs are also known as “Alternative Indexing” (Jacobs & Levy, 2015). A “Smart Beta” ETF is classified as quasi-active as it encompasses hybrid active/passive strategies and is one of the newest forms to offer the potential for enhanced index returns. As a “rules-based” investment strategy, Smart Beta is similar to active ETFs in that it goes beyond limitations of a traditional, market-cap weighted index to add value to returns, by divorcing a security’s weight in an index from its capitalization weight in a market (Philips, Bennyhoff, Kinniry, Schlanger, & Chin, 2015). The difference between active and SB ETFs is based on the fact, that no individual is making the decisions, since they are fully accomplished by computer models consisting of decades of historical data on which to base investment strategies. In other words, this investment approach, also known as “strategic beta”, passively follows an index but uses alternate criteria to select the weight each security will have in a fund. Still considered as a relatively new concept, there is no clear-cut definition for Smart Beta. A common belief is that using Smart Beta strategies enables achieving greater-than-market returns. According to Malkiel (2014) one has to tilt (or flavour) the portfolio in a specific direction, such as value versus growth, smaller versus larger companies, relatively strong stocks versus weak, and low-volatility stocks versus high volatility ones. Based on this, there are also portfolios blending value and small size and those that mix several of the above-mentioned flavours. SB strategies can be related to multi-factor models of asset pricing. Assuming the CAPM to be an incomplete measure of risk, the flavours listed above might be considered as additional risk factors. Arnott, Hsu &, Moore (2005) were amongst the first ones who publicized the core idea of a SB strategy. Defining the traditional cap-weighted indexes as the “Wall Street” of the size of an enterprise, characteristics such as gross revenue, equity book value, gross sales, gross dividends or and cash flow are classified as “Main Street” measures. As proponents of alternative weighting schemes, the authors describe a group of fundamentals-based market portfolios whose construction method is based on selection and weighting with metrics of company size[6] other than cap weighting. They show that they provide higher returns and lower risks than the traditional cap-weighted equity market indices while retaining many of the benefits of traditional indexing. Their analysis led them to the outcome, that the resulting portfolios outperformed the S&P 500 by an average of 1.97 percentage points in the period between 1962 and 2004. Consequently, Arnott et al. (2005) suggest that indices constructed using Main Street measures of company size are significantly better than the cap-weighted Wall Street indices. Further advocates, such as Kahn & Lemmon (2016) describe the Smart Beta era as a “Disruptive Innovation” and urge investors to shift partly to Smart Beta besides pure alpha managed products, especially because of their lower fee structure. A survey conducted by FTSE Russell (2017a) outlines the strong trend in “cost savings” as an investment objective. On top of that, proponents of the strategy, argue that in the long-run SB should enable investors to capture factor premiums more efficiently because cap-weighted indexes may lack diversification due to holding large positions in potentially overpriced stocks and small positions in potentially under-priced stocks. Amenc, Goltz, Le Sourd, & Lodh (2015) reveal that the main reasons investors are reluctant to choose SB strategies are “lack of familiarity” and “doubts over robustness” because benchmarks and index methodology are often chosen based on attractive back-tested performance history. Aiming for good value for money, an increasing number of asset owners are seeing Smart Beta as a useful tool to hold down costs. Their survey concludes that the shift in assets is due to disappointment in active asset performance, but cost is still top of mind for many asset owners today. This rationale is also

6 The size measures include book value, revenues, dividends and others.

confirmed by Hsu (2004) who argues that cap-weighted portfolios are suboptimal relative to their non-cap-weighted counterparts. The aforementioned findings are affirmed by Arnott, Kalesnik, Moghtader, & Scholl (2010) who find that non-cap-weighted strategies offer superior performance which is proved in their historical dominance in terms of return and/or risk-adjusted returns. The latter authors build their analysis by using a fundamental index (weighted economic scale of the companies), an equal weighted index, a minimum variance portfolio and combinations of these methods. FTSE Russell (2018) just recently published insights of their Smart Beta global asset owner survey for the fifth time. Once again, they observed increasing global growth trends and Smart Beta to be a widely recognized set of new tools with great potential for helping market participants to achieve their goals. Recent institutional surveys suggest that SB ETFs are poised to grow over the next years, because 60% of decision-makers expect to increase their use by contemplating further allocations (FTSE Russell, 2018). This confirms the survey suggestions of Russell Investments (2014), expecting the SB ETFs to grow over the next three years. Nevertheless, it must not be ignored that at the same time, the average cost of SB ETFs measured by the asset-weighted expense ratio was 70 percent higher (0.41 percent vs. 0.24 percent) than that of traditional cap-weighted non-SB ETFs, implying that SB product providers as a whole are charging investors an additional $370 million in fees per year (Glushkov, 2016). Challenging SB strategies to be really "smart", also Arnott, Kalesnik, & Wu (2017) expect really *Smart Beta* strategies to be designed to optimally capture return premiums and to be able to deliver them to investors *after* trading costs. Furthermore, the aforementioned authors highlight the fact that Smart Beta strategies cannot be replicated with simple factor tilts. In fact, a Smart Beta strategy *has* tilts and offers more than just breaking the link between price and portfolio weights. It rather delivers different return and portfolio characteristics from those simple factor tilts and it delivers alpha net of the factor tilts and net of the FF four- or five-factor regressions.

Throughout the literature review of this thesis a landscape of controversial opinions related to the evolution of SB ETFs quickly emerges. This trend seems omnipresent throughout articles by referring to "dumb beta" (Steward, 2014) or "Smart Beta is Not Monkey Business" (Amenc et al., 2015). Even though there is no universal consensus on what Smart Beta is, surveys conducted by Spence Johnson (2014) in Europe and the United Kingdom; Bank of Montreal (BMO, 2014) in Canada; Pensions & Investments (2013) in Asia; AXA Investment Managers

(2013) in Australia; and Cogent Research (now Market Strategies International) (2014) in the United States of institutional consultants and investors across the world agreed on the growing global interest in exploring Smart Beta strategies. Especially since the beginning of 2009, SB continues to attract a steadily increasing share of net flows relative to the rest of the ETF market, reaching an all-time high of nearly 35 percent of all net flows into US-domiciled ETFs in 2013 (Glushkov, 2016). Steward (2014) sheds dubious light on this new investment strategy, as he critically outlines the fact, to distinguish between "smart investing" and "smart trading". According to the latter author, "dumb beta" portfolios keep breaking the relationship between price and weight. Hence, they trade against the mean reversion. This is the main reason why most Smart Beta strategies outperform. By severing the link between the weight and the market price through systematic rebalancing to non-price target weights, many SB strategies may benefit from mean reversion in the factor premiums (Hsu, 2014). To confirm this finding with further literature, Plyakha, Uppal, & Vilkov (2014) also provide a matching insight, that it is not the initial weights of the SB strategy, but rule-based rebalancing that is responsible for the excess return it earns relative to the cap-weighted portfolio. Amongst others, Ung & Luk (2015) suspect the frequent rebalancing to cause increasing transaction costs. Summing it up, attitudes towards Smart Beta beliefs have their roots in the contention about market efficiency versus market inefficiency. Evolved into a euphemism for "beat the market", conforming to Malkiel (2014) SB indexed ETFs attract and count on "dumb" investors to hold portfolios producing inferior returns. On top of that, Malkiel (2014) concludes that many SB ETFs have failed to produce reliable excess return, whereas only a few have "beaten the market" over the lifetime of the funds. In fact, the excess returns should be interpreted as a reward for assuming extra risk, since investors are taking on a different set of risks when departing from the market portfolio. The author doubts Smart Beta portfolios as they represent a "sophisticated mousetrap for investors" (Malkiel, 2004, p. 133). Hence, this strategy can be seen as the object of considerable marketing hype, resulting in *smart marketing*, rather than smart investing.

Given the explosive growth in assets managed and flows received by SB ETFs and the continuous controversial opinions, Glushkov (2016) tries to sound the hypothesis that SB ETFs outperformed their raw (total return) benchmarks out. In order to see how smart SB really is, he uses a sample of 164 domestic equity US-domiciled SB ETFs categorized into fourteen factor-theme categories during May 2003 to December 2014 which he then compares with passive index benchmarks[7]. After concluding that the performance of SB is not great, he also states that none of the SB categories is able to outperform their risk-adjusted benchmarks. Generally spoken, there is no conclusive empirical evidence that the factor exposures have led to robust relative outperformance. The latter author emphasizes, that it is important to keep in mind, that performance of various SB strategies may only be partially attributed to intended factor tilts, as their excess returns might be influenced through potentially unintended factor bets. This thought is further expanded by Jacobs (2015) who is convinced that only a diversity of factors can provide for more consistent performance, as investments are distributed across a range of factors. Thus, SB portfolios, focusing on only one or a few factors, are likely to underperform, sometimes over protracted periods, when the factors on which they focus underperform. The author agrees with the aforementioned findings of Glushkov (2016) and finds that ignoring factors that may be related to the targeted factor can lead do unintended risk exposures. Summarized, Jacobs (2015) states that Smart Beta is not a good alternative to active, dynamic, multi-factor portfolio management. Sharpe (2014) makes clear to be bothered by the numerous Smart Beta definitions and doubts that many of these strategies will be winners in the future. Bogle (2015) is of opinion that SB is a ploy by active managers to recapture assets lost to indexers. If we look at all these negative statements above, Jacobs & Levy (2015) find a well-balanced conclusion when saying that SB strategies may be a useful addition to the range of investment approaches available to investors, but not being a magic formula for increasing returns while reducing risks. As with most controversial topics in life, investors would be better served by a more realistic consideration of the pros and cons of Smart Beta investing.

3.2 Traditional Factors

When evaluating different factors, two main categories are distinguished throughout this work: traditional and alternative strategies. Traditional strategies are used to be more generally accepted and characterized by a longer track record. Considered as the main factors are those used in the three and four-factor model as described by Fama & French (1992) and Carhart (1997) respectively. Despite that, all other factors can be classified as alternative strategies for the purpose of this paper.

[7] Glushkov (2016) uses three types of benchmarks: A self-declared benchmark by the ETF provider; a risk-adjusted version of the self-declared benchmark; a blended benchmark constructed as an annually re-balanced combination of passive existing funds representing the broad stock market and various factor exposures (size and value).

3.2.1 Fama and French Three-Factor Model

Several deviations or "anomalies" from the CAPM were discovered during the 1980s and 1990s. The seminal work of Fama & French (1992, 1993) undoubtedly marked a great turning point in the development of factor models. Their work motivated interest in studying cross-sectional return patterns (Harvey et al., 2016) and has had a substantial impact to the academic literature and professional practice, reflected in citations to these papers and the frequent utilization of their data. The next section of this paper focuses on this topic in more detail.

Of course, it was the work of Fama and French that was originally motivated by the poor empirical performance of the CAPM. However, the empirical contradictions of the CAPM documented in existing literature at that time triggered their interest to study the cross-sectional variation in average stock returns (Fama & French, 1992), and later to present the three-factor model (Fama & French, 1993). Especially, the strong negative relation between *size* (market capitalization) and average return (Banz, 1981); the positive relation between earnings-price ratios (E/P) and average return (Basu, 1983); the positive relation between book-to-market equity (B/M) and average return (Rosenberg et al., 1985); and the positive relation between leverage and average return (Bhandari, 1988).

Fama & French (1992) based their work on using the cross-sectional regression approach of Fama & MacBeth (1973) to investigate the cross-sectional variation in average stock returns associated with the aforesaid variables. The main findings of the paper can be concluded as follows. First, the positive relation between average return and the market beta predicted by the CAPM (Black et al., 1972; Fama & MacBeth, 1973) disappears during the 1963-1990 period for US stocks. Second, the cross-sectional variation in average stock returns associated with size, B/M, E/P, and leverage is captured by size and B/M. In other words, size and B/M are significant in explaining the cross-sectional variation in average return on the US stock market for the data in question, whereas E/P and leverage are redundant.

In the light of the above findings, Fama & French (1993) used the time-series regression approach of Black et al. (1972) to further investigate the explanatory power of size and B/M in capturing the cross-section of average returns. The established FF model well explained 95% variation of the excess return wherein they added two additional factors: SMB, the difference between the return on a portfolio of small stocks and the return on a portfolio of large stocks, and HML, the difference between the return on a portfolio of high book-to-market ratio and the return on a portfolio of low book-to-market equity. Relying their test on 25 portfolios sorted on size and book-to-market equity from 1963 to 1990, they could summarize that size and BE/ME ratio as the two mimic risk factors play important roles in capturing variation in returns. Supporting evidence was also shown by Chui & Wei (1998) who extended their research to five Pacific-Basic emerging markets[8]. Their results led them to the conclusion that a weak relationship was shown between the market beta and average returns while BE/ME gives powerful explanations of cross-sectional variation of expected returns in Hong Kong, Korea and Malaysia, whereas the size effect is significant in all stock markets except Taiwan.

The latter results gave rise to the famous Fama-French three-factor model, which adds two new factors to the CAPM:

$$\underbrace{R_{it} - R_{ft} = \alpha_i + \beta_i\left(R_{Mt} - R_{ft}\right)}_{\text{CAPM}} + s_i SMB_t + h_i HML_t + \varepsilon_{it} \tag{4}$$

where R_{it} is the return of asset or portfolio i at time t, R_{ft} is the risk-free rate of return, R_m is the return of the value-weighted market portfolio, $R_{Mt} - R_{ft}$ is the excess market return, SMB_t is the difference between returns on diversified portfolios and big size stocks, HML_t is the difference between the returns on diversified portfolios of high and low B/M stocks, the coefficients β_i, s_i and h_i are the asset's sensitivity to each of the factors, α_i is the intercept and ε_{it} is the error term at time t.

Fama & French (1993, 1996) have interpreted their three-factor model as evidence for a risk premium, or a "distress premium". The intuition behind this is that contrary to a market-cap-weighted index, which is mainly influenced by large-cap companies, a size-tilted portfolio aims for small-cap companies. From an economic point of view, small-cap companies are seen as more susceptible to financial distress, which is why small stocks have higher betas and higher average returns than large stocks. Breaking down HML, "High" refers to companies with a high book value to market value ratio. Vice versa, "Low" refers to companies with a low book

8 Hong Kong, Malaysia, Taiwan, Korea and Thailand.

value to market value ratio. It is also known as the “value factor” or “value versus growth factor”, since companies with a high book to market ratio are typically considered “value stocks”, whereas companies with low market to book value are classified as “growth stocks”. As research has demonstrated that value stocks outperform growth stocks in the long run, with buying value stocks and shorting growth stocks in a portfolio, investors can earn returns in excess of the market.

Nevertheless, there is considerable debate about the power of the FF three-factor model. Opponents argue that the “distress premium” found in the three-factor model is the result of survivor bias and data snooping. According to Kothari, Shanken, & Sloan (1995) average returns on high book-to-market portfolios are clearly overstated as the data set is more likely to include distressed firms that survive whilst missing distressed firms that fail. Fama & French (1996) refute these arguments, but indeed, do not provide an empirical explanation why “distress risk” is priced. Another criticism is that the model is motivated by a purely empirical conception and the selection of the factors is ad hoc (Bailey, 2005, p. 186). Nonetheless it is important to outline that the majority of the factor models in the literature are subject to similar and other criticisms.

Moreover, the model has difficulties to explain the continuation of short-term returns, known as the momentum effect. The work of Fama and French (1996) lacks on providing a rational risk-based explanation for the momentum effect. However, they argue, that it may be the result of data snooping or survivor bias.

3.2.2 Carhart Four-Factor Model

Jegadeesh and Titman (1993) pioneered in research regarding the momentum effect as they concluded that a significant abnormal return would be realized between 1965 and 1989 if one implements a strategy aiming at buying past winners and selling past losers. This strategy states that by buying stocks that recently have exhibited high returns (winners) and simultaneously selling (shorting) stocks that have exhibited poor returns (losers) based on stock performance in the past three to twelve months, investors can earn returns in excess of the market as a whole. In a later paper Carhart (1997) created a factor mimicking portfolio for the momentum effect based on the work of Fama & French in 1996 for the size and value factor. In his paper, he proposed a four-factor model by extending the three-factor model with the momentum factor:

$$E(R_i) - R_f = b_i\left[E(R_m) - R_f\right] + s_i(SMB) + h_i(HML) + m_i(MOM) \qquad (5)$$

To sum the research on this section concerning the traditional factors up, one still needs to keep in mind, that the described models need more time and further investigation before they can be accepted as a credible theory-based model. Thus, several researchers and academics started to test the three-factor model for areas other than the US and for different time periods. Amongst them, Arshanapalli, Coggin, & Doukas (1998) test the TFM in 18 stock markets, of which 10 are based in Europe, from 1975 to 1995. Their results confirm that size and B/M risk factors are relevant in explaining stock returns both in the US stock exchanges and other markets. Griffin (2002) shows that TFM performs better if risk factors are defined domestically rather than internationally, including the US, Canada, Japan and the UK. According to Moerman (2005), who analyses a sample of stocks coming from 11 countries investigated from 1991 to 2001, highlights the fact, that TFM seems to work well in the European stock markets and confirms the aforementioned finding of Griffin (2002), that the FF risk factors are country-specific. Emerging markets are analysed by Al-Mwalla & Karasneh (2011), who also confirm that size and B/M factors help in explaining variations in stock returns. Last but not least, Fama & French (2012) show along other works (Kothari et al., 1995; Daniel & Titman, 1997; Davis et al., 2000; Taneja, 2010; Eraslan, 2013; Foye, Mramor, & Pahor, 2013, Sehgal & Balakrishnan, 2013; Sharma & Mehta, 2013) that excess returns are explained well by the Sharpe-Lintner model and market beta is always positive since R^2 often exceeds 60%. Furthermore, SMB and HML alone are significantly related to excess returns, but the explanatory power of the model without the market risk premium is significantly lower. Last, the model that provides the best fit is the one including all three risk premiums. In the majority of the above studies, except of Daniel and Titman (1997), R^2 is greater than 90%.

3.3 Factor Zoo

Since the publication of Carhart's four factor model in 1997 extensive research resulted in a sharp increase in equity return factors. Within the last 20 years, the number of allegedly premium-bearing factors skyrocketed from five to several hundred whereas the exact number seems impossible to determine. Hsu, Kalesnik, & Viswanathan (2015) mention approximately 250 factors while they expect an annual increase of 40 factors. A year later, Harvey et al. (2016) reviewed 313 articles with respect to cross-sectional return patterns whereupon they catalogue

316 equity return factors. Irrespective of the exact number of factors out there, the recent developments seem unusual to say the least. Assuming that the EMH holds, all information should be included in pricing financial assets so that one would expect the number of factors to decrease over time rather than to increase. However, empirical literature does find stock mispricing quite frequently (Doukas, Kim, & Pantzalis, 2010; Bird, Menzies, Dixon, & Rimmer, 2011). Keeping this in mind, several questions arise. Firstly (1), what drives recent spikes in identifying new factors? Secondly (2), is it possible that several hundred factors are robust and implementable factors to predict above average risk-adjusted returns? Lastly (3), to what extent can (and will) Smart Beta strategies that focus on one or more of the above factors outperform the capitalization-weighted market portfolio over time?

In order to comprehend the evolution of equity return factors, it is inevitable to understand how the research environment in economics has changed. Back in 1973 when Fama & French published the market model, little research has previously been conducted on cross-sectional expected stock returns. Consequently, finding a factor with a t-statistic of 2.57 sufficiently exceeded the usual cut-off value of 2.0. More than 40 years later, though, many newly added factors found in cross-sectional regression studies lack explanatory power. Hence, the cut-off value of factors is not sufficiently high as Harvey et al. (2016) argue. This likely is an outcome of countless back tests and predictive regressions conducted by thousands of finance professors, professionals and students which forcibly detect positive outliers. The bulk of research papers likely considered overlapping sample periods since it is obviously impossible to choose another data set of historical stock returns than the realized past stock returns. Whereas other research fields have the luxury of generating new observations economics does not. Additionally, to make matters worse, replication studies are less apparent in finance and economics so that there is a bias towards publishing new factors rather than verifying already published factors (Harvey et al., 2016). A shockingly high amount of published results can simply not be replicated due to varying databases and construction methodologies (Hsu et al., 2015).

Out of the 316 factors that are evaluated in the paper by Harvey et al. (2016) most factors are likely to be false discoveries. Only a handful of factors does indeed prove significantly including the market, book to market and momentum factors. In contrast, Green, Hand, &

Zhang (2014) argue that most commonly used factors (size, book-to-market and momentum) miss important economic aspects of cross-sectional stock returns. Rather, firm characteristics such as earnings yield industry-adjusted momentum, stock turnover and quarterly earnings better explain cross-sectional variation in equity returns.

Generally, many prominent researchers agree that (most of) recent publications reveal a staggering number of factors which do not provide future premia due to their data-mined heritage. Instead, many of these factors are simply significant by chance. Alarming voices against a factor proliferation are becoming louder and louder in academia (Harvey et al., 2016; Hsu & Kalesnik, 2014). The severity of this development is best described by Levi & Welch (2014) who assessed 600 factors both from the academics and practitioner literature. They found that 49% of these factors delivered zero to negative premia out-of-sample which makes investment strategies based on these factors just slightly more attractive than tossing coins.

Hence, the proliferation of allegedly uncorrelated factors is as much as alarming issue as it needs to be seriously addressed. Luckily, researchers have already attempted to provide guidance with respect to identifying true factors in the factor zoo. Both Levi & Welch (2014) as well as (Hsu et al., 2015) present similar criteria to sort the wheat from the chaff. For a factor to be declared robust and implementable in terms of Smart Beta investing all of the following characteristics should apply: (i) the factor should be vetted by top academic journals independent of timespan and geographies; (ii) it should not differ as a result of slight changes in the factor definition/construction; (iii) it should have an economically sound reason to produce a consistent premium and lastly (iv) it should exceed higher t-stat significance threshold (>3.5 instead of 2.0) to avoid data snooping among other issues. Summing up, at least partial answers to the previously raised questions during the beginning of the section will be stated.

Firstly (1), recent spikes in the publication of new factors are mainly driven by the higher presence of similar studies in the financial environment paired with technological advances that substantially drove down the cost of research. Moreover, most studies focus on identifying new factors instead of pursuing replication studies.

Secondly (2), as a result of lacking validation of existing equity factors in economics, most factors do not produce a consistent positive premium in the future. Rather, they are data-mined artefacts from historical equity data. Moreover, even if a factor's existence is empirically proven, the implementation of respective factor does not automatically ensure an attractive Smart Beta strategy. More specifically, a passive Smart Beta strategy can provide above average risk-adjusted returns only if excess returns are available in liquid assets with limited portfolio turnover and trading while allowing large in- and outflows. Otherwise active funds are better equipped to exploit these factor-based investment strategies (Hsu et al., 2015). For example, momentum based strategies seem to be better suited for active fund managers because getting ahead of the crowd (uninformed investors) is key in the short holding horizon. In contrast, value and low beta strategies usually require annual turnover of less than 20 percent combined with a low signal decay. Thus, these factors are better qualified and much more likely to be captured in low cost Smart Beta ETFs (Hsu & Kalesnik, 2014). In different words, it is not possible that hundreds of factors are sufficiently robust and implementable to outperform traditional ETFs. Only a fraction of the several hundred factors likely find its way in Smart Beta strategies.

It is impossible to give one straightforward answer to the last question (3). As a starting point, a useful approach seems to rule out less qualified factors. Remember, not all factors look attractive to be implemented as Smart Beta strategies. An even larger number of factors does not seem to produce consistent positive premia. Literature suggests contradictory findings of significant factors. More importantly, factor robustness across time spans seems to pose a key issue inherent in Smart Beta strategies. As a preliminary answer, it looks like the odds for Smart Beta outperformance are lower than one would hope for as an investor. Whether the few remaining factors prove to outperform traditional cap weighted ETFs remains unanswered for the time being. The imminent analysis attempts to shed light on this question by investigating potential equity factors.

3.4 Universe of Smart Beta Factors (Alternative Factors)

In order to fully grasp the complexity of the aforementioned factor zoo, it is inevitable to outline the universe of alternative factors in contrast to the rather well-known established equity factors such as the market factor, size, book-to-market and momentum (section 3.2). We refer to these as traditional factors throughout the analysis due to their long existence and empirical verification. Albeit not all researchers entirely agree on the true nature of these allegedly positive risk premium carrying factors per se, more recently published factors often perceive a greater extent of criticism for contradictory empirical findings. Nonetheless, advocates of SB investing argue that factor premiums can be captured more efficiently in the long run because cap-weighted counterparts may lack diversification due to large holdings of overpriced stocks versus unproportioned holdings of under-priced securities. In other words, "*breaking the link between price and weight creates diversification*" (Steward, 2014, p.4). Whether this in fact applies to SB factors will be examined at a later stage.

In the following, the most frequently used Smart Beta strategies are outlined to provide a first picture about the universe of currently used equity factors. According to Morningstar (2017), 18 different Smart Beta strategies are currently available to investors. Several of them will be outlined in the first part, followed by an introduction to more exotic factors which seem to be hardly used in financial markets (yet?). Note that each factor that is included in the empirical analysis will be shortly presented.

3.4.1 Growth versus Value

In contrast to the value effect, a growth strategy focuses on usually young companies with attractive opportunities. As a result, stocks of these companies are expected to perform well. According to Fidelity (2016) growth can be measured as a surge in revenues, earnings or cash flow. Irrespective of the factor construction, Chan & Lakonishok (2004) discuss the performance difference between value and growth investing and mention advantages and disadvantages of the growth factor. On one hand, it is possible that growth stocks underlie a higher risk level (higher volatility) which is rewarded through higher rewards. On the other hand, it remains questionable whether value or growth stocks are riskier. Again, timing seems to be a main driver here. Whereas investor sentiment reached optimistically high levels towards growth stocks in the 1990s, the dot com bubble lowered expectations substantially. As a result, growth stocks were relatively richly priced compared to value stocks. In recent years, growth stocks are shown to suffer more during recessions and economically challenging periods. Paired with behavioural biases towards growth stock expectations, it is possible that a growth strategy might be rewarded nowadays. However, Kwag, & Whi (2006) find that value portfolios consistently outperform growth portfolios throughout the business cycle. Literature provides evidence for the value effect Fama & French (1998) whereas the growth factor is highly controversial (Arnott, Hsu, Kalesnik, & Tindall, 2013). Summing up, it seems unlikely that growth strategies pay off because a growth tilt is rarely found in Smart Beta products. According to Arnott et al. (2013) the weak performance of growth biased portfolios justifies the low availability of pure growth tilts. Based on our analysis, however, this trend seems to be somewhat reverse with a total of 33 growth tilted SB ETFs that represent 9.73 percent of the whole sample.

3.4.2 Low Volatility

There is a broad arsenal of studies analysing the low-volatility anomaly, stating that contrary to CAPM methodology, excess return can be achieved (Haugen & Baker, 1996). As stated by the risk-return principle, securities' exhibiting low volatility should be associated with lower returns. Nevertheless, several academic papers outline the fact, that a portfolio of stocks with low volatility can indeed outperform stocks with high volatility on a risk-adjusted basis. The positive (negative) abnormal returns of portfolios composed of low-beta (high-beta) stocks - also referred to as the beta anomaly - are one of the most persistent and widely studied anomalies in the field of empirical research of security returns. The lottery demand effect provides a potential explanation for the area of contemporary asset pricing. Bali, Brown, Murray, & Tang (2017) describe the lottery phenomenon in which high volatility stocks are more likely to give a high, lottery-like payoff. This results in an upward price pressure on stocks with high probabilities of large up moves, which are driven by sensitivity to the market portfolio. When riskier payoffs are preferred by a majority of investors, the demand for high volatility stocks will be reflected in the price. Consequently, a positive alpha is generated for a portfolio that is long low-beta stocks and short high-beta stocks. Based on the hypothesis that investors exhibit risk-loving preferences when it comes to the selection of stocks, low volatility stocks tend to be undervalued, leading to an increase in their potential return. Hence, empirical research shows that low-volatility securities tend to generate higher risk-adjusted returns over the longer-term as they usually fall less in down markets. In a study of several Smart Beta funds on the European market, minimum variance Smart Beta portfolio achieved the second highest excess return of 5.47 percent. Only the size portfolio performed better with an excess return of

6.92 percent (De Meyer, 2016). Thus, it is interesting to assess to what extent the low volatility strategy is similarly successful in the US market.

3.4.3 Dividend Yield

Although Miller & Modigliani (1961) argue that dividends should not determine equity prices or equity returns, several studies such as O'Shaughnessy (2012) found that a long-short portfolio based on dividend yield beat all US stocks by 1.6 percent (annualized). The top decile of dividend yield outperformed the bottom decile 71% since 1931 on the basis of rolling five-year periods. Litzenberger and Ramaswamy (1982) suggest that investor require a larger compensation for dividend paying stocks due to a relatively higher tax rate.

Boyadzhiev et al. (2017) claim that the prevailing (low) interest rate environment triggers the demand for dividend yielding portfolios. Fisher (2013) links the good performance of high dividend yielding stocks towards an inherent value tilt. In his opinion investors should directly invest in value and high earnings tilted portfolios to capture the outperformance induced by high dividend stocks. Thus, the analysis investigates whether the popularity of dividend yield tilts reward Smart Beta investors appropriately. Assuming, however, investors are turning towards dividend paying equities as a result of the low (or negative) real interest rates, the price appreciation of these equities should lower the risk/adjusted returns over time.

3.4.4 Fundamentals Based

The underlying intuition of fundamentals-based investing challenges the economic concept that a stock price represents a company's true value. Preferred measures include company sales, earnings, dividends or book value. Ambassadors of such a strategy argue that stock prices are often influenced by speculation, momentum trading or institutional investors who buy and sell equities for undisclosed reasons which are not related to the actual stock performance (Arnott et al., 2005). Interestingly, the fundamental factoring fails to eliminate the capitalization driven reliance on large companies. Due to a high correlation of fundamental attributes with liquidity and capitalisation variables, Arnott et al. (2005) find that fundamental strategies often invest considerably in large companies, similar to traditional capitalisation-based ETFs. Nonetheless, the fundamental index constructed in their study consistently outperformed a cap/weighted market index by more than 2 percent each year between 1962 and 2004. Note, however, that construction of a fundamentals-based SB strategy highly differs according to Ung & Luk

(2016), similarly for dividend and low volatility tilts. This complicates a comparison between various Smart Beta strategies.

3.4.5 Momentum

Like value, momentum trading in the stock market is a rather old phenomenon with academic evidence dating back to the early 1990s. It is based on the idea that stocks tend to maintain recent price trends in the future, whereof the momentum strategy takes advantage of. Extensive research providing significant results in terms of excess returns includes Jegadeesh & Titman (1993) who monitored this phenomenon on the U.S. stock market between 1965 and 1989. The authors provide evidence for trading strategies that buy stocks which have performed well in the past (winners) and sell stocks that have performed poorly in the past (losers) have generated significant positive returns over 3- to 12-month holding periods in the respective time period. They showed that following the strategy of buying winner stocks and shorting losing stocks produced excess returns compared to the benchmark index over the same period. Even though the momentum factor proved to have historical success, the strategy of following trends is a short-horizon game. As the effect reportedly dissipates in less than two years, Jegadeesh & Titman (1993) suggested a three-to twelve-month holding periods for securities. Consequently, momentum strategies require frequent rebalancing, which might offset some of the returns in trading costs. Jegadeesh and Titman (1993) suggest that a large investor liquidating a large block of stock artificially reduces the stock price, which partially reverses over the next month. Movements in the past are related to future returns and can exploit a cognitive bias associated with news information. Hence, they propose that investors can profit from these anomalies when the market corrects such pricing errors. Note that Carhart (1997) was the first academic who described the momentum factor with his four-factor model. The author led to the belief that fund managers who seem to have a "hot hand" are mostly just the lucky beneficiaries of momentum in stock returns.

3.4.6 Quality

There is no consistent opinion throughout literature whether quality companies tend to yield excess returns. However, according to Campbell, Polk, & Vuolteenaho (2009) cash flow fundamentals control stock prices more than macroeconomic variables. In other words, a successful firm can gain a competitive advantage through careful capital management.

Especially during bad times, quality stocks tend to outperform, since investors become risk-averse and aim to invest in stocks with sound capital management when macroeconomic conditions start to deteriorate. Namely, this effect of moving capital away from riskier investments to safer ones, is known as the so-called "flight-to-quality" (Asness, Frazzini, & Pedersen, 2019). There are several research papers dealing with approaches to define quality as a factor. Amongst others, Novy-Marx (2013) found that firms with high gross profitability earned returns in excess to the market benchmark over longer periods. Nevertheless, the quality factor construction seems very inconsistent as other "quality metrics" (return on equity, earnings stability, dividend growth stability, cash flows, etc.) are often mentioned as components throughout literature.

3.4.7 Earnings-Weighted

The earnings-weighted strategy seeks to track the investment results of an earnings-weighted index to reflect the proportionate share of the aggregate earnings each portfolio company has generated. Consequently, companies with greater earnings use to have larger weights within portfolios, as greater earning power of the company is speculated to lead to higher value of its stock. This strategy seems to be not very established in the SB ETF market yet and therefore little empirical evidence is found.

3.4.8 Equal-Weighted

As the name already indicates, the equal-weighted strategy assigns the same (uniform) weight to each component in the index regardless of its market capitalization and incorporates perhaps the simplest of strategic beta methodologies. Hence, equal weight investing is based on unbiased exposures to all stocks, making all stocks equally important (Malkiel, 2014). In other words, this strategy increases the importance of smaller stocks and requires no information about the risk and return characteristics of the portfolio constituents (FTSE Russell, 2017b). Its advocates argue that by breaking the relationship between price and position size, equal-weighted approaches more likely avoid a structural overweight to overvalued securities. On the other hand, one of the strategy's drawbacks includes unintended factor concentrations that could be arbitrary driven by the number of securities that happen to be listed under a particular sector, industry, or country.

3.4.9 Non-traditional

Non-traditional is a SB strategy for which there are only few information available and nearly no publications, studies or reviews with consistent definitions. We expect this category to mainly invest in non-traditional commodities and equities. Based on the underlying dataset of this thesis, there are only a few ETFs within this category tracking mainly non-traditional commodities, such as mining, gold and renewable energy. Hence, this category does not seem to be widely spread yet but might become more popular in the future.

3.4.10 Multi-factor

Lastly, multi-factor strategies experienced a substantial surge in popularity. These strategies attempt to combine a variety of factors (value, growth, momentum and quality for example) in order to improve risk-adjusted performance. The rising availability of multi-factor SB ETFs suggests a saturation towards plain vanilla funds. In different words, many traditional and single factor tied ETFs seem to be less demanded nowadays, perhaps financial markets participants agree that most of the single anomalies have been exploited (Boyadzhiev et. al, 2017). By combining different factors, diversification benefits arise because certain factors behave rather distinctly to market conditions. As a result of their low correlation, excess returns can be achieved while keeping volatility at minimum levels. Naturally, the multi-factor category as a whole is difficult to compare due to largely different construction criteria.

3.5 Microeconomic versus Macroeconomic Factors

In contrast to the rather well-known factors many studies argue that stock returns are strongly dependent on macroeconomic states. Thus, macroeconomic variables are more frequently incorporated as explanatory variable factors of the variation in equity returns (Bilson, Brailsford, & Hooper, 2001). However, most research in Smart Beta products focusses on excess returns with respect to microeconomic effects such as dividend yields and price to earnings ratios. Philips, Bennyhoff, Kinniry, Schlanger, & Chin (2015) show that time-varying exposures to certain risk factors such as size and style partly explain excess return on Smart Beta strategies. Glushkov (2016) assessed whether Smart Beta investors are provided with the intended exposures towards well-established factors. He concludes that only the value factor yielded risk-adjusted excess returns different from zero. Hsu et al. (2015) show that value, momentum, low beta and illiquidity are robust factors in the US market. Unfortunately, only value and low beta factors seem to be robust across regions. To make matters worse, Ung & Luk (2016) correctly state that factor construction is seldom homogeneous. Most strategies are difficult to compare because the strategies apply vastly different stock selection processes. In particular, low volatility and dividend factors are stated to be quite dissimilar. Nonetheless, even prominent factors such as value and momentum (which are constructed in a more uniform manner) are stated to have considerably little explanatory power with respect to SB funds excess returns. Whereas many of these factors are not empirically replicated, well-established factors such as value and size seem to be more accepted factors among researchers in explaining excess returns.

Even so, it seems to be necessary to set a new direction in the context of assessing Smart Beta strategies. As previously mentioned, microeconomic variables often fail to explain Smart Beta (out)performance albeit being widely used in existing studies. Thus, a more macroeconomic perspective is applied in the remainder of this paper. A first stance on macroeconomic influences on Smart Beta performance can be found both in the studies of Ung & Luk (2016) and Glushkov (2016). On one hand, Ung & Luk (2016) divided the sample period into simplified periods of bullish, bearish and market conditions in order to identify superior SB strategies depending on the state of the economy. They found that momentum and growth strategies perform best in bear periods whereas value and dividend yield produce the highest risk-adjusted returns during recovery periods. On the other hand, Glushkov (2016) assigned up ("good" times) and down ("bad" times) periods based on the top and bottom quintile of percentage change in implied volatility index VIX. Consequently, dividend and volatility focussed strategies perform best during market declines, however underperformed during bullish periods. In contrast, equal weighted, revenue-weighted and multi-factor beat the benchmark during "good" markets by taking on larger market risk.

While these findings are particularly relevant for Smart Beta investors, it remains up for discussion to which extent changes in macroeconomic variables can explain excess returns of Smart Beta strategies. Therefore, a more detailed analysis of relevant macroeconomic performance measures is inevitable. It is expected that new variables such as unemployment, interest rates, or consumption growth are able to explain a fairly high degree of Smart Beta excess returns. More importantly, it sheds light on the relationships between macroeconomic changes and various SB strategy returns. Moreover, time-varying results are likely to be found with respect to various SB strategies. First, however, it is important to elaborate on previous findings of macroeconomic studies on equity returns to identify desirable factors for the purpose of this study.

3.6 Macroeconomic Factors

Economic news is well believed to have a particular effect on asset prices. While this hypothesis expresses a strong intuitive appeal, it lacks empirical support. Since macro changes simultaneously affect many firm’s cash flows and risk-adjusted discount rate, though, macro variables are attractive candidates for pricing extra market risk. Financial theory proposes that certain macroeconomic variables such as the spread of long and short interest rates, expected and unexpected inflation, the spread between high- and low-grade bonds and industrial production systematically influence stock markets (Chen, Roll, & Ross, 1986). More importantly, for the context of this study, it matters whether a higher exposure towards macroeconomic changes are separately priced. In different words, are SB tilts towards certain macroeconomic variables rewarded as a result of bearing extra market risk?

It is noteworthy that the key problem in assessing macroeconomic factors lies in the inconsistent application of coherent research methodologies. This does not come as a surprise based on the previously outlined (similar) shortcomings in microeconomic factoring studies. Birz & Lott Jr (2011) argue that weak evidence with respect to linking real economic news and stock returns has been presented among researchers. The reason lies in the heterogeneous approaches towards interpreting macroeconomic information in different economic contexts.

Albeit results of many macroeconomic studies deliver rather mixed results, most have outlined significant relationships between macroeconomic variables and equity returns (Tangjitprom, 2012). Most widely used methodologies include multiple regression techniques, vector autoregression techniques or the co-integration technique. Another angle is chosen in many studies that apply the GARCH model or GARCH-family models to capture time-varying volatility of equity returns. Lastly, in contrast to time-series data, a group of papers chooses event-studies in order to study the announcement effect of macroeconomic news.

Bodie (1976), and Geske & Roll (1983) indicate that both expected and unexpected inflation next to money growth has a negative impact on equity values. However, Geske & Roll (1983) warn that the empirical phenomenon does not indicate causality and might be an empirical illusion which is proven through inducing a spurious causality by a combining of (i) a reversed adaptive inflation expectations model and (ii) a reversed money growth/stocks model. Aforementioned study of (Chen et al., 1986) assessed empirical evidence to extend other (macroeconomic) risk factors beyond the equity market risk premium inherent in the introduction in the CAPM (Sharpe, 1964; Lintner, 1965; Mossin, 1966). Whereas the spread of long and short interest rates, expected and unexpected inflation, the spread between high- and low-grade bonds and industrial production are found to systematically affect stock returns, no evidence is found for oil price changes or aggregate consumption to be separately priced. Moreover, Flannery, & Protopapadakis (2002) estimate a GARCH model of daily equity returns in which realised returns and conditional volatility derive from 17 macro series' announcements. In total, they conclude that three nominal (Consumer Price Index, Producer Price Index and a Monetary Aggregate) and three real (Balance of Trade, Employment Report and Housing Starts) series qualify as priced equity factors.

A different stance is taken by McQueen & Roley (1993) who blame the failure in finding significant macro factors to a shortcoming of the constant-coefficient models being estimated. They state that based on the business cycle the announcement surprise impacts the equity returns differently. As an example, a rise in employment can be a bullish sign when the economy just comes out of a recession whereas it very well can be bearish sign close to a cyclical peak. Likewise, Boyd, Jagannathan, & Hu (2001) present their results of time-varying effects of macro news on equity returns. In a study of equity returns between 1948 and 1995 they find that surprisingly high unemployment raises stock prices during economic expansion but lowers it during contraction.

Altogether, macroeconomic variables and stock returns certainly influence each other. However, researchers have not yet agreed on the direction of the relationship. Whereas some studies conclude that macroeconomic variables explain future stock returns (Asai & Shiba, 1995; Hondroyiannis & Papapetrou, 2001) the opposite effect is found, namely that stock returns can be used to predict future macroeconomic variables (Tangjitprom, 2011; Henry, Olekalns, & Thong, 2004). For the purpose of this study, the logic that expected returns depend upon macroeconomic risk factors applies. Thus, a unidirectional relationship is assumed from macroeconomic variables to Smart Beta returns.

Many macroeconomic variables capture cyclical market conditions. Therefore, Tangjitprom (2012) categorises these variables into four groups depending on their effect on either (i) economic conditions, (ii) interest rate and monetary policy, (iii) price levels and (iv) international activities. Throughout this study, several variables will be presented for each of the above categories in order to assess the explanatory power of certain macroeconomic variables on Smart Beta returns. Ideally, yet unanticipated sensitivities of SB returns towards macroeconomic changes can be derived in order to provide new evidence to better understand differences and similarities among SB strategies.

Whereas the previously outlined (microeconomic) factors can and should be tested on a global basis, macroeconomic variables are rather "local" variables that vary from country to country. Thus, it would be highly interesting to address the macroeconomic analysis towards Smart Beta ETFs domiciled in different countries. However, note that SB ETFs only recently started to take off in countries outside of the US. Consequently, historical return information is not sufficient to find empirically verifiable results with respect to differences in macroeconomic factors among several countries. In different words, we do not feel comfortable to use the infrequently reported macroeconomic data in combination with Smart Beta returns of these rather new markets to draw conclusions. The results would be indicative at best so that this analysis purposely investigates the impact of macroeconomic variables in the universe of US domiciled SB strategies only. A comparison between different markets will be relevant for future research once the data quality allows for such an assessment.

3.6.1 Economic Conditions

Most obviously, perhaps, seems to be the GDP which measures the growth rate for domestic products or national output. It is argued that this growth represents both the change in economic activities and the population's income level. Singh, Mehta, & Varsha (2011) find that GDP affected stock returns positively across all constructed portfolios in the study on the Taiwanese stock market. They conclude that investors stand a chance to develop profitable investment strategies based on macroeconomic changes. In contrast, Birz & Lott Jr (2011) claim that albeit being widely used in macroeconomic studies, GDP or Gross National Product (GNP) is hardly ever producing statistically significant relationships between changes in GDP/GNP and stock returns. Likewise, Flannery, & Protopapadakis (2002) did not find significant effects of GDP on stock returns.

Furthermore, including the industrial production index, for example, serves as a good proxy for economic conditions. The growth tends to be consistent with the average growth of companies' sales and cash flow (Chen et al., 1986). This finding is strengthened by Humpe & Macmillan (2009) who showed a positive relationship between the industrial production index and stock prices in Japan and the US. Flannery & Protopapadakis (2002), however, fails to find a significant relationship.

Another popular measure of economic prosperity is the unemployment rate. Boyd, Hu, & Jagannathan (2005) outline that the announcement of surging unemployment positively affects equity prices during economic expansions and negatively in times of contractions. This is due to the bundling property of the unemployment variable: information about future interest rates, equity risk premium and corporate earnings and dividends are all interrelated to unemployment news. The timing of the information bundle is highly relevant in the context of the economy's state. During expansions information about interest rates dominate while information about future corporate dividends prevails during contractions. Flannery & Protopapadakis (2002) show that unemployment does influence stock returns' conditional volatility only. In contrast, Pearce & Roley (1984) find no significant stock price effects with regard to employment/unemployment.

3.6.2 Interest Rate and Monetary policy

Widely used as an economic indicator the interest rate is said to be particularly relevant for commercial bank stocks (Tangjitprom, 2012). Other studies extend the applicability towards the general stock market (Li, Iscan, & Xu, 2010). Whereas, many studies use yields on government securities as the proxy of the interest rate level, others focus on the Federal Reserve announcement of discount rate changes. For example, Prather and Bertin (1999) show that general stock return movements can be predicted using the announced discount rate changes. In contrast, Durham (2003) finds a weak and insignificant relationship between the discount rate of 16 countries and the respective stock returns. Alam & Uddin (2009) study the relationship between interest rate changes and stock prices and find a significant, negative relation for several countries. Based on the literature we choose the changes in the US Treasury bill rate (3 months) as a proxy for interest rate levels.

Another popular measure of economic activity is money supply (M2) growth rate as applied by Chen (2007) and (Bilson et al., 2001). Again, Flannery, & Protopapadakis (2002) find that money supply affects both the level of volatility and volatility of equity returns. In another macroeconomic study by Chancharat, Valadkhani, & Havie (2007) the impact of money supply (M2) among other macro factors is tested on Thai stock returns. Their results are derived from a GARCH-M model and indicate no impact of money supplies on stock returns.

3.6.3 Price Levels

Often chosen as one of the most prominent variables affecting price levels is the Consumer Price Index because it includes information about general price levels and inflation. In many studies, the inflation is broken into expected versus unexpected inflation. For example, Chen et al. (1986) use the percentage change of the CPI to measure actual inflation. The difference between expected and actual inflation obviously leaves unexpected inflation. The expected inflation is less straightforward to determine and is derived from inflation forecasting based on other economic factors. In the study, they find a negative impact of both inflation variables and stock returns. Similarly, Pearce & Roley (1984) and Adrangi, Chatrath, & Raffiee (1999) find a negative relationship between inflation and stock returns where the latter finding concerns the Korean stock market. Lastly, an asymmetric model constructed by Kolluri & Wahab (2008) indicates that the relationships between inflation and stock returns changes. During low/high inflation regimes the relationship is negative/positive respectively.

Another prominent variable that is often seen as a proxy for price levels is the change in oil prices. The popularity certainly derives from substantial importance for consumption and production processes (Tangjitprom, 2012). The result presented by Faff & Brailsford (1999) indicate that oil price changes are clearly affecting equity returns in many Australian industries where the direction of the relationship depends on the particular industry. Oil and gas industry express a significant positive relationship whereas it is negative in case of the paper and packaging and transport industries. Park & Ratti (2008) find statistically significant results between oil price shocks and equity returns in the US and 13 European countries between 1986 and 2005. They conclude that oil price shocks increase stock returns' volatility by statistically significant 6 percent (median result). Increased volatility of oil prices significantly depress equity returns for many European countries (not the US). Lastly, Fedorova, & Pankratov (2010) study the similar relationship in Russia and find Brent oil prices are the macroeconomic factor that influence stock returns the most. Consequently, linking the industry specific finding of Faff & Brailsford (1999) to the context of Smart Beta strategies, it is likely that oil price changes as a macro factor loads quite differently depending on industry holdings of their respective Smart Beta strategies.

3.6.4 International Activities

A very popular indicator of international activities is given by the exchange rate especially for countries that depend on international trade (Tangjitprom, 2012). However, this analysis purely focuses on the US market so that exchange rates seem less relevant. Apart from this, FDI is often used. Both studies of Adam & Tweneboah (2008) and Mohammed et al. (2009) find show that FDI plays an important and significant role with respect to stock price movements in Ghana and Pakistan respectively. As a result, FDI and all aforementioned variables are included in the analysis to assess its relevance in the US Smart Beta ETF market.

3.6.5 Choosing the Appropriate Amount of Equity Factors

After having presented a vast amount of both micro- and macroeconomic factors that influence equity returns the key question lies not only in identifying but also choosing the adequate number of factors. In theory, it is possible to add factors for the sake of increasing R^2 of multi-factor regressions as it often has been done in previous studies. However, as pointed out earlier, it is virtually impossible that each factor is useful on a stand-alone basis. Datamining and correlation among factors have fuelled the development towards "factor breeding" especially in the context of microeconomic factors such as market or value factors. Despite lacking prior research in the area of macroeconomic influences within the Smart Beta environment, general advice can be found on a suitable number of factors in the context of equity factors. Prominent researchers argued that usually a handful of factors sufficiently explain equity returns. The source of contention is addressed by Trzcinka (1986) who identifies five dominant factors with respect to US equity returns. Likewise, Cho (1984) states that the number of factors lies in between two and five in an inter-battery factor analysis on a selection of US industries. Cho et al. (1986) reports an adequate amount of one to five factors on a similar study at the international level for eleven industrial economies.

However, the exact number of factors is subject to much criticism due to arbitrary selection processes and subjectivity. According to Fama (1991) this remains an unavoidable problem in this area of research. As a result, this study incorporates all eight macroeconomic factors as possible explanatory variables. Later, few macroeconomic variables are selected in a stepwise multiple regression in order identify significant factors. In another attempt to mitigate the dimensionality issue addressed above principal component analysis (PCA) is conducted. The principal component loadings will provide a better comparison between Smart Beta strategies and the respective sensitivity towards varying macroeconomic variables. The next section elaborates on the methodology and research method used.

4 Methodology

This section provides an overview of existing approaches in prior research which are later applied in the context of this study. Presenting the mathematical and statistical properties of these approaches provides a fundamental understanding of the research method. As a result, an empirical framework attempts to introduce the reader into the relevant research methods. Note, however, that the first subsection (section 4.1) focuses on a rather theoretical presentation of relevant research methods followed by discussing necessary assumptions and possible implications (section 4.2). Thereafter, certain performance measures will be touched upon. These will be applied in section 4.3 to assess the Smart Beta ETF performance. Next, data quality issues are mentioned. Lastly, the aforementioned research methods in the SB ETF context will be outlined in depth (section 4.4).

4.1 Empirical Framework

Throughout this chapter, we aim to provide mathematical insights for the models used in our analysis. The simple linear regression model serves as a starting point in our study when evaluating the excess returns of SB ETFs on their respective benchmark. Based on this, the multiple linear regression model expands the CAPM by more than just one explanatory variable and allows to extend the model by the FF factors and our chosen macroeconomic variables. Including up to eleven factors in the multiple regression model, we follow suggestions throughout literature and believe that we can reduce the model to three factors. This step is conducted using PCA, introduced and derived as a last step of this section.

4.1.1 Linear Regressions

Linear regressions are a useful tool when one seeks to predict a quantitative response. Approved to be a useful and widely used statistical learning method, it serves as a good starting point in understanding more extensive approaches. This section begins with the *Simple Linear Regression*, before introducing the model of *Multiple Regressions.*

The model estimation relies on Ordinary Least Squares (OLS) based on multivariate statistical regressions, named as *multiple linear regression model* in the course of this thesis. Amongst several academia, Campbell, Lo, & MacKinlay (1997) inspire the use of OLS as a consistent estimation procedure for the market-model parameters, but also Cuthbertson & Nitzsche

(2004). OLS estimate sets of explanatory variables (coefficients) for multiple regressions when aiming to minimize the dependent variable residual from the fitted model. Before continuing with multiple regression models, we will introduce the simple linear regression model as a starting point before extending the (simple) model. Inspired by Campbell et al. (1997), the theoretical foundation will be based on his explanations.

Known as a very straightforward approach for predicting a quantitative response *Y* on the basis of a single predictor variable *X*, it presumes a linear relationship between *X* and *Y*. In mathematical terms, this can be expressed as

$$Y \approx \alpha + \beta_1 X \tag{6}$$

where α and β_1 are two unknown constants representing the *intercept* and *slope* terms in the linear model. Alpha shows the outperformance of the fund in terms of excess returns as predicted by the regression model. Similarly, a negative alpha indicates how much worse it did. Beta, the slope, is the degree of change in the outcome variable for every one-unit of change in the predictor variable, *holding all other predictors fixed.* The beta coefficient can be negative or positive. A positive beta coefficient implies that for every one-unit increase in the predictor variable, the outcome variable will increase by the beta coefficient value. Vice versa, if it is negative, for every one-unit increase in the predictor variable, the outcome variable will decrease by the beta coefficient value.

These constants are also known as the model *coefficients* or *parameters* and are based on estimates leading to the following equation:

$$\hat{y} = \hat{\alpha} + \hat{\beta}_1 x \tag{7}$$

where $\hat{y}$ indicates a prediction of Y (denoted by the hat symbol) on the basis of $X = x$. When estimating the coefficients, we aim to obtain parameters $\hat{\alpha}\ and\ \hat{\beta}_1$ such that the linear model fits the available data well, lying as close as possible to the data points. Defining

$$\hat{y_\imath} = \hat{\alpha} + \hat{\beta}_1 x \tag{8}$$

to be the prediction for *Y* based on the *i*th value of *X*, then $e_i = y_i - \hat{y}_i$ represents the *i*th *residual*; which is the difference between the *i*th observed response value and the *i*th response value predicted by the linear model. The *residual sum of squares* is defined as

$$RSS = e_1^2 + e_2^2 + \cdots + e_n^2 \tag{9}$$

or equivalently as

$$RSS = \left(y_1 - \hat{\alpha} - \hat{\beta}_1 x_1\right)^2 + \left(y_2 - \hat{\alpha} - \hat{\beta}_1 x_2\right)^2 + \cdots + \left(y_n - \hat{\alpha} - \hat{\beta}_1 x_n\right)^2. \tag{10}$$

In order to measure *closeness*, the least squares approach chooses $\hat{\beta}_0\ and\ \hat{\beta}_1$ to minimize the RSS. The minimizes are calculated as

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}((x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{11}$$

and

$$\hat{\alpha} = \bar{y} - \hat{\beta}_1 \bar{x}. \tag{12}$$

where $\bar{y} \equiv \frac{1}{n}\sum_{i=1}^{n} y_i\ and\ \bar{x} \equiv \frac{1}{n}\sum_{i=1}^{n} x_i$ are the sample means. Summing it up, the above equation defines the *least squares coefficient estimates* for simple linear regression. Assuming the relationship not to be fully linear, an error term has to be included, as there may be other variables that cause variation in *Y* possibly resulting in measurement error. Having an error term that is independent of *X*, we can rewrite the equation as

$$Y = \alpha + \beta_1 X + \epsilon. \tag{13}$$

Standard errors can also be used to perform hypothesis tests on the coefficients. SE tells us the average amount that the estimated mean differs from the actual value of the mean. The most common hypothesis test involves testing the null hypothesis, stating that the coefficients are insignificant against the two-sided alternative that they are not. Providing a specific example for the beta coefficient, mathematically corresponds to testing

$$H_0\colon \beta_1 = 0 \tag{14}$$

versus

$$H_A: \beta_1 \neq 0, \tag{15}$$

where $\beta_1 = 0$ would reduce the model to

$$Y = \alpha + \epsilon \tag{16}$$

withdrawing the relationship between X and Y. To test the null hypothesis, we need to determine if $\hat{\beta}_1$ is far from zero. Therefore, we compute a *t-statistic* given by

$$t = \frac{\hat{\beta}_1 - 0}{SE(\hat{\beta}_1)} \tag{17}$$

Small values of $\hat{\beta}_1$ provide strong evidence that $\beta_1 \neq 0$, hence $\hat{\beta}_1$ must be large in absolute value in order to reject the null hypothesis. Having no relationship, we expect to have a t-distribution with $n - 2$ degrees of freedom. The probability of of the null hypothesis to be statistically significant is named the p-value, where a small p-value (typically p<0.05) indicates an association between the predictor and the response. When the regression is conducted, an R^2 statistic (coefficient of determination) is computed in order to assess the percent of variance in the outcome variable that is explained by the set of predictor variables. An R^2 statistic close to 1 indicates that a large proportion of the variability in the response has been explained by the regression, whereas a value close to 0 does not.

4.1.2 Multiple Regressions

Over time, the one factor model, introduced as the CAPM was treated more sceptically as more researchers started to be interested in analysing the influence of multi-factor models on stock excess returns. Thus, according to Bilson et al. (2001) factors that influence equity returns turned out to be a source of much contention.

Ross (1976) developed the Arbitrage Pricing Theory, a one-period model in which every investor believes that the stochastic properties of returns of capital assets are consistent with a factor structure. Based on the idea that an asset's return can be predicted using the linear relationship between the asset's expected return and a number of macroeconomic variables that capture systematic risk, several academia continued their research based on this finding. In a later work, Roll & Ross (1984) applied factor analysis to 42 groups of 30 stocks using data between 1962 and 1972. After running a first-pass regression they find that for most groups about five “factors” provide a sufficiently good statistical explanation of R_{it}. In the second-pass regression they conclude that the number of factors can be reduced to three factors providing sufficient explanation. According to Dhrymes, Friend, & Gultekin (1984) the number of statistically significant factors appear to increase as more securities are included in the analysis, causing difficulties in interpreting results from factor analysis. In the following years, Sharpe (1982), Chen (1983), Roll & Ross (1984), and Chen et al. (1986) examine a wide variety of factors that might influence expected return in the first-pass time series regression. Inspired by Ross (1976) these academics used the estimated coefficients from the first-pass regressions as cross-section variables in the second-pass regression for each month. Providing statistically significant results, they could support a multi-factor APT model. Overall, the early empirical work on the APT pioneered the movement that more than one factor is important in determining asset returns. This is approved by continuous studies on US data (e.g., Shanken, 1992; Shanken & Weinstein, 1990) and UK data (Clare & Thomas, 1994; Poon & Taylor, 1991) using improved variants of the two-stage regression tests.

When facing several factors, an extension of the simple linear regression model is required, instead of fitting a separate simple linear regression model for each predictor. Contributing each predictor, a separate slope coefficient and assuming to have p distinct predictors, the multiple linear regression model takes the form

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \epsilon \tag{18}$$

where X_j represents the jth predictor and β_j quantifies the association between that variable and the response.

As explained for a simple linear regression setting, the regression coefficients $\alpha, \beta_1, \dots, \beta_p$ are unknown and must again be estimated. Predictions are made by using the formula

$$\hat{y} = \hat{\alpha} + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_p x_p. \tag{19}$$

Choosing the least squares approach to minimize the sum of squared residuals leads us to

$$RSS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} \left(y_i - \hat{\alpha} - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \cdots - \hat{\beta}_p x_{ip}\right)^2. \tag{20}$$

Again, we aim to identify if there is a relationship between the response and predictors. Recall that in the simple linear regression setting we can simply check if the coefficient beta equals zero. However, in the context of a multiple regression with several predictors, we need to test if all regression coefficients are zero, i.e. whether $H_0 = \beta_1 = \beta_2 = \ldots = \beta_p = 0$ versus the alternative that at least one β_j is non-zero. This type of hypothesis test is performed by computing the F-statistic,

$$F = \frac{\left[\frac{TSS - RSS}{p}\right]}{\left[\frac{RSS}{n - p - 1}\right]} \tag{21}$$

where $TSS = \Sigma(y_i - \bar{y})^2$ and $RSS = \Sigma(y_i - \hat{y}_i)^2$.

Again, an F-statistic resulting in a p-value below 0.05, allows to reject the Null-Hypothesis. After computing the F-statistic and examining the associated p-value that eventually allows to conclude that at least one of the predictors is related to the response, the next step is to identify the guilty ones. As described for the simple setting, an R^2 statistic explains how many percent of variance in the outcome variable is explained by the set of predictor variables. After evaluating the F-value and R^2, the t-test will be used to determine the significance of each predictor and beta coefficients will be used to determine the magnitude of prediction of each independent variable. For (significant) predictors, every unit increase in the predictor causes the dependent variable to increase or decrease by the value of its unstandardized beta coefficients.

The common test statistics are applied in order to assess statistical significance based on respective distributions and defined by its degrees of freedom and confidence levels respectively. If applicable, the t-statistics account for heteroscedasticity as White's heteroscedasticity-consistent (HC) standard errors and Newey-West's heteroscedasticity- and autocorrelation-consistent (HAC) standard errors are applied in relevant cases. Walt (1980) presents a covariance matrix estimator of the ordinary least squares coefficients that is consistent in presence of conditionally or unconditionally heteroscedastic error terms. Newey & West (1986) similarly present consistent standard errors with respect to autocorrelation and heteroscedasticity. Croux, Dhaene &, Hoorelbeke (2004) state that these robust standard errors will not be robust against OLS outliers since they are derived from the OLS estimator. Thus, it is important to point out that outliers have been removed from the data set.

4.1.3 Principal Component Analysis

Principal component analysis is a method of data processing which extracts a small number of synthetic variables, named principal components, from a large number of input variables. PCs are a sequence of projections of the data, mutually uncorrelated and ordered in variance, which are obtained as linear manifolds approximating a set of N points (James, Witten, Hastie, & Tibshirani, 2013). PCA conveniently summarizes a large set of observations of possibly correlated variables with a smaller number of representative variables that collectively explain most of the variability in the original set. Replacing the original larger set of variables with principal components as predictors in a regression model in the next step, one can perform a principal components regression. The PC regression substantially reduces the dimensionality of the model which can then be compared to the multiple regression including all explanatory variables. Ideally, the fit of the model decreases marginally.

Understanding this concept in a more empirical way (in mathematical terms), the starting point is that one is supposed to want to visualize n observations with measurements on a set of p features, $X_1, X_2, \ldots, X_p$, as part of an exploratory data analysis. Since examining the data in form of a two-dimensional scatterplot is too complex, unless p is small, an alternative model in form of a low-dimensional representation of the data capturing as much of the information as possible is clearly required. PCA aims for exactly this approach based on the idea, that each of the n observations live in a p-dimensional space, but with an unequal distribution of interest along these dimensions. With a small number of dimensions that are as interesting as possible, measured by the amount that the observations vary along each dimension, each of the dimensions found by PCA is a linear combination of the p features. A set Ω is transformed to PCs which are uncorrelated and ordered in a way that the first component captures most of the variation present in the original variables. For every nxK matrix X where n and k represent observations and explanatory variables respectively, PCA searches for linear combinations of X in a lower dimensional space. Thus, orthogonal transformation yields a linear representation of the highest possible variance in the original data set. In order to remove the issue of scale dependence from PCA, standardization is suggested. PCA ensures minimal information loss while choosing the most relevant variables to represent multivariate data in lower dimensions (James et al., 2013; Jolliffe, 2011).

Algebraically, the first principal component, y_1, is a linear combination of $x_1, x_2, \dots, x_p$; i.e.,

$$y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_1 x_p = \sum_{i=1}^{p} a_{1i} x_i. \tag{22}$$

such that the variance of y_1 is maximized given the constraint that the sum of the squared weights is equal to one (i.e., $\sum_{i=1}^{p} a_{1i}^2 = 1$). PCA finds the optimal weight vector $(a_{11}, a_{12}, \dots, a_{1p})$ and the associated variance of y_1 which is denoted by λ_1. Similarly, the second PC, y_2, involves finding a second weight vector $(a_{21}, a_{22}, \dots, a_{2p})$ such that the variance of

$$y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p = \sum_{i=1}^{p} a_{2i} x_i. \tag{23}$$

is maximized subject to the constraints that it is uncorrelated with the first principal component and $\sum_{i=1}^{p} a_{2i}^2 = 1$. Hence, y_2 is then having the next largest sum of squared correlations with the original variables. As a logical consequence, the sum of squared correlations with the original variables, or equivalently, the variances of the principal components get smaller as successive principal components are extracted. The first two PCs together provide the highest possible sum of squared multiple correlations, i.e., $\sum_{i=1}^{p} R_{xiy_1,y_2}^2$, with p variables. The process can be continued until as many components as variables have been calculated; till the *k*th stage where a linear function y_k is found, which has the maximum variance, also subject to being uncorrelated with $y_1, y_2, \dots, y_{k-1}$. The *k*th derived variable y_k is the *k*th PC. However, the first two PCs usually account for most of the variation in the variables and put them at the centre of our focus, although small components still can provide information about the data structure (Dunteman, 1989). The general goal is that most of the variation in x will be accounted for by m PC's where $m < p$.

After defining PCA, the question arises of how to calculate them. Defining x as a vector of p random variables and a_k as a vector of p constants, we can then rewrite as follows, $a_k' x = \sum_{i=1}^{p} a_{ki} x_i$. Top of all, the concept of covariance matrix Σ of vector x needs to be introduced. This is the matrix whose *(i, j)*th element is the covariance between the *i*th and *j*th elements of

x when $i \neq j$, and the variance of the jth element of x when $i = j$. When seeking to derive the first PC, $a_1'x$, in which vector a_1 maximizes the variance, $var(a_1'x) = a_1'\sum a_1$. Since a_1 could be infinitely large without constraint, a normalization constraint, namely $a_1'a_1 = 1$ (unit length vector) is chosen, which indicates that the sum of squares of elements of a_1 equals 1.

In order to maximize $a_1'\Sigma a_1$ subject to $a_1'a_1 = 1$, the standard approach is to use Lagrange multipliers technique, that is to maximize $a_1'\Sigma a_1 - \lambda(a_1'a_1 - 1)$, where λ defines the Lagrange multiplier. Differentiating the latter equation with respect to a_1 results in $\Sigma a_1 - \lambda a_1 = 0$ or $(\Sigma - \lambda I_p)a_1 = 0$, where I_p is the (pxp) identity matrix[9]. Accordingly, λ is an Eigenvalue of Σ and a_1 is the corresponding Eigenvector. Solving the question which eigenvector to choose, we recognize that the quantity to be maximized is $a_1'\Sigma a_1 = a_1'\lambda_1 a_1 = \lambda_1 a_1'a_1 = \lambda_1$ and therefore, we should choose λ_1 to be as big as possible. So, calling λ_1 the largest eigenvector of Σ and a_1 the corresponding eigenvector then the solution to $\Sigma a_1 = \lambda_1 a_1$ is the 1st principal component of x.

The second PC, $a_2'x$, maximizes $a_2'\Sigma a_2$ subject to being uncorrelated with $a_1'x$. The un-correlation constraint equals $cov(a_1'x, a_2'x) = 0$, where cov(x, y) denotes the covariance of the random variables x and y and can be expressed by using any of these quotations; $cov(a_1'x, a_2'x) = a_1'\Sigma a_2 = a_2'\Sigma a_1 = a_2'\lambda_1 a_1' = \lambda_1 a_2'a = \lambda_1 a_1'a_2 = 0$. Choosing the last out of these we can write an Lagrangian to maximize a_2 as follows, $a_2'\Sigma a_2 - \lambda_2(a_2'a_2 - 1) - \phi a_2'a_1$, where λ, ϕ are the Lagrange multipliers. Similarly, to before, again differentiation of this quantity with respect to a_2 needs to be taken before setting the result to zero and results in $\Sigma a_2 - \lambda_2 a_2 - \phi a_1 = 0$. Multiplying the left-hand side with a_1 leads to this expression: $a_1'\Sigma a_2 - \lambda_2 a_1'a_2 - \phi a_1'a_1 = 0$. Keeping the non-correlation constraint in mind, we transform the latter equation into $0 - 0 - \phi 1 = 0$, where we can see that ϕ must be zero. Therefore, we are left with $\Sigma a_2 - \lambda_2 a_2 = 0$, or equivalently $(\Sigma - \lambda I_p)a_2 = 0$, clearly another eigenvalue equation concluding that λ is an Eigenvalue of Σ, and a_2 the corresponding Eigenvector. Analogical to the first PC, the strategy of choosing a_2 to be the eigenvector associated with the second largest eigenvalue yield the second PC of x, namely $a_2'x$.

9 Also referred to as a unit matrix of size n, represents an n*n square matrix with ones on the main diagonal and zeros elsewhere.

This process can be repeated for $k = 1 \dots p$ yielding up to p different eigenvectors of Σ along with the corresponding eigenvalues $\lambda_1, \dots \lambda_p$; with λ_p being the smallest Eigenvalue, respectively. Furthermore, the variance of each of the PCs is given by $Var(a_k' x) = \lambda_k$.

4.2 Implication and Assumptions

Concluding the statistical methodology of model testing, it is inevitable to discuss important assumptions and implication that underlie the methodology of this study. The following assumptions apply throughout the entire analysis.

4.2.1 Linearity Assumption

To begin with, the relationship between the response variable y to the predictors $x_1, x_2, x_3 \dots x_n$ is assumed to be linear in regression parameters (Chatterjee & Hadi, 2012). As Osborne & Waters (2002) correctly state, unmodelled non-linearity can be identified by plotting residuals against predicted values of y. In case of non-linearity of in the parameters, the calculated coefficients likely lead to erroneous conclusions about strength and nature of the relationships among variables. As a result, various plots of the predictor and response variables including the fitted regression slope were applied (Appendix I) to assess the presence of linearity. Besides partial violations, the linearity assumption is said to hold throughout the statistical analysis. The assumption of linearity is said to hold in order to implement the data analysis.

4.2.2 Stationarity & Autocorrelation Assumption

Next, stationarity is a requirement when running OLS regressions. Autocorrelation occurs when the residuals are not independent from each other. In other words when the value of $y(x_{t+1})$ is not independent from the value of $y(x_t)$. This for instance typically occurs in equity prices, where today's price is not independent from yesterday's price (Myers, 1990). Granger and Newbold (1974) show that regression results with non-stationary data will be spurious. If data is not stationary, i.e. there is autocorrelation across points in time, the resulting estimators and hypotheses tests are likely to be misleading, to say the least. Distributions of the estimators will not be asymptotically normal, and the hypothesis testing is inappropriate. Under the null hypothesis of the Dickey Fuller test, the data-series has a unit root. Note that the null hypothesis is rejected in each instance so that stationarity is accepted for each variable's series. Neither an intercept nor a trend is included in the test regression. The critical values are taken from Hamilton & Susmel (1994) and Dickey & Fuller (1981). Autocorrelation will be accounted for using Durbin-Watson statistic through Newey-West's heteroscedasticity- and autocorrelation-consistent (HAC) standard errors (Appendix II). However, autocorrelation in equity returns implies rejection of the EMH in its weakest form; hence stationarity is assumed to hold.

4.2.3 Homoscedasticity Assumption

Likewise, the variation in the residuals (amount of error in the model) is supposed to be similar at each point across the model. Differently said, variance σ^2 can be viewed as the inverse of information. Thus, a small variance implies that all data points lie close to the regression line whereas a large variance indicates a larger vertical spread about the regression line. It is assumed that each point of contributes equally to total information, hence has the same variance σ^2. As mentioned above, the vertical spread of the data serves as an estimate for the error in the model and is assumed to be constant. This assumption of homoscedasticity can be assessed by the scatter plot of (standardized) residuals versus (standardized) predicted values (see Appendix III) and is accounted for using Newey-West's heteroscedasticity- and autocorrelation-consistent (HAC) and White's heteroscedasticity-consistent (HC) standard errors. A random spread indicates constant variance, otherwise heteroscedasticity is present and the assumption is violated (Casson & Farmer, 2014). The assumption of homoscedasticity is said to hold despite violations in the data set.

4.2.4 Normality Assumption

Next, the assumption of normality is often inherent in statistical tools used throughout this study. Density plots in Appendix IV show the return distribution of each strategy, its respective benchmark plus the explanatory macroeconomic variables. According to Royston (1982), the Shapiro-Wilk test compares the scores in the sample to a normally distributed set of scores with the same mean and standard deviation under the null hypothesis is that "sample distribution is normal." A significant test implies a non-normal distribution. Whereas small samples often pass the test, in case of larger samples significant results are derived with slight deviations from normality. However, this does not affect the results of a parametric test (Öztuna, Elhan &

Tüccar, 2006). The Shapiro-Wilk test suggests non-normal distributed data but as so often, the data sample of historic returns is not exactly normally distributed for each sub sample. Based on the central limit theorem, this does not pose a severe threat to the analysis though since it is the residuals that are required to be normally distributed (Osborne and Waters, 2002). Likewise, Williams, Grajales & Kurkiewicz (2013) argue that essentially only the assumption of normally distributed errors is relevant to multiple regression. It may be assumed that errors are normally distributed for any combination of values of predictor variables. Non-normal errors may imply that coefficient, t and F statistic do not follow t and F distributions. In contrast, in case of multiple regression it is not required that the normally distributed error assumption holds in order to determine unbiased and consistent coefficients. As the sample increases in size, inferences about coefficients will likely become more reliable. Based on the central limit theorem, the sampling distribution of coefficients tends to approach a normal distribution as the sample size increases (even in case of non-normally distributed errors). Thus, the regression results can be said to be relatively robust to the assumption of normally distributed error terms. Both Q-Q plots and histograms are used to check for normality of the residuals (Appendix IV). Consequently, the assumption is said to hold.

4.2.5 Multicollinearity Assumption

Other potential problems include multicollinearity. Multicollinearity implies correlations between predictor variables. Whereas in severe cases such as perfect correlation no unique least squares solution to the regression can be found, it is more common to receive unstable estimates of individual predictor coefficients (Williams, Grajales & Kurkiewicz, 2013). More specifically, the standard errors and confidence intervals are likely inflated. However, in prediction uses of the response variable, the results are not harmfully distorted (Montgomery et al., 2001). Multicollinearity can be checked against through a correlation matrix among independent variables or the Variance Inflation Factor (VIF) (see Appendix V). As a rule of thumb, a covariate should consist of at least 10 data points (Casson & Farmer, 2014). Chatterjee & Hadi (2012) propose principal component regression as a response to multicollinearity. Hence, a PCA approach is included in the analysis of macroeconomic variables.

4.2.6 Outliers

Lastly, outliers can strongly affect regression results by having unusually high/low values. While this does not necessarily pose a problem for the analysis, excluding such observations from the analysis has to be carefully considered (Williams, Grajales & Kurkiewicz, 2013). The criteria of excluding observations can often be rather subjective besides well-known diagnostics such as Cook's distance (Cohen et al., 2003). Any values exceeding one are likely to be significant outliers and should be removed from the dataset. Nonetheless, it is often argued that unless there is a justifiable argumentation for believing possible outliers represent erroneous data, they should be included (Casson & Farmer, 2014). Outliers will be removed from the data set in a similar fashion combined with visual inspection of the data sample (Appendix VI).

4.3 Evaluation: Performance Measures & Risk Management

In order to assess the performance of the Smart Beta portfolios over its underlying benchmark portfolios, selected performance metrics will be computed. Both the absolute and relative performance will be analysed on a risk-adjusted basis through several of the following metrics. Note that absolute performance considers Smart Beta portfolio on its own whereas the relative performance applies the benchmark ETF portfolio as a reference. Risk-adjusted performance is an essential component in assessing to what extent the riskiness of a particular strategy is appropriately rewarded. Furthermore, different Smart Beta strategies can be compared based on their level of risk. Note that the metrics are always derived from monthly data but for the purpose of this study all metrics are annualised.

Table II: Absolute Performance Metrics

Metric	*Definition*	*Formula*
Annualized Return	*Annualized Return equals the return of the ETF converted into an annual rate. It is computed by compounding monthly ETF Price returns over a 12-month investment horizon and annualizing the rate as a last step.*	$$Rm_{ETF,t} = \frac{P_{ETF,t} - P_{ETF,t-1}}{P_{ETF,t-1}}$$ $$Annualised\ Return, R_{ETF,t} = \left(\prod_{i=t-11}^{t} (1 + Rm_{ETF,i})\right) - 1$$ *Where $P_{ETF,t}$ is the monthly adjusted ETF price at moment t and $Rm_{ETF,t}$ denotes monthly ETF return at moment t.*
Annualized Volatility	*Annualized volatility equals the annualized standard deviation of the ETF returns. Volatility, defined as standard deviation of returns, measures dispersion of SB ETF strategy returns around their mean. It is computed as standard deviation of monthly returns over a 12-months investment horizon and annualized (multiplied by square root of 12)*	$$Annualised\ Volatility,$$ $$\sigma_{ETF,t} = \sqrt{12}\left(\frac{\sum_{i=t-11}^{t}(Rm_{ETF,i} - \overline{Rm_{ETF,t}})^2}{11}\right)$$ *Where $Rm_{ETF,t}$ denotes monthly ETF return at moment t and $\overline{Rm_{ETF,t}}$ is average monthly ETF return during period considered (from t-11 to t)*
Sharpe ratio	*Sharpe Ratio (SR) measures risk-adjusted performance computed by dividing the excess return of ETFs over the risk-free rate by its standard deviation of ETF returns. In other words, the return per unit of risk. Investors tend to prefer higher SRs due to risk-aversion. The higher the SR, the higher the risk-adjusted performance.*	$$SR_{ETF,t} = \frac{R_{ETF,t} - rf_t}{\sigma_{ETF,t}}$$ *Where $R_{ETF,t}$ denotes the annualized ETF return, $\sigma_{ETF,t}$ denotes the annualized ETF volatility and rf_t presents the geometrically annualized risk-free return at moment t*
Annualized Downside Deviation	*Downside Deviation (DD) measure only deviations below a specified benchmark, in our case we take the mean return of a SB ETF strategy. It is computed from monthly data and annualized (multiplied by square root of 12)*	$$DD_{ETF,t} = \sqrt{12}\sqrt{\frac{\sum_{i=t-11}^{t} d_{ETF,1}^2}{11}}$$ $$where\ \ d_{ETF,1} = \min\left[\left(Rm_{ETF,t} - \overline{Rm_{ETF,t}}\right), 0\right]$$ *Where $Rm_{ETF,t}$ denotes monthly ETF return at moment t and $\overline{Rm_{ETF,t}}$ denotes the mean monthly ETF return*

Sortino Ratio	*Sortino Ratio (SoR) is a measure of the reward of an ETF adjusted by its downside risk. In other words, it is the ratio of the return of an ETF in excess of a minimum required return (return below which investor does not accept to drop) over the standard deviation of the returns that are below this minimum required return. In our case the mean SB ETF return is chosen as minimum required return.* *A large SoR implies a lower probability of realizing large losses. Many investors do not mind upside volatility so that SoR often is the preferred measure (particularly in case of highly volatile portfolios)*	$SoR_{ETF,t} = \frac{R_{ETF,t} - rf_t}{DD_{ETF,t}}$ *Where $R_{ETF,t}$ denotes annualized ETF return at moment t, rf_t denotes the annualized risk-free rate at moment t and $DD_{ETF,t}$ denotes annualized downside deviation at moment t*

Table III: Relative Performance Metrics

Metric	*Definition*	*Formula*
Tracking Error	*Tracking Error (TE) is defined as divergence between the price behavior of a portfolio and the price behavior of its benchmark. Tracking error is reported as standard deviation percentage difference. In other words, it reports the difference between the return an investor receives and that of the benchmark they were attempting to imitate. It is also known as residual risk.*	$TE_{ETF,t} = \sqrt{12}\sqrt{\frac{1}{11}\sum_{i=t-11}^{t}\left[\left(Rm_{ETF,i} - RBm_{ETF,i}\right) - \left(\overline{Rm_{ETF,t} - RBm_{ETF,t}}\right)\right]^2}$ *Where $Rm_{ETF,t}$ is the monthly ETF return at moment t and $RBm_{ETF,t}$ denotes monthly benchmark ETF return at moment t, and $\overline{Rm_{ETF,t} - RBm_{ETF,t}}_{ETF,t}$ denotes the mean of the relative returns during the period considered (from t-11 to t)*

Information Ratio	*The Information Ratio (IFR) measures a portfolio's returns beyond the returns of a benchmark compared to the volatility of those returns.* *It is often regarded as a measure of a portfolio manager's level of skill and ability to generate excess returns relative to a benchmark. Consistency of performance is incorporated through tracking error.*	$Annualised\ Volatility, \sigma_{ETF,t} = \sqrt{12}\left(\frac{\sum_{i=t-11}^{t}(Rm_{ETF,i} - \overline{Rm_{ETF,t}})^2}{11}\right)$ *Where* $Rm_{ETF,t}$ *denotes monthly ETF return in moment t and* $\overline{Rm_{ETF,t}}$ *is average monthly ETF return during period considered (from t-11 to t)*

In particular, three ratios are used to compare the performance, namely Sharpe Ratio (SR), Sortino Ratio (SoR) and Information Ratio (IFR). Sharpe (1966, 1994) developed the so-called Sharpe Ratio which presents the excess return per unit of risk and consequently is widely popular metric to compare assets with different return and risk profiles. Simply said, the higher the SR the better the risk-adjusted performance. Due to its computational simplicity and interpretability it is one of the most referenced risk-adjusted performance measures in finance. In contrast, Sortino & Van der Meer (1991), Sortino & Price (1994) and Sortino & Forsey (1996) present another metric called the Sortino Ratio that is constructed in a similar fashion as aforementioned SR. The main difference lies in the fact that excess return is only adjusted by its downside risk instead of its total risk. Since upside volatility is not necessarily a bad thing for investors, SoR often is preferred for highly volatile portfolios. Only downside deviations (when realised returns lie below a minimum required return) are included in the metric. Hence, a large SoR indicates a considerably small change of realising a large loss. Lastly, the IFR was presented by Sharpe (1994) to compute risk on a relative return basis. IR effectively eliminates market risk, leaving only risk resulting out of active management. In one simple number, it is possible to assess a portfolio manager performed per unit of active risk taken. A high ratio implies good performance efficiency.

4.4 Results Quality

On an ending note of the empirical framework, the data quality should be critically discussed to indicate possible shortcomings of the results. With respect to both the construction of the

Smart Beta portfolios (including benchmark portfolios) and the explanatory variables in the regression and principal component analysis it can be said that output is only as good as input. Incorrect estimates of return, risk and correlation, inconsistent data reporting and the barrier to access preferred databases (at times), can cause the analysis to reach misleading and/or wrong conclusions. Nonetheless, as this presents a preliminary analysis of macroeconomic factors in explaining Smart Beta performance, we have chosen to not go deeper into back testing for these issues other than already specified in the assumptions section above. This presents an opportunity for further research and will be resumed in the limitations section.

5 Data Analysis

The following section specifically outlines the research method used throughout this study. Initially, the focus lies in data collection and processing after which the reader will be guided through the exact implementation of this Smart Beta analysis. Data is collected from a variety of databases including Datastream, Bloomberg, Morningstar, the Federal Reserve Economic Database[10] and the ETF database[11]. The subsequent analysis is conducted in Excel, Stata and R.

5.1 Data Collection

The different raw data sources are shortly presented while elaborating both on the treatment process and important considerations in the study specific context.

5.1.1 Constituent List

As a starting point, the universe of Smart Beta ETFs and their benchmarks is explored. Given the resources at hand, a frequently updated list of SB ETFs is identified and chosen from the ETF database. A total of 894 SB ETFs represent our sample. Note that the list has been growing to over 900 SB ETFs over the course of this study and continues to do so. These SB ETFs track both domestic and international markets. In line with previous statements, most of the sample consists of US domiciled SB ETFs. The database includes information about underlying benchmarks, expense ratios, inception dates and ratings which was cross-referenced with Datastream and Bloomberg to check for data inconsistencies. We found the data to be quite trustworthy. An overview of the largest SB ETFs can be found in Appendix VII. Since the list of SB ETFs does not state the relevant SB category on which the fund is based the dataset is enriched by Bloomberg information to categorise all funds correctly according to their investment factor. The same is true for the assets under management of each SB ETF at each point in time. These numbers are extracted from Datastream and added to the sample. This is

10 Source: https://fred.stlouisfed.org.
11 Source: https://etfdb.com/themes/smart-beta-etfs.

motivated by the fact that not only equally but also size weighted SB portfolios are constructed to assess robustness of the results.

5.1.2 Time Window

The rather recent increase in popularity of SB ETFs implies these financial asset classes are quite new and thus lack a long record. As this does not allow to analyse data prior to 20 years of historical data the time window of this study was set to include every SB ETF between January 2000 and December 2018. Since the aim of the study is to use as large a sample as possible in order to compute robust and reliable results the analysis focuses on US domiciled SB ETFs only. The European SB market is simply not established enough to fulfil the criteria of reliable results, in particular with the macroeconomic focus of this study. Macroeconomic variables are reported less frequent (either monthly or quarterly) so that SB ETF data is gathered monthly correspondingly.

5.1.3 Data-series

With respect to the return calculation, historical price data is downloaded from Bloomberg. Similarly, Datastream (adjusted) prices and return values are used to verify the data quality supplied by Bloomberg. Ultimately, monthly return data is based on Bloomberg SB ETF data. Monthly excess return is determined by subtracting the risk-free rate from monthly SB ETF returns. Remember that monthly SB ETF returns are computed the following way:

$$r_{m_{ETF,t}} = \frac{P_{ETF,t} - P_{ETF,t-1}}{P_{ETF,t-1}} \tag{24}$$

Risk-free rate is proxied by the US one-month Treasury rate and is extracted from the Fama-French homepage together with their value and size factors. Note that the value and size factors will be used as a reference point in the multiple regression later on. This issue will be addressed subsequently. Furthermore, each of the SB ETFs was (manually) assigned a benchmark ETF based on its prospectus and/or factsheet (see Table IV for an overview of assigned benchmark ETFs). The decision to assign benchmark ETFs manually is motivated by the fact that returns of benchmarks declared by SB fund providers do not fully account for trading costs and management expenses. The identified benchmark ETFs are carefully chosen to alleviate this concern and allow for a risk-adjusted comparison of cost-effective passive vehicles. All benchmark ETFs below are assessed respective to the assigned benchmark ETF return below.

Table IV: Overview of Assigned Benchmark ETF for each SB ETF category

Benchmark	Dividend	Earnings-Weighted	Equal-Weighted	Fundamentals	Growth	Low Volatility	Momentum	Multi-Factor	Non-traditional	Quality	Value	Total by benchmark
EWSC Invesco S&P SmallCap 600 Equal			3									3
First Trust Dow Jones Global Select Dividend	14		7	2				5				28
Invesco DWA Momentum							6					6
Invesco Dynamic Energy Exploration & Production									4			4
Invesco Fundamental High Yield				17								17
Invesco High Yield Equity Dividend Achievers	16											16
Invesco Russell 2000 Equal			21									21
Invesco Russell 2000 Pure Value											3	3
Invesco Russell MidCap Equal			1									1
Invesco Russell Top 200 Pure Value											3	3
Invesco S&P 500 Pure Growth					3							3
Invesco S&P 500 Quality										5		5
Invesco S&P SmallCap High Dividend Low Volatility	3		1									4
Invesco Zacks Multi-Asset Income		7						56				63
iShares Core S&P 500			3					14				17
iShares Edge MSCI Intl Quality								2		3		5
iShares Edge MSCI Min Vol EAFE						9		2				11
iShares Edge MSCI Min Vol USA						2						2
iShares Edge MSCI USA Momentum							1					1
iShares MSCI EAFE											30	30
iShares MSCI EAFE Growth					10							10
iShares Russell 1000				6				1			3	10

iShares Russell 1000 Growth					6							6
iShares Russell 2000											1	1
iShares Russell 2000 Growth					3							3
iShares S&P Mid-Cap 400 Value				1							5	6
iShares S&P SmallCap 600 Value											6	6
JHancock Multi-factor Large Cap								12				12
ProShares S&P 500 Dividend Aristocrats ETF												
SPDR S&P 1500 Momentum							2					2
SPDR S&P 500 Fossil Fuel Reserves Free									1			1
iShares Core S&P 500											10	10
SPDR® Portfolio S&P 500 Growth					11							11
SPLV Invesco S&P 500 Low Volatility						5		1				6
Total per SB category	33	7	36	26	33	16	9	93	5	8	61	327

Similarly, monthly benchmark returns are computed by downloading monthly adjusted price information from Datastream. In the next step, excess returns are computed by subtracting the risk-free rate. As a last step, market value data from Datastream allows to scale SB ETF categories based on their relative market value. Market value is proxied by AUM.

Likewise, macroeconomic information is extracted from both Datastream (if available) and the Federal Reserve Economic Database. More specifically, GDP, interest rate (proxied by US 3-month Treasury rate), consumer price index, and unemployment data are extracted from Datastream. It is important to point out that GDP data is reported quarterly and beginning of respective months whereas both the consumer price index and unemployment numbers are reported mid-monthly. Interest rate is reported beginning of each month. Likewise, the remaining four macroeconomic variables foreign direct investment, industrial production index, money supply and oil price data are supplied by the Federal Reserve Economic Database. Again, FDI is reported quarterly similar to GDP. IPI, money supply and oil prices are reported beginning of each month. In order to rely on monthly values for both GDP and FDI, averages between the differences of quarters are computed and added to the sample for $t + 1$ and $t + 2$ accordingly. The exact alignment of these eight macro variables will be discussed in the section 5.4 (Research Method).

5.2 Data Filtering

Proceeding from the initial sample of 894 SB ETFs, several screening steps are necessary. For example, a minimum record of twelve monthly observations is imposed because their short histories might lead to inconsistent conclusions. Consequently, all SB ETFs with an inception date in 2018 are removed from the sample. This reduced the sample to 797 funds. Note that 97 SB ETFs emerged in 2018 alone, thereby underlining the rise in popularity of these funds. Beyond that, this study focuses on equity returns so that all non-equity SB ETFs are removed from the sample lowering the sample size to 665 SB ETFs. Unfortunately, no data was available for another 175 funds resulting in 490 remaining SB ETFs. The last filtering step removes 163 SB ETFs which invest in equities outside of the US. Since the macroeconomic analysis relies on US data, we decided to exclude this group of funds from the sample because analysis results are likely to be distorted. The focus of this study lies strictly on US domiciled funds that invest within the US. In sum, this leaves 327 SB ETFs in the sample. Interestingly, 27 out of these 327 funds died during the time window considered (2000-2018). In contrast to several SB studies such as Glushkov (2016) and Meyer (2016), we decided to include these SB ETFs in our sample to base the analysis on a survivorship bias free dataset. Generally, SB ETFs are being closed as a result of lacking investor interest and/or poor fund performance (Glushkov, 2016). Dead SB ETFs are spread across most SB strategies and after individual inspection no reason to exclude them is found (see Appendix VIII for an evolution of dead Smart Beta ETFs between 2000 and 2018). Hence, including all SB ETFs works in favour of finding no evidence of significant benchmark-adjusted outperformance. This suggests that the probability of detecting robust outperformance of SB strategies would likely be higher by excluding the dead SB ETFs.

Remember that Morningstar (2018) currently lists 18 different Smart Beta strategies. Our final sample, however, only consists of 11 strategies, namely (1) dividend, (2) earnings-weighted, (3) equal-weighted, (4) fundamentals, (5) growth, (6) low volatility, (7) momentum, (8) multi-factor, (9) non-traditional commodity, (10) quality and (11) value. The remaining categories are removed due to their considerably low presence. We set a minimum amount of 5 SB ETFs for each category which explains why none of the remaining categories are investigated. Moreover, Glushkov (2016) argues that many of the well-known value SB ETFs (e.g. Vanguard, iShares, Schwab) became popular long before the birth of SB ETFs and as a result distort the results. Yet, we choose to include all (value) SB ETFs to be consistent. We argue that no credible reason that supports the exclusion is found. This opinion is supported in recent academic research, for example both Amenc (2013) and Meyer (2016) support our approach. In sum, aforementioned screening steps reduced the dataset to a final sample of 327 SB ETFs and 34 distinct benchmark ETFs.

5.3 Creating Portfolios

A crucial last step includes the construction of portfolios. Two different weighting schemes will be applied in order to assess the robustness of results. On one hand, equal weighted portfolios are built in the following way:

$$r_P = \sum_{n}^{1} \frac{1}{n} * r_i \tag{25}$$

where r_P is the portfolio return, r_i the specific SB ETF return and n denotes the number of funds at the end of each month. Consequently, eleven equally-weighted SB portfolios are built. In theory, this portfolio is used to give each fund the same weight and thus can tell about the average SB performance depending on each investment factor. In contrast, it can be argued that extreme risk and return characteristics put too much weight on small SB funds and too little weight on prominent and established SB funds. Likewise, results can be similarly distorted. Therefore, AUM are used to represent the SB universe in terms of its market value. The sum of AUM values for each SB category indicate the market size of each strategy where the fraction of each strategy's total fund size represents the weight of each category in the value weighted portfolio (see Table VII for evolution of asset value for each SB category). Note that the last month's fund size is used as weights to determine the value weighted portfolio for the upcoming month respectively. The latter portfolios are expected to more closely mirror the actual universe of SB ETFs returns whereas the equally-weighted portfolios strictly assess the average SB ETF performance. Mathematically, the value-weighted portfolio is calculated as follows:

$$x_i = \frac{AUM\ of\ i}{Total\ AUM\ All\ Securities} = \frac{AUM_i}{\sum_j AUM_j} \tag{26}$$

With

$$r_t = \sum \left(\frac{w_{i,t-1} * r_{i,t}}{w_{i,t-1}} \right) \tag{27}$$

Where r_t is the value weighted portfolio return, $w_{i,t-1}$ is the weight of security i at time t-1 and $r_{i,t}$ is the return of security i at time t.

5.4 Research Method

A simple regression will be applied initially to test for statistically significant out/underperformance of SB categories. With respect to monthly excess returns, the model can be written as:

$$r_{SB\ ETF_t} = \alpha + \beta_{SB\ Benchmark_t}(r_{SB\ Benchmark}) + \epsilon \tag{28}$$

where $r_{SB\ ETF_t}$ is the realised return for the ith category at time t and $\beta_{SB\ Benchmark_t}$ is the exposure to the respective monthly benchmark ETF returns at time t.

Furthermore, multi-factor models have been developed to assess the variation in ex-post security returns beyond what is provided by standard market models (Bilson et al., 2001). Most multi-factor models assume perfect integration where security returns are modelled as a linear relation to selected global risk sources (e.g. Ferson & Harvey, 1994; Dumas & Solnik, 1995). A critical problem of multi-factor models lies in determining these risk sources. Remember that earlier macroeconomic multi-factor studies argued that the dominant amount of equity factors lies between one to five factors. However, much contention surrounds this dilemma. Particularly difficult is the choice of initial factors that tend to be based on arbitrary and subjective justification. Fama (1991) refers to as an unavoidable problem in this field of research. Individual judgement about the relevance of previously included factors often serves as a preferred starting point. However, research suggests multitudinous relevant factors. Variables such as goods prices, real activity, interest rates, political risks and oil prices among many others are frequently mentioned. In the aforementioned study of Harvey et al. (2016) hundreds of published papers since 1967 are studied in order to catalogue a total of 313 equity factors. These underlines both the direction of where current research is heading and the underlying difficulty of choosing dominant equity factors. In light of above considerations and balancing theoretical arguments and prior studies, the most frequently mentioned factors among the categories *economic conditions, interest rate and monetary policy, price levels, international activities* are selected. Albeit findings with respect to the selected macroeconomic factors are rather contradictory, we chose to focus on well-established factors instead of screening for new ones to not further fuel the growth of Cochrane's "factor zoo". Thus, a total of eight explanatory variables (1) GDP[12], (2) industrial production index[13], (3) unemployment rate[14], (4) interest rate[15], (5) money supply (M2)[16], (6) consumer price index[17], (7) oil price[18], (8) FDI[19] are chosen. When these variables are included, the testable model initially contains eight factors in total.

Foremost, however, time delays with respect to the generation of macroeconomic variables require further investigation. The transmission and incorporation of macroeconomic information into stock and market prices is not always instantaneous. It is possible that due to reporting delays the incorporation of macroeconomic variables into equity prices creates a lag.

12 Gross domestic product (GDP) is the market value of goods and services produced by labor and property in the United States, regardless of nationality; GDP replaced gross national product (GNP) as the primary measure of U.S.
production in 1991. Series are in billions of dollars and seasonally adjusted at annual rates.

13 The Industrial Production Index (INDPRO) is an economic indicator that measures real output for all facilities located in the United States manufacturing, mining, and electric, and gas utilities (excluding those in U.S. territories). It measures movements in production output and highlights structural developments in the economy. Growth in the production index from month to month is an indicator of growth in the industry.

14 The Current Population Survey (CPS; household survey) provides information on the labor force, employment, and unemployment. It is a sample survey of about 60,000 eligible households conducted by the U.S. Census Bureau for the U.S. Bureau of Labor Statistics (BLS). The reference period is generally the calendar week that contains the 12th day of the month. The sample is selected to reflect the entire civilian non-institutional population. (Source: http://www.bls.gov/cps/cps_htgm.htm).

15 United States, Treasury Bill Rate - 3 Month (Source: Federal Reserve, United States (http://www.federalreserve.gov/).

16 M2 includes a broader set of financial assets held principally by households. M2 consists of M1 plus: (1) savings deposits (which include money market deposit accounts, or MMDAs); (2) small-denomination time deposits (time deposits in amounts of less than $100,000); and (3) balances in retail money market mutual funds (MMMFs) and is measured in billions of US$.

17 The Consumer Price Index (CPI) is a measure of the average change over time in the prices of consumer items - goods and services that people buy for day-to-day living. The quantity and quality of these items are kept essentially unchanged between major revisions so that only price changes will be measured. All taxes directly associated with the purchase and use of items are included in the index. (Source: http://www.bls.gov/cpi/cpi_methods.htm).

18 Defined as the US$ spot price of crude oil per barrel (Source: West Texas Intermediate (WTI) - Cushing, Oklahoma: DCOILWTICO).

19Rest of the world foreign direct investment in U.S measured in millions of US$ (For a detailed description, how this series is constructed, see: https://www.federalreserve.gov/apps/fof/SeriesAnalyzer.aspx?s=FA263092001&t=).

A contemporaneous model, on the other hand, measures all variables at time t and connotes assumptions about contemporaneous association. Thus, if applicable, the empirical model in this study lags explanatory variables to incorporate delays in the publication of (macroeconomic) information. International Monetary Fund (IMF) data dissemination standards are used to minimise data issues in regard to time delays[20]. Based on IMF advice, GDP, IPI and FDI are shall be lagged by two months. All remaining explanatory variables are supposed to be lagged by one month. Note that unemployment rates and consumer prices are reported mid-monthly though. Due to inconsistent reporting dates, we additionally used a vector autoregression (VAR) model in order to be able to empirically justify the optimal lag length for each variable.

The VAR model was provided by Sims (1980) and is regarded as the extension of the univariate autoregressive model to multivariate time series. Most importantly, the model contributes a useful macro econometric framework to capture the dynamic behaviour of economic and financial time series. An eight-variable VAR is estimated using monthly U.S. data on the percentage change of IPI, MS, IR, OP, UR, CPI, FDI, and GDP from 2000 to 2018. At this point, we examine the model as a way to display the VAR toolkit, whereas criticism is reserved for the limitations section at the end. Starting with the reduced form VAR we express each variable as a linear function of its own past values, the past values of all other variables being considered and a serially uncorrelated error term. Hence, in our setup, the VAR involves eight equations: current IPI as a function of past values of IPI, MS, IR, OP, UR, CPI, FDI, and GDP; and equivalently for the remaining seven equations. Each equation is estimated by ordinary least squares regression. The number of lagged values to include in each equation be determined by a number of different methods, and we will use one lag[21] in our example (see Appendix IX).

20 Respective standards can be found on IMF homepage: https://dsbb.imf.org/.

21 Frequently, the AIC or BIC information criteria are used; for a discussion, see Lütkepohl (1993, Chapter 4). Please note that there are several research papers based on the question on which lag length selection criterion to employ. The AIC is one of the leading selection methods but is best applicable when the sample size is small (Hurvich & Tsai, 1989). This is confirmed by Liew (2004), who claims AIC and FPE to be superior than the other criteria in the case of small sample (60 observations and below), whereas HQC is best applicable to identify the true lag length with relatively large samples (120 or more observations). In our case, FPE, AIC, HQIC and BIC provide a common lag length of one, which allows us to be indifferent within a choice of a particular selection criterion.

Note that as a preliminary check, we verify that our eight series are stationary (see Appendix IX). We include four lagged differences to eliminate serial correlation in the error term of the Dickey-Fuller regression. In all eight cases, we comfortably reject the presence of a unit root in the series at a 5% significance level. Therefore, we conclude that VAR analysis can be performed on the eight series without differencing. As mentioned above, we assess an optimal lag length of one, since all of the criteria support this lag length to choose. Assessing the validity of our VAR, we test for stability and for autocorrelation of the residuals. We can confirm the stability of our system, since all eigenvalues lie inside the unit circle (see Appendix IX). However, when applying the Lagrange-multiplier test, we can reject the null of no residual autocorrelation at orders 1 through 3 at any conventional significance level, so we have evidence to contradict the validity of our VAR. Only with a lag level of four we obtain the desired result of no autocorrelation (see Appendix IX). Nevertheless, please regard this weakness of our model as part of the limitations, since the appropriate lag length is one of the most complex steps to estimate (Hatemi-J, 2003). Hence, we will stick to our previously defined lag length of one in the course of this analysis. This is in line with academic methods and yielded the most robust model results. Nonetheless, further research can dig deeper in the right choice of lags with respect to each variable (see Study Limitations in section 8). Consequently, using monthly return intervals, the macroeconomic model can be formulated as:

$$r_{it} = \alpha_i + \beta_1 GDP_{t-1} + \beta_2 IPI_{t-1} + \beta_3 UR_{t-1} + \beta_4 IR_{t-1} + \beta_5 MS_{t-1} + \beta_6 CPI_{t-1} + \beta_7 OP_{t-1} + \beta_8 FDI_{t-1} + \epsilon_{it} \qquad (29)$$

where $r_{it}$is the realised return for the ith category at time t, $GDP_{t-1}$is the percentage change in the US' GDP at time t-1, $IPI_{t-1}$is the percentage change in the Industrial Production Index for the US at time t-1, $UR_{t-1}$is the percentage change in the unemployment rate in the US at time t-1, $IR_{t-1}$is the percentage change in the Fed's interest rate at time t-1, $MS_{t-1}$is the percentage change in US' money supply variable at time t-1, $CPI_{t-1}$is the percentage change in the US' consumer price index at time t-1, $OP_{t-1}$is the percentage change in the US' oil price at time t-1, $FDI_{t-1}$is the percentage change in foreign direct investment at time t-1.

In addition, research suggests that popular SB strategies deliver added value mainly due to its tilt towards the value and/or size factor (Asness, 2006). Similarly, Chow (2011) investigates

RAFI indices and concludes that value premium is able to explain most or all of its performance. Lastly, Plzakha, Uppal, & Vilkow (2014) find that a higher systematic return of equal weighted strategies relative to cap-weighted derives from a comparatively high exposure to value, size and market factors. As a result, the macroeconomic model is extended to ten factors by including value and size factors. The analysis of a strictly macroeconomic eight factor model versus an extended ten factor model that includes well-documented size and value factors allows to assess to what extent macroeconomic variables can explain SB excess returns versus value and size factors. Therefore, the extended model can be written as:

$$r_{it} = \alpha_i + \beta_1 Size_{t-1} + \beta_2 Value_{t-1} + \beta_3 GDP_{t-1} + \beta_4 IPI_{t-1} + \beta_5 UR_{t-1} + \beta_6 IR_{t-1} \\ + \beta_7 MS_{t-1} + \beta_8 CPI_{t-1} + \beta_9 OP_{t-1} + \beta_{10} FDI_{t-1} + \epsilon_{it} \tag{30}$$

In the final stage of this analysis, a PCA is conducted to verify the impact of each macroeconomic variable and substantially reduce the dimensionality of the model. The a priori selection of variables for the PCA analysis commences with the earlier identified macroeconomic variables. The central empirical finding that only a handful of factors can explain a large fraction of the variance of many macroeconomic series has been confirmed by many studies (Giannone, Reichlin, & Sala, 2004; Watson, 2004).

This PCA involves three steps. Firstly, for the US market of SB ETFs, a number of (macroeconomic) variables are identified that are able to explain a significant portion of return variation. Naturally, the same eight macroeconomic variables are used. Secondly, from these variables a number of principal components are extracted through PCA. Thirdly, the equity returns for the US market are then regressed against these synthetic variables (PCs) to assess to which underlying macroeconomic forces the SB strategies have similar sensitivities. The PCA aims to confirm the results of the multiple regression while reducing the complexity of the model. Ideally, the “reduced” models perform relatively robustly without losing much explanatory power.

5.5. Expectations

As highlighted in the literature review there is a relationship between macroeconomic variables and stock returns. Further evidence is provided by Homa & Jaffee (1971) and Boudoukh & Richardson (1993). Stock levels are said to be positively related to future levels of real activity such as GDP and industrial production. This finding seems logical as returns are a function of future cash flows. Prominent researchers such as Fama (1981), Harris & Oppler (1990) and James (1985) find a positive relationship between real activity and equity returns. Thus, a priori, changes in GDP and industrial production are expected to positively affect SB ETF returns. The last economic activity measure, unemployment, possesses a bundling property with respect to future interest rates, equity risk premia and corporate earnings according to Boyd, Hu, & Jagannathan (2005). Thus, the announcement of surging unemployment positively affects equity prices during economic expansions and negatively in times of contractions so that SB ETF returns are expected to be positively/negatively affected during bullish/bearish markets.

Common stock is often regarded as a hedge against inflation because equity can be viewed as a contingent claim on the realm assets of a firm. Inflation will drive the value of these contingent claim upwards. Day (1984) argues that proportionate price increases should not affect real rates of return on equity. However, monetary assets should be independent of price level fluctuations so that it is only the real component of the firm's return that will be hedged against changes in inflation (Hong, 1977). Previous tests have found a negative relationship between inflation and nominal equity returns (Fama & Schwert, 1977; Gultekin, 1983). Thus, changes in inflation are expected to negatively influence SB ETF returns. Similarly, oil price are another popular measure of price level stability while changes in oil prices clearly seem to affect equity returns. Remember that Faff & Brailsford (1999) indicate that oil price changes are clearly affecting Australian equity returns where the direction of the relationship depends on the particular industry. Therefore, it seems impossible to specify the direction of the relationship since little is known about the exact portfolio holdings. It is hypothesised that SB strategies are positively or negatively affected based depending on respective (unknown) portfolio holdings.

Monetary portfolio theory implies that changes in money supply affect the equilibrium position of money, which in turn alters the composition and price of an investor's portfolio (Cooper, 1974). Furthermore, changes in money supply affect real economic variables and thus have a lagged impact on stock returns (Rogalski & Vinso, 1977). Summing up, the relationship between changes in money supply and stock returns is expected to be positive. With respect to interest rates as a second monetary policy measure, Blanchard (1981) states that interest rate changes lead to changes in anticipated profits and discount rates. Interest rate changes are usually exploited to steer inflation levels. Since a decrease in the interest rate increases money supply, the relationship should be reverse to the hypothesised money supply effect, thus changes in interest rates are expected to negatively affect SB ETF strategies.

The last explanatory variable FDI affects equity returns because it is a source of capital, complements domestic private investment and boosts job opportunities. Adam and Tweneboah (2008) observe a triangular causal relationship in which FDI stimulates economic growth, economic growth stimulates stock market development, and FDI fuels stock market development. Both Adam and Tweneboah (2008) and Mohammed et al. (2009) find a positive relationship between FDI changes and stock market development in Ghana and Pakistan respectively. Although the results apply to developing countries only the relationship between FDI changes and SB ETF returns is expected to be positive in the United States as well. Find a summary of expected relationships between changes in macroeconomic variables and SB returns in the Table V.

Table V: Expectations for macroeconomic relation with SB ETF excess return in multi-factor models

Expectations	*GDP*	*IPI*	*UR*	*IR*	*MS*	*CPI*	*OP*	*FDI*
Direction of relationship	+	+	+ *(bullish)* - *(bearish)*	-	+	-	+/- *(depending on industry)*	+

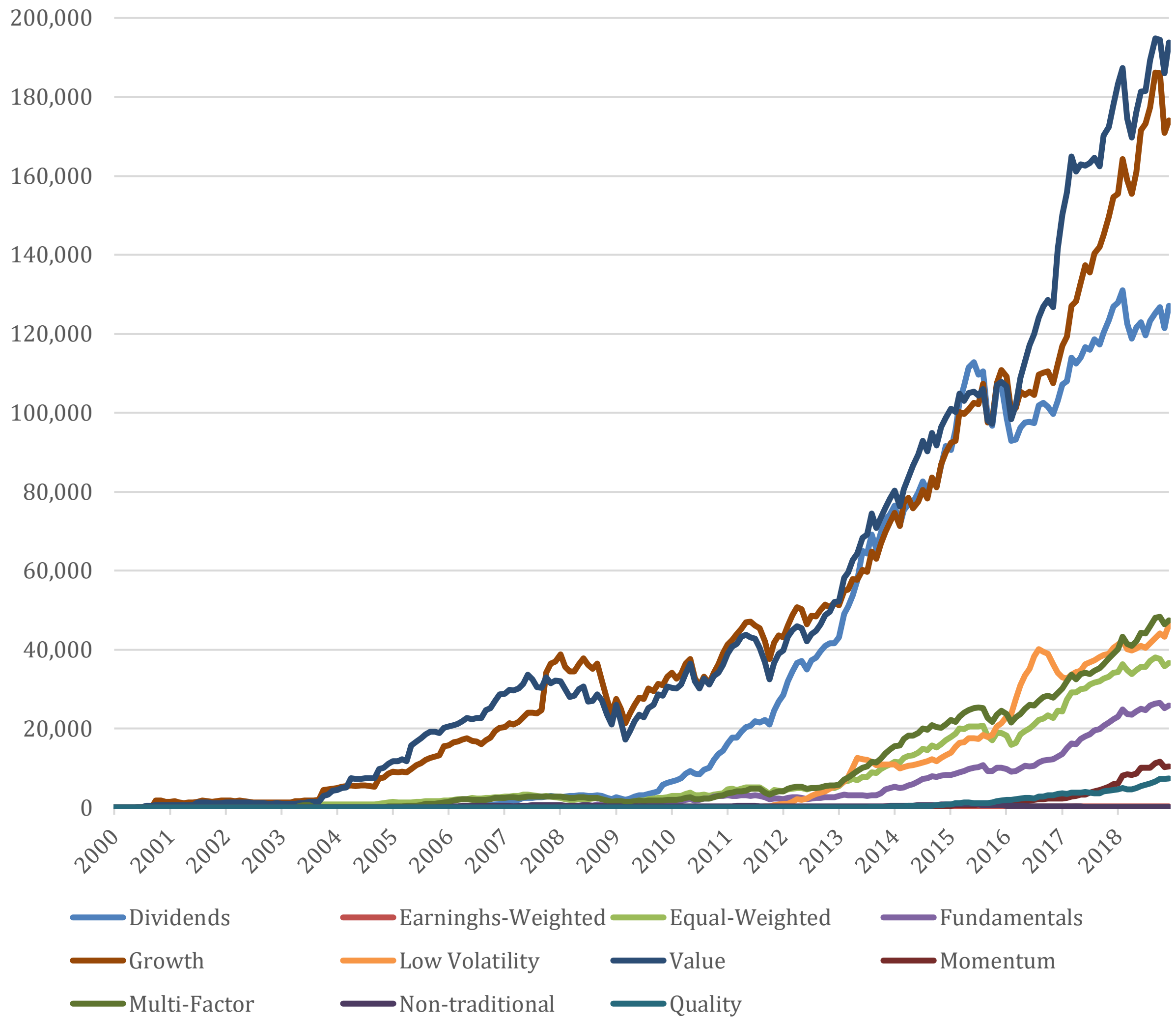
6 Results

The following section is divided into descriptive statistics, risk-adjusted performance (simple regression), multi-factor models and lastly PCA.

6.1 Descriptive Statistics

We use a sample of 327 US-domiciled SB ETFs targeting only domestic markets in order to analyse asset flows into this segment of the ETF market. Figure 1 demonstrates the steady growth in AUM per Smart Beta strategy between 2000 and 2018 increasing its shares from nearly zero in 2000 to nearly 200 billion in AUM.

Figure I: Time-series of AUM per Smart Beta Strategy between 2000 and 2018 (in millions)



Having a closer look at the absolute number of active SB ETFs in total, particularly since 2009, SB funds have been attracting an increasing share of assets into the ETF market simply due to a growth in available funds. Their absolute number nearly tripled from 106 active SB ETFs at year-end 2009 to 310 end of 2017 with a slight decline in the year 2018 (see Table VI). Interestingly, this marks the first decline in available SB ETFs since inception. Recalling the period after the financial crisis 2008 this finding seems reasonable, since investors mainly invested into non-SB products during the financial crisis but switched over to the allegedly "safe investment harbour" when investing in SB ETFs post-crisis. The recent decline in SB ETFs can be linked to an unusual high amount of funds being closed in recent years paired with generally poor performances on financial markets. Find an overview of closed funds per year for each SB strategy in Appendix VIII.

Table VI: Number of active SB ETFs per category, between 2000 and 2018

Smart Beta factor	*Dividend*	*Earnings-Weighted*	*Equal-Weighted*	*Fundamentals*	*Growth*	*Low Volatility*	*Momentum*	*Multi-Factor*	*Non-traditional*	*Quality*	*Value*	*Total*
2000	*0*	*0*	*0*	*0*	*6*	*0*	*0*	*0*	*0*	*0*	*7*	*13*
2001	*0*	*0*	*0*	*0*	*7*	*0*	*0*	*0*	*0*	*0*	*8*	*15*
2002	*0*	*0*	*0*	*0*	*7*	*0*	*0*	*0*	*0*	*0*	*8*	*15*
2003	*0*	*0*	*1*	*0*	*7*	*0*	*0*	*0*	*0*	*0*	*8*	*16*
2004	*0*	*0*	*1*	*0*	*10*	*0*	*0*	*0*	*0*	*0*	*13*	*24*
2005	*1*	*0*	*1*	*0*	*14*	*0*	*0*	*2*	*0*	*0*	*18*	*36*
2006	*12*	*0*	*4*	*2*	*17*	*0*	*0*	*2*	*0*	*0*	*20*	*57*
2007	*16*	*0*	*5*	*8*	*17*	*0*	*1*	*23*	*2*	*0*	*23*	*95*
2008	*18*	*0*	*5*	*10*	*17*	*0*	*1*	*24*	*2*	*0*	*23*	*100*
2009	*18*	*0*	*5*	*10*	*19*	*0*	*1*	*25*	*2*	*0*	*26*	*106*
2010	*18*	*0*	*7*	*11*	*21*	*0*	*1*	*25*	*2*	*0*	*32*	*117*
2011	*21*	*1*	*8*	*11*	*22*	*4*	*1*	*28*	*2*	*0*	*33*	*131*
2012	*24*	*2*	*13*	*11*	*26*	*5*	*2*	*35*	*2*	*0*	*39*	*159*
2013	*24*	*3*	*16*	*19*	*26*	*8*	*3*	*37*	*2*	*1*	*41*	*180*
2014	*25*	*4*	*20*	*19*	*29*	*8*	*3*	*40*	*2*	*1*	*43*	*194*
2015	*29*	*6*	*26*	*23*	*29*	*11*	*6*	*59*	*2*	*1*	*48*	*240*
2016	*31*	*6*	*24*	*25*	*31*	*11*	*7*	*77*	*2*	*4*	*52*	*270*
2017	*33*	*5*	*31*	*24*	*32*	*14*	*9*	*90*	*5*	*8*	*59*	*310*
2018	*33*	*5*	*30*	*22*	*32*	*14*	*9*	*83*	*5*	*8*	*59*	*300*

Table VII provides an overview of the relative weight of active SB ETFs per category, whereas Appendix X visualizes the time-series of AUM per SB strategy between 2000 and 2018.

Table VII: Relative weight of active SB ETFs per category (measured in AUM) between 2000 and 2018 *(Note that due to rounding not all rows sum to 100%)*

Smart Beta factor	*Dividend*	*Earnings-Weighted*	*Equal-Weighted*	*Fundamentals*	*Growth*	*Low Volatility*	*Momentum*	*Multi-Factor*	*Non-traditional commodity*	*Quality*	*Value*
2000	*0%*	*0%*	*0%*	*0%*	*73.6%*	*0%*	*0%*	*0%*	*0%*	*0%*	*26.4%*
2001	*0%*	*0%*	*0%*	*0%*	*59.2%*	*0%*	*0%*	*0%*	*0%*	*0%*	*40.8%*
2002	*0%*	*0%*	*0%*	*0%*	*55.9%*	*0%*	*0%*	*0%*	*0%*	*0%*	*44.1%*
2003	*0%*	*0%*	*7.2%*	*0%*	*48.4%*	*0%*	*0%*	*1.5%*	*0%*	*0%*	*42.9%*
2004	*0%*	*0%*	*6.1%*	*0%*	*40.8%*	*0%*	*0%*	*1.0%*	*0%*	*0%*	*52.1%*
2005	*0.1%*	*0%*	*4.5%*	*0%*	*40.0%*	*0%*	*0%*	*3.4%*	*0.2%*	*0%*	*51.8%*
2006	*1.7%*	*0%*	*4.8%*	*0.1%*	*36.5%*	*0%*	*0%*	*4.3%*	*0.7%*	*0%*	*51.9%*
2007	*3.5%*	*0%*	*3.3%*	*0.4%*	*47.4%*	*0%*	*0%*	*3.5%*	*0.7%*	*0%*	*41.3%*
2008	*4.2%*	*0%*	*2.6%*	*0.9%*	*47.1%*	*0%*	*0%*	*2.9%*	*0.4%*	*0%*	*42.0%*
2009	*8.2%*	*0%*	*3.5%*	*1.9%*	*43.4%*	*0%*	*0%*	*2.4%*	*0.3%*	*0%*	*40.3%*
2010	*14.5%*	*0%*	*3.6%*	*3.0%*	*39.3%*	*0%*	*0%*	*3.1%*	*0.2%*	*0%*	*36.2%*
2011	*22.2%*	*0%*	*3.5%*	*1.9%*	*36.1%*	*0.6%*	*0%*	*3.3%*	*0.2%*	*0%*	*32.2%*
2012	*25.4%*	*0%*	*3.0%*	*1.6%*	*31.6%*	*3.2%*	*0%*	*3.3%*	*0.1%*	*0%*	*31.7%*
2013	*27.8%*	*0%*	*4.1%*	*1.8%*	*27.1%*	*4.1%*	*0.1%*	*5.5%*	*0.1%*	*0.1%*	*29.3%*
2014	*26.8%*	*0%*	*5.0%*	*2.4%*	*26.3%*	*3.9%*	*0.1%*	*6.2%*	*0.2%*	*0.2%*	*28.9%*
2015	*26.5%*	*0%*	*4.7%*	*2.5%*	*27.4%*	*5.3%*	*0.3%*	*6.0%*	*0.1%*	*0.5%*	*26.7%*
2016	*22.2%*	*0.1%*	*5.3%*	*2.8%*	*24.2%*	*7.3%*	*0.5%*	*6.2%*	*0.0%*	*0.7%*	*30.6%*
2017	*20.9%*	*0%*	*5.6%*	*3.7%*	*25.5%*	*6.7%*	*1.0%*	*6.4%*	*0.0%*	*0.7%*	*29.4%*
2018	*19.0%*	*0%*	*5.5%*	*3.9%*	*26.0%*	*6.9%*	*1.6%*	*7.1%*	*0.0%*	*1.1%*	*29.0%*

Differently to the results of Glushkov (2016), there is a dominance of growth- and value-oriented funds followed by dividend-oriented funds in the entire sample of domestic SB ETFs as confirmed by the evidence presented in Table VII. In terms of overall AUM, the first two categories command around 35% of all assets in domestic equity SB space, with Dividend strategy representing roughly 20% of the domestic equity Smart Beta products, respectively. Together, these three categories command around 75% of the entire SB ETF market by year-end 2018.

Getting deviating results than the one from Glushkov (2016) is explained by the amount of Smart Beta ETFs included in the underlying data set. Contrary to the latter author, we consider all SB ETFs as they have been labelled as such. Consequently, we have a higher share of growth and value stocks accounting for the aforementioned results. However, as can be seen from Appendix X, there is a diminishing trend in growth and value SB ETFs over time. Since the breakthrough of Fama and French three-factor model (1993) further research (despite its controversy) has led to a shift towards more exotic SB strategies throughout the last years.

Exemplarily, dividend is reaching its peak with a relative weight of 27.8% of all active SB ETFs in 2013, before declining steadily to 19% in 2018. Following the strong annualized excess return (up to 60%) in 2010, the dividend strategy reaches its highest market share in the SB ETF market in both the equally- and size-weighted portfolio. However, since 2018 excess returns steadily declined and are negative by the end of the year. Moreover, multi-factor strategies combine a variety of factors such as value, size, momentum et cetera in order to improve risk-adjusted performance. The popularity of combining several factor tilts fuels demand for this SB strategy. Whereas the market share of multi-factor strategies keeps increasing to 7.1% in 2018, the number of active funds in this category tripled from 2011 to 2017, reaching its peak in 2017 with a total of 90 active SB ETFs.

When analysing Table VIII and IX of absolute and relative return performances of both equally-weighted and size-weighted domestic SB ETFs over their underlying benchmark, it appears that seven out of eleven SB categories are able to outperform their respective raw benchmarks since inception (equally-weighted scheme). Apparently, four categories being earnings-weighted, momentum, non-traditional and quality fail to outperform. Remarkably, quality SB ETFs are characterized by a short track record of active funds (starting in 2013). Moreover, earnings-weighted, momentum and non-traditional account for the smallest fraction of active SB ETFs by year-end 2018. Taking this into consideration, the result of an underperformance might be attributable to a selection bias. However, earnings-weighted performs drastically worse than its respective raw benchmark, likely due to the subjective choice of the underlying benchmark and/or fund-specific anomalies.

Table VIII: Absolute and relative return performance of equally-weighted domestic equity Smart Beta ETFs and their benchmarks by category, as of Dec 31, 2018

Smart Beta category	Return type	1 year	3 years	5 years	Since Inception
Dividend	Smart Beta	-15.60%	2.65%	-0.16%	-0.67%
	Declared Benchmark	-6.54%	3.92%	2.84%	-4.12%
	Benchmark-adjusted	-9.20%	-1.28%	-3.01%	3.48%
Earnings-Weighted	Smart Beta	-24.73%	-8.57%	-19.15%	-13.60%
	Declared Benchmark	-4.87%	2.17%	-3.87%	-0.82%
	Benchmark-adjusted	-20.18%	-10.83%	-15.35%	-12.82%
Equal-Weighted	Smart Beta	-15.14%	-0.86%	-2.40%	-2.36%
	Declared Benchmark	-4.30%	4.54%	-2.80%	-2.94%
	Benchmark-adjusted	-11.01%	-5.44%	-5.22%	0.57%
Fundamentals	Smart Beta	-15.78%	3.07%	0.63%	0.78%
	Declared Benchmark	-4.15%	1.83%	1.01%	-0.30%
	Benchmark-adjusted	-11.82%	1.26%	-0.38%	1.10%
Growth	Smart Beta	-9.66%	5.91%	4.72%	0.25%
	Declared Benchmark	1.32%	7.47%	7.30%	-0.99%
	Benchmark-adjusted	-11.15%	-1.56%	-2.59%	1.26%
Low Volatility	Smart Beta	-7.19%	4.72%	3.89%	6.05%
	Declared Benchmark	-1.98%	3.56%	4.16%	5.38%
	Benchmark-adjusted	-5.30%	1.17%	-0.27%	0.67%
Value	Smart Beta	-15.36%	1.83%	-0.14%	2.08%
	Declared Benchmark	-6.06%	3.94%	2.27%	0.44%
	Benchmark-adjusted	-3.45%	-2.13%	-2.42%	1.67%
Momentum	Smart Beta	-9.74%	4.76%	5.83%	2.61%
	Declared Benchmark	2.61%	7.87%	8.52%	4.69%
	Benchmark-adjusted	-12.55%	-3.14%	-2.70%	-2.08%
Multi-Factor	Smart Beta	-13.55%	2.94%	2.45%	3.12%
	Declared Benchmark	-4.25%	5.02%	-0.55%	-1.04%
	Benchmark-adjusted	-10.77%	-1.00%	3.10%	4.62%
Non-traditional	Smart Beta	-24.72%	-13.83%	-20.61%	-11.32%
	Declared Benchmark	-8.82%	-14.08%	-13.71%	-7.31%
	Benchmark-adjusted	-16.16%	0.26%	-6.93%	-4.03%
Quality	Smart Beta	-7.92%	5.40%	5.99%	3.69%
	Declared Benchmark	-2.44%	6.53%	7.48%	4.24%
	Benchmark-adjusted	-5.57%	-1.14%	-1.50%	-1.20%

Table IX: Absolute and relative return performance of domestic size-weighted equity Smart Beta ETFs and their benchmarks by category, as of Dec 31, 2018

Smart Beta category	Return type	1 year	3 years	5 years	Since Inception
Dividend	Smart Beta	-11.74%	2.51%	0.22%	-0.01%
	Declared Benchmark	-5.42%	4.86%	3.75%	-3.49%
	Benchmark-adjusted	-6.41%	-2.37%	-3.55%	3.52%
Earnings-Weighted	Smart Beta	-26.07%	-7.49%	-20.37%	-13.85%
	Declared Benchmark	-4.80%	3.52%	-3.49%	-0.61%
	Benchmark-adjusted	-21.62%	-11.07%	-16.96%	-13.28%
Equal-Weighted	Smart Beta	-13.01%	2.60%	2.02%	3.39%
	Declared Benchmark	-2.05%	5.32%	4.43%	3.09%
	Benchmark-adjusted	-11.14%	-2.73%	-2.43%	0.36%
Fundamentals	Smart Beta	-16.51%	3.01%	0.47%	-1.94%
	Declared Benchmark	-4.48%	1.27%	0.97%	-1.39%
	Benchmark-adjusted	-11.85%	1.77%	-0.50%	-0.24%
Growth	Smart Beta	-8.19%	6.27%	5.74%	-0.58%
	Declared Benchmark	0.72%	7.26%	6.78%	-0.88%
	Benchmark-adjusted	-9.05%	-0.99%	-1.04%	0.31%
Low Volatility	Smart Beta	-4.87%	1.79%	1.69%	3.79%
	Declared Benchmark	-1.88%	1.95%	2.86%	4.61%
	Benchmark-adjusted	-3.02%	-0.16%	-1.17%	-0.82%
Value	Smart Beta	-12.42%	1.15%	-0.55%	1.66%
	Declared Benchmark	-1.49%	4.93%	3.95%	0.40%
	Benchmark-adjusted	-11.07%	-3.81%	-4.51%	1.26%
Momentum	Smart Beta	-6.83%	8.09%	8.69%	5.07%
	Declared Benchmark	3.97%	11.21%	11.45%	5.49%
	Benchmark-adjusted	-10.87%	-3.15%	-2.77%	-0.42%
Multi-Factor	Smart Beta	-13.97%	1.78%	1.06%	1.74%
	Declared Benchmark	-5.49%	4.64%	-1.22%	-0.72%
	Benchmark-adjusted	-8.52%	-2.81%	2.37%	2.59%
Non-traditional	Smart Beta	-32.66%	-16.81%	-21.91%	-13.07%
	Declared Benchmark	-9.55%	-13.12%	-13.14%	-7.03%
	Benchmark-adjusted	-23.50%	-3.71%	-8.81%	-6.06%
Quality	Smart Beta	-9.01%	4.93%	5.70%	7.74%
	Declared Benchmark	0.62%	7.95%	8.34%	10.04%
	Benchmark-adjusted	-9.79%	-3.04%	-2.64%	-2.31%

Interestingly, in the equal-weighted portfolio the growth strategy consistently underperforms its benchmark in the short term (one, three and five years) respectively but outperforms since inception. This accounts as well for value and dividend. A potential explanation for this phenomenon is surpassing excess returns after the financial crisis, pushing the performance since inception enormously.

Howbeit, within the seven outperforming categories the highest outperformance is exhibited by the multi-factor portfolio. The outperformance of 3.1% on five years and 4.62% since inception explains why investors have leaned towards multi-factor strategies over single-factor products in recent years. However, short-term performance results indicate that multi-factor strategies fail to deliver similar returns lately. Note that all categories underperform in the short-term horizon of one year due to the bad equity performance in 2018, the worst year for equity returns since the financial crisis in 2008 (Isidore, 2018). Due to the increasing demand within the multi-factor category paired with poor recent stock performances, the bad performance in the short-term horizon of one year (2018) seems reasonable. Similarly, dividend historically represented a highly popular SB strategy that attracted many investors. Attributable to the low interest environment investors were looking for dividend paying equities. However, after fund-inflows peaked SB dividend performance deteriorated compared to its respective benchmark.

Looking at the size-weighted portfolio we can find similar results. Five out of eleven categories are able to outperform their underlying raw benchmarks since inception. Besides the categories mentioned in the equal-weighted portfolio, also fundamentals and low volatility fail to outperform. Again, earnings-weighted accounts for the worst performance since inception, whereas dividend and multi-factor are the best performers.

Generally, both tables indicate that overall around half of SB ETF categories are able to outperform their respective raw benchmark since inception, but nearly all of them fail in doing so in the short term when looking at a one-, three-, or five-year horizon. This is consistent with the evidence provided above as extraordinary excess returns in the years before the financial crisis 2008 and afterwards (2010-2012) cause the long-term performance to be relatively high before declining to negative results lately.

6.2 Results Risk-adjusted Performance

Additionally, to the relative return analysis, Table X and XI provide a risk-adjusted analysis with an overview of essential performance metrics (defined in Table II) with further insights into the return performance between different categories, again for both equally-weighted and size-weighted portfolios. Alpha, the residual return that is not attributable to benchmark coefficient (beta), is the highest for multi-factor and lowest for earnings-weighted strategy when evaluating equally-weighted portfolios. Likewise, growth SB ETFs outperform its benchmark by 2.41% within size-weighted portfolios (exhibiting highest SR and SoR), whereas earnings-weighted and non-traditional strategy perform worst with a statistically significant underperformance of -3.72% and -2.32% respectively. Again, one needs to keep in mind that these categories consist of rather few funds, thus potentially distorting the validity of results.

Table X: Return performance of equally-weighted domestic equity Smart Beta ETFs vs. their benchmarks by category, Jan 2000 - Dec 2018

		Excess Return		SB Return	Alpha	Beta	Sharpe Ratio		Sortino Ratio		IR
Smart Beta category	# of Obs	Smart Beta	Declared Benchmark	Net of declared benchmark	Smart Beta	Smart Beta	Smart Beta	Declared Benchmark	Smart Beta	Declared Benchmark	Smart Beta
Dividend	146	-0.67%	-4.12%	3,48%	1.77%	0.70 (***)	0.10	-0.04	0.11	-0.03	0.61
Earnings-Weighted	82	-13.60%	-0.82%	-12.82%	-3.72% (**)	1.39 (***)	-0.20	0.02	-0.16	0.03	-0.32
Equal-Weighted	177	-2.36%	-2.94%	-0.57%	0.30%	0.84 (***)	0.27	0.28	0.37	0.32	-0.03
Fundamentals	142	0.78%	-0.30%	1.10%	1.29% (***)	1.09 (***)	0.17	0.08	0.22	0.09	0.20
Growth	211	0.25%	-0.99%	1.26%	2.11% (***)	0.93 (***)	0.40	0.19	0.50	0.22	0.39
Low Volatility	79	6.05%	5.38%	0.67%	0.77% (*)	0.89 (***)	0.44	0.40	2.47	2.46	0.12
Value	211	2.08%	0.44%	1.67%	1.86% (***)	0.91 (***)	0.42	0.22	0.53	0.28	0.31
Momentum	125	2.61%	4.69%	-2.08%	0.45%	0.61 (***)	0.27	0.31	0.36	0.38	-0.11
Multi-Factor	156	3.12%	-1.04%	4.62%	3.39% (***)	0.86 (***)	0.37	0.06	0.52	0.07	0.61
Non-traditional	122	-11.32%	-7.31%	-4.03%	-1.70% (***)	0.92 (***)	-0.10	-0.05	-0.13	-0.05	-0.14
Quality	54	3.69%	4.24%	-1.20%	0.27% (*)	0.94 (***)	0.29	0.29	2.10	6.16	0.31

*Significance Codes: *** (<0.01); ** (<0.05); * (<0.1)*

Table XI: Return performance of size-weighted domestic equity Smart Beta ETFs vs. their benchmarks by category, Jan 2000 - Dec 2018

		Excess Return		SB Return	Alpha	Beta	Sharpe Ratio		Sortino Ratio		IR
Smart Beta category	# of Obs	Smart Beta	Declared Benchmark	Net of declared benchmark	Smart Beta	Smart Beta	Smart Beta	Declared Benchmark	Smart Beta	Declared Benchmark	Smart Beta
Dividend	146	-0.01%	-3.49%	3.52%	1.65%	0.65 (***)	-0.02	-0.21	-0.02	-0.17	0.21
Earnings-Weighted	82	-13.85%	-0.61%	-13.28%	-3.72% (**)	1.54 (***)	0.03	0.22	0.04	0.28	-0.28
Equal-Weighted	177	3.39%	3.09%	0.36%	0.90% (**)	0.93 (***)	0.31	0.27	0.35	0.31	0.14
Fundamentals	142	-1.94%	-1.39%	-0.24%	0.53%	1.11 (***)	0.01	0.12	0.02	0.13	0.06
Growth	211	-0.58%	-0.88%	0.31%	2.41% (**)	0.08 (***)	0.89	0.64	10.74	3.01	-0.12
Low Volatility	79	3.79%	4.61%	-0.82%	0.13%	0.81 (***)	0.39	0.30	2.59	0.40	-0.08
Value	211	1.66%	0.40%	1.26%	1.33% (***)	0.89 (***)	0.20	0.45	0.24	0.54	0.20
Momentum	125	5.07%	5.49%	-0.42%	1.25% (*)	0.66 (***)	0.34	-0.11	0.42	-0.14	-0.05
Multi-Factor	156	1.74%	-0.72%	2.59%	2.15% (***)	0.81 (***)	0.28	0.04	0.38	0.06	0.38
Non-traditional	122	-13.07%	-7.03%	-6.06%	-2.32% (**)	0.92 (***)	-0.03	0.45	-0.04	3.04	-0.18
Quality	54	7.74%	10.04%	-2.31%	0.14%	0.92 (***)	0.30	0.13	9.91	0.17	-0.01

*Significance Codes: *** (<0.01); ** (<0.05); * (<0.1)*

In terms of performance measures, low volatility, growth and value account for the highest SRs when assessing equally-weighted portfolios and earnings-weighted for the lowest. Similarly, growth (again) provides clearly the highest SR within size-weighted portfolios, whereas non-traditional and dividend strategies the lowest. In the equal-weighted as well as the size-weighted analysis, growth and multi-factor ETFs have significantly higher SRs relative to their benchmark. Since a higher skill level of a portfolio manager generates higher excess returns relative to the benchmark, growth and multi-factor categories account for the highest information ratios.

Whereas the highest Sortino Ratio can be found for low volatility and quality in the equally-weighted setup, growth and quality account for the highest ratio within size-weighted portfolios. In the equal-weighted portfolio, earnings-weighted and non-traditional provide negative values, implying a higher probability of large losses. Interestingly, dividend has a very low Sortino Ratio in both setups, being even negative in the size-weighted one.

Naturally, all beta coefficients are significant at the 0.01 level for both portfolios and indicate to track their respective benchmark somewhat closely. Nevertheless, alpha coefficients in the equally-weighted portfolio exhibit less coherent results, since only eight categories are characterized by a significant alpha, whereof earnings-weighted ETFs significantly underperform. As expected after the preceding analysis, the growth, value and multi-factor strategies exhibit statistically significant outperformance in both equal- and size-weighted portfolios. In contrast, when assessing the size-weighted portfolio, seven out of eleven categories provide significant alpha coefficients, whereof earnings-weighted significantly underperforms. Again, multi-factor ETFs account for the highest outperformance (2.15%) at a 0.01 significance level.

Interestingly, dividend ETFs outperformance is statistically insignificant in both the equal-weighted and size-weighted analysis resulting in low to negative information ratios, even though they outperform their respective raw benchmark. Furthermore, given the above analysis, value and growth belong to the best performing categories, accounting for the highest SRs and SORs relative to their benchmark due to their significant outperformance.

6.3 Results Multi-Factor Models

The previous return-based analysis does not support SB ETF ambassadors' view that SB ETFs consistently outperform risk-adjusted benchmarks by titling their portfolios towards aforementioned factors. The variation in relative performance seems to be the result of different risk profiles of the underlying factor in which only few SB strategies are suggested as good investments. Therefore, numerous macroeconomic variables are introduced to detect the relative sensitivity towards these explanatory variables. Please find the results of the multiple regression for each SB ETF category's returns below. First, the complete model (Appendix XI) can be compared with the strictly macroeconomic model (Table XII) in case of the equally-weighted portfolio. Likewise, the complete model for the size-weighted portfolio can be found in Appendix XII whereas the macroeconomic model is presented in Table XIII.

Table XII: Multiple regression of equal-weighted domestic equity Smart Beta ETFs vs. 8 macroeconomic factors by category, Jan 2000 – Dec 2018

Smart Beta category	Intercept	GDP	IPI	UR	IR	MS	CPI	OP	FDI	R^2
Dividend	-0.04	19.07 (*)	6.07 (***)	-0.78 (**)	-0.01 (*)	-2.14	-0.01 (***)	-0.00	0.00	0.3502
Earnings-Weighted	-0.05	3.81	1.04	0.46	0.00	0.47	0.00	0.12	0.00	-0.0012
Equal-Weighted	-0.02	18.61 (**)	4.38 (**)	-0.72	-0.01	-1.07	-0.01 (*)	0.07	0.00	0.2498
Fundamentals	-0.03	18.36	7.05 (***)	-0.71 (*)	-0.01	-3.63	-0.02 (***)	0.01	0.00	0.3243
Growth	-0.02	25.89 (***)	5.34 (**)	-1.59 (***)	-0.00	-4.18	-0.02 (***)	0.02	-0.00	0.3088
Low Volatility	0.03 (*)	-0.49	0.33	-0.41 (**)	-0.00	0.66	0.00	-0.00	-0.00	0.0129
Value	-0.03	26.32 (***)	4.94 (**)	-1.14 (***)	-0.00	-3.88	-0.02 (**)	-0.02	0.00	0.3319
Momentum	-0.01	13.32	4.07 (***)	-0.87 (**)	-0.01 (**)	-1.22	-0.01	0.01	-0.00	0.2267
Multi-Factor	-0.00	0.14	5.68 (***)	-1.00 (***)	-0.01	-1.81	-0.02 (***)	0.02	0.00	0.2760
Non-traditional	-0.09	18.28	4.10	-0.52	-0.02	-0.74	-0.01	0.14	0.00	0.1007
Quality	0.02	0.94	-0.10	-0.31 (**)	-0.00	-0.45	-0.00	-0.02	-0.00	0.0100

*Significance Codes: *** (<0.01); ** (<0.05); * (<0.1)*

As indicated above, several factors do not show any significant exposure towards SB ETF excess returns. Interestingly, the established factors value and size are never statistically significant in the complete model with the exception of the growth and quality strategies, suggesting a mild significance at best. Consequently, the macroeconomic model without these two factors results in slightly better models (higher R squared) for the majority of SB categories.

Table XIII: Multiple regression of size-weighted domestic equity Smart Beta ETFs vs. 8 macroeconomic factors by category, Jan 2000 – Dec 2018

Smart Beta category	Intercept	GDP	IPI	UR	IR	MS	CPI	OP	FDI	R^2
Dividend	-0.04	17.18 (*)	5.56 (***)	-0.84 (**)	-0.01 (*)	-1.89	-0.01 (***)	-0.03	0.00	0.3561
Earnings-Weighted	-0.05	3.58	1.10	0.47	0.00	1.01	0.01	0.16	0.00	0.0038
Equal-Weighted	-0.02	18.08 (*)	4.98 (***)	-0.97 (***)	-0.01	-1.24	-0.01 (***)	0.00	0.00	0.3072
Fundamentals	-0.04	22.46 (**)	6.82 (***)	-0.30	-0.01	-6.41	-0.03 (***)	-0.02	0.00	0.3213
Growth	-0.03	22.86 (***)	4.26 (**)	-1.58 (***)	-0.00	-3.69	-0.02 (***)	0.00	-0.00	0.2966
Low Volatility	0.02	-0.60	0.21	-0.24	-0.00	0.58	0.00	-0.01	0.00	-0.0017
Value	-0.03	24.24 (***)	3.64 (**)	-1.04 (***)	-0.01	-2.95	-0.01	-0.04	0.00	0.3193
Momentum	0.01	11.82	4.03 (***)	-1.09 (***)	-0.01	-1.01	-0.01 (***)	0.02	-0.00	0.2487
Multi-Factor	-0.01	12.72	4.56 (***)	-0.85 (***)	-0.01	-1.26	-0.01 (**)	0.02	0.00	0.2601
Non-traditional	-0.09	19.99	3.99	-0.51	-0.02	-1.30	-0.02	0.15	0.00	0.1093
Quality	0.02	0.96	-0.07	-0.29 (**)	-0.00	-0.47	-0.00	-0.02	-0.00	0.0046

*Significance Codes: *** (<0.01); ** (<0.05); * (<0.1)*

In case of the size-valued portfolio, neither SMB nor HML factor exhibits statistically significant coefficients in any of the eleven SB models. Hence, the model fit (R^2) increases similarly for many size-weighted SB categories. As a result of the unexpected poor performance of both SMB and HML factors (both small coefficient and t-statistic), the complete model is neglected throughout the remainder of the discussion. Hence, the argument that size and value factor drive many SB ETFs (excess) returns is refuted. Instead a stepwise regression approach based on the macroeconomic model attempts to find the best model for each category by removing statistically insignificant explanatory variables. Since some criticism is found in academia with respect to stepwise regression such as (1) incorrect use of degrees of freedoms in computer packages, (2) possibly not identifying the best variable in a data set, (3) unusual small standard errors or (4) non-replicable results (Thompson, 1995) the results should always be judged complementary to the original model. Appendix XIII present the findings of the stepwise regression for the equally- and size- weighted portfolio respectively. The coefficients and respective p-values seem to be closely related to the full

model, so that above mentioned concerns seem to be partly mitigated. Nonetheless, the macroeconomic model serves as a benchmark throughout the results and discussion section.
Interestingly, four macroeconomic variables indicate statistically significant sensitivities among many SB strategies including GDP, IPI, UR and CPI. Notably, GDP shows the highest significant coefficients among the various models ranging from 18.61 to 26.32 followed by IPI with a coefficient value ranging between 4.07 and 7.05. The GDP coefficient peaks at 26.32 for the value strategy (equally-weighted portfolio) and implies by far the highest sensitivity towards a macroeconomic effect among all models. In terms of interpreting the output, the model suggests for example that for each percentage increase in GDP, SB value excess returns increase by 26.32%, assuming all other parameter stay constant. In contrast, FDI, OP and MS are never exhibiting statistically significance among any SB categories in either the equally-or size-weighted portfolio which suggests the low influence towards SB ETF returns. This is further supported by the rather low coefficients of these explanatory variables, in particular in case of FDI and OP. Lastly, it seems relevant to point out that IR is only statistically significant for dividend (equally- and size-weighted portfolio) and momentum strategies (equally-weighted portfolio). This finding is supported by the interest alike features of dividend-paying equities which caused a large inflow in dividend SB ETFs during the recent low-interest rate environment. Hence, interest rates and dividend SB ETFs are related in a way that a decrease in interest rates negatively affects SB ETF excess returns.

In terms of comparing the equally- and size-weighted portfolio models, interestingly, SB excess returns exhibit larger exposure towards (significant) GDP and IPI coefficients in case of an equal weighting style. This is true for the dividend, equal-weighted, growth and value SB strategies and indicates a particularly high sensitivity of “small” (low AUM) SB ETF’s towards GDP and IPI changes compared to “large” (high AUM) funds. In contrast, UR has a higher (significant) coefficient when SB ETFs are weighted equally for growth, value, multi-factor and quality strategy whereas the coefficient is smaller when SB ETFs are weighted according to size for the dividend and momentum strategy; thus the sensitivity towards unemployment rate changes is less/more severe for small funds depending on the SB ETF strategy. CPI shows only minor deviations between equal and size weighted results which can be attributed mainly to the generally low sensitivity of SB ETF excess returns towards changes in consumer prices.

Among the best fitting models are dividend, fundamentals and value models for both the equal and size-weighted portfolio. The respective R squared is strictly higher than 30%, which is not the case for most of the remaining SB factor models. The worst performing models are clearly earnings-weighted, low volatility, non-traditional and quality in which R squared even reaches values below zero. Logically, this implies the model is explaining SB excess returns worse than a horizontal line and none of the eight macroeconomic variables seem to describe the SB ETF observations well. Note, however, that the substantially lower amount of observations compared to the dividend, fundamentals and value strategies poses a key problem. The combination of few funds and few observations leads to the conclusion that these SB ETF strategies cannot be assessed from a macroeconomic perspective and hence, the analysis focuses on the better-fitting models.

Table XIV: Significant coefficient direction of equal-weighted domestic equity Smart Beta ETFs vs. 8 macroeconomic factors by category, Jan 2000 - Dec 2018

Smart Beta category	GDP	IPI	UR	IR	MS	CPI	OP	FDI
Dividend	+	+	-	-		-		
Earnings-Weighted								
Equal-Weighted	+	+				-		
Fundamentals		+	-			-		
Growth	+	+	-			-		
Low Volatility			-					
Value	+	+	-			-		
Momentum		+	-	-				
Multi-Factor		+	-			-		
Non-traditional								
Quality			-					

Furthermore, remember the expected direction of the coefficients in Table V whereas the actual sign of the relationship between SB excess returns and macroeconomic variables can be found in Table XIV and XV for both the equally- and size-weighted portfolio. In fact, the economic conditions indicators GDP and IPI are often times highly statistical and have a positive relation with SB ETF returns. The low volatility marks the only exception since the sign of the coefficient is indeed negative (but statistically insignificant). This seems somewhat plausible given the desire to invest in assets with low (systematic) risk exposure which often causes low volatility assets to exhibit countercyclical return characteristics. This might be seen as an attractive hedge against recessions as low volatility SB ETFs imply a slight outperformance compared to its underlying benchmark. Surprisingly, the results display a negative relationship between the UR coefficient with all SB ETF excess returns except the statistically insignificant coefficient of the earnings-weighted strategy. The sign was expected to depend on either bullish (positive relation) or bearish markets (negative relation). However, it is impossible to generalise the period 2000-2018 to be either of bearish or bullish nature. The issue will be explained more in-depth in section 7 (Discussion). Apart from this, CPI suggests - as expected - a negative relationship among SB ETF returns except statistically insignificant coefficients of the low volatility and earnings-weighted model. Furthermore, OP and FDI show varying relationships which was expected for oil prices since its sensitivity depends on specific industry holdings whereas it was not hypothesised for FDI. Changes in MS are found to be mostly negatively related to SB ETF excess returns which acts contrary to the expectation of having a positive relationship. Note, however, that coefficients of OP and FDI are close to zero and strictly statistically insignificant, thus lowering empirical validity of its influence on SB excess returns. Likewise, MS is hardly indicating significant results, however, the coefficient is relatively large so that MS seems to somewhat affect SB ETF excess returns.

Table XV: Significant coefficient direction of size-weighted domestic equity Smart Beta ETFs vs. 8 macroeconomic factors by category, Jan 2000 - Dec 2018

Smart Beta category	GDP	IPI	UR	IR	MS	CPI	OP	FDI
Dividend	+	+	-	-		-		
Earnings-Weighted								
Equal-Weighted	+	+	-			-		
Fundamentals	+	+				-		
Growth	+	+	-			-		
Low Volatility								
Value	+	+	-					
Momentum		+	-			-		
Multi-Factor		+	-			-		
Non-traditional								
Quality			-					

On a last note it is important to point out that common significance levels are used. As outlined earlier, this does pose certain problems in economic research since appropriate and verifiable factors likely need to exceed t-statistics beyond three in order to be regarded robust factors. However, raising the required threshold to t>3 results in only five significant coefficients in the equally-weighted portfolio (CPI coefficient for Fundamental, Growth, Multi-Factor; UR coefficient for Growth and Value) and seven significant coefficients in the size-weighted portfolio (CPI coefficient for Fundamental, Growth, Momentum; UR coefficient for Growth, Momentum and Value; GDP coefficient for Value).

6.4 Principal Component Analysis

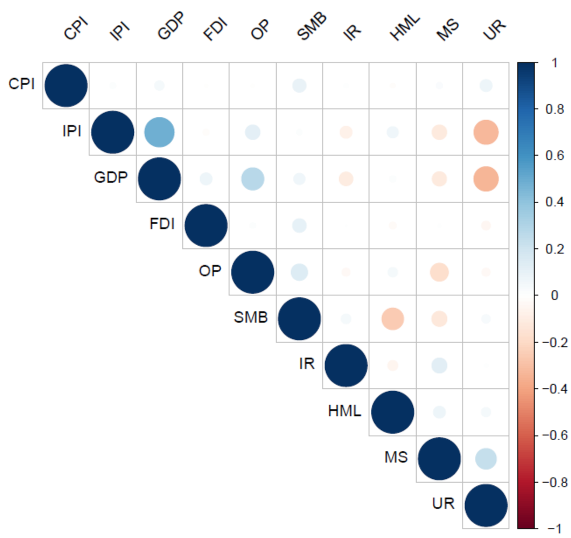
With respect to the PCA, all eight macroeconomic variables are included. With the aim to reduce dimensionality of multivariate data PCA assumes that a small number of representative variables is enough to collectively explain most of the variation present in the original variables. As can be seen in Table XVI the data set is transformed to principal components (PCs) that are uncorrelated and ordered in a way the first component captures most of the variation.

Table XVI: Individual standard deviation and cumulative portion of variance among macroeconomic variables, Jan 2000 - Dec 2018

	PC1	*PC2*	*PC3*	*PC4*	*PC5*	*PC6*	*PC7*	*PC8*
Standard Deviation	1.4063	1.1274	1.0280	0.9827	0.9797	0.84697	0.76204	0.68613
Proportion of Variance	0.2472	0.1589	0.1321	0.1207	0.1200	0.08967	0.07259	0.05885
Cumulative Portion	0.2472	0.4061	0.5382	0.6589	0.7789	0.86857	0.94115	1.00000

Find a correlation plot below in Figure II. The correlation between the independent variable are in the range of -0.33 to 0.48. PCA is often used to account for multicollinearity. Note, though, that variance-inflation factors are all within reasonable ranges (<2) in this case. Based on the overall Kaiser-Meyer-Olkin measure of sampling adequacy of 0.59 the results are accepted given that Kaiser (1974) recommends accepting values greater than 0.5. Additionally, Bartlett's test of sphericity is highly significant (<0.001) so that factor analysis seems appropriate.

Figure II: Correlation plot for independent variables in multiple regression



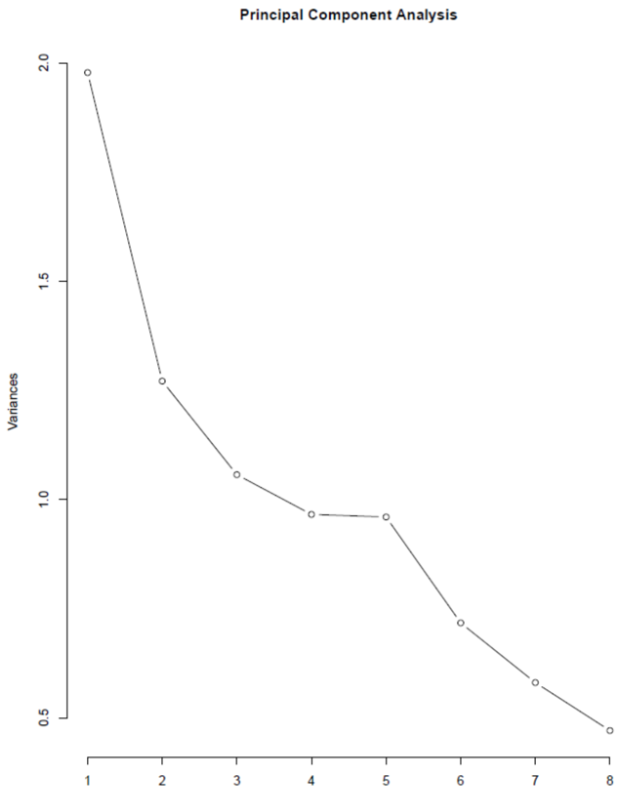
Following, orthogonal transformation yields a total of eight PCs in which the first PC explains almost a quarter of the variation in the data set (24.72%). Combining the first three PCs results in a cumulative portion of 53.82% of total variation. Standard deviation and loading vectors are retrieved, the latter imply the degree to which each original variable contributes to the PCs. Exemplary, PC1 depends largely on the variables GDP, UR and IPI, whereas PC2 is mostly driven by FDI. Lastly, PC3 depends mainly on IPI, IR and MS (find a detailed rotated component matrix in Table XVII).

Table XVII: Rotated Component Matrix based on eight macroeconomic variables, Jan 2000 – Dec 2018

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
IPI	0.5192	-0.0227	0.3174	-0.2259	0.0537	-0.0909	-0.5630	-0.4999
MS	-0.3072	0.0614	0.5432	–0.3551	0.4592	0.3767	0.3075	-0.1818
IR	-0.1593	-0.0595	0.5946	0.7289	0.0787	-0.2268	-0.1404	0.0932
OP	0.2933	-0.1397	-0.3891	0.4126	0.6449	0.0952	0.2182	-0.3163
UR	-0.4542	-0.0404	-0.2402	-0.1595	0.4997	-0.1635	-0.6167	0.2278
CPI	-0.0449	-0.7021	0.0783	-0.2568	0.0425	-0.5888	0.2844	-0.0573
FDI	0.1263	0.6782	0.0366	-0.1251	0.2232	-0.6289	0.2443	0.0435
GDP	0.5479	-0.1339	0.1839	-0.1184	0.2504	0.1416	0.0038	0.7423

Naturally, the maximum number of explanatory variables can never exceed the number of explanatory variables for which the cumulative variance amounts to 100%. The first three PCs are chosen as a condensed version of the original data set since they explain more than 50% of the variance. Note that the captured variance of each PC is shown and plotted in Figure III based on which three PCs were identified as the most suitable amount. Extending the number of PCs beyond three increases the cumulative variance only linearly. This finding is supported by the low to negative correlation among explanatory variables as supposed by the opposing direction of the respective variables in the bi plot in Appendix XIV.

Figure III: "Elbow Plot" for Principal Component Analysis: Cumulative explained variance



Consequently, the question arises to what extent these three PCs perform in a factor analysis. A so-called principal component regression is conducted for both the equally- and size-weighted portfolio in the proceeding step. Find the regression results in Table XVIII and XIX. The results are in line with the previous findings. Again, earnings-weighted, low volatility and quality models seem to have a bad fit with R squared exhibiting negative values. In contrast, the remaining models indicate that the PCs are quite accurate in explaining SB excess returns.

Table XVIII: Principal Component Regression: Regressing first three PC's on equal-weighted SB ETF monthly excess returns

Smart Beta category	Intercept	PC1	PC2	PC3	R^2
Dividend	0.01	0.06 (***)	0.01	0.02 (**)	0.3068
Earnings-Weighted	-0.03 (***)	0.01	-0.01	0.00	-0.0039
Equal-Weighted	0.03	0.05	0.01	0.02	0.2080

	(***)	(***)			
Fundamentals	0.02 (***)	0.06 (***)	0.01 (*)	0.01	0.2788
Growth	0.04 (***)	0.07 (***)	0.01	0.02 (*)	0.2582
Low Volatility	0.03 (***)	0.01 (**)	-0.00	0.00	0.0111
Value	0.04 (***)	0.07 (***)	0.01	0.02 (**)	0.2862
Momentum	0.03 (***)	0.04 (***)	0.00	0.02 (**)	0.1949
Multi-Factor	0.04 (***)	0.05 (***)	0.01	0.02 (**)	0.2437
Non-traditional	-0.03 (**)	0.04 (***)	0.01	0.02	0.0809
Quality	0.02 (***)	0.00	-0.00	0.00	-0.0016

*Significance Codes: *** (<0.01); ** (<0.05); * (<0.1)*

Naturally, PC1 is statistically significant for each of the categories excluding the above-mentioned categories earnings-weighted, low volatility and quality. This is in line with previous findings because PC1 is largely driven by FDI, UR and IPI which were mostly statistically significant in the previous multiple regression. Based on the fact that PC2 is affected mainly by FDI and CPI, it comes as no surprise that PC2 is never statistically significant except a mildly significant coefficient for the fundamental model. Note that FDI did not exhibit any statistical significance in the multiple regressions before thereby supporting the finding. Lastly, PC3 which is driven by IPI, IR and MS is statistically significant in five and six models with respect to the equally-weighted and size-weighted portfolio respectively (equally-weighted portfolio: dividend, growth, value, momentum, multi-factor model; size-weighted portfolio: dividend, equal-weighted, growth, value, momentum, multi-factor model). This, again, confirms the importance of IPI and MS as explanatory variables with respect to SB ETF excess returns.

Table XIX: Principal Component Regression: Regressing first three PC's on size-weighted SB ETF monthly excess returns

Smart Beta category	Intercept	PC1	PC2	PC3	R^2
Dividend	0.01 (*)	0.05 (***)	0.01	0.02 (**)	0.3054
Earnings-Weighted	-0.03 (***)	0.01	-0.01	0.00	-0.0033
Equal-Weighted	0.04 (***)	0.01 (***)	0.01	0.02 (***)	0.2759
Fundamentals	0.01	0.06 (***)	0.01 (*)	0.01	0.2580
Growth	0.03 (***)	0.06 (***)	0.00	0.01 (*)	0.2460
Low Volatility	0.02 (***)	0.00	-0.00	0.00	-0.0016
Value	0.03 (***)	0.06 (***)	0.01	0.02 (**)	0.2736
Momentum	0.05 (***)	0.04 (***)	0.01	0.01 (*)	0.2111
Multi-Factor	0.03 (***)	0.04 (***)	0.01	0.02 (**)	0.2293
Non-traditional	-0.03 (**)	0.05 (***)	0.01	0.02	0.0846
Quality	0.02 (***)	0.00	-0.00	0.00	-0.0011

*Significance Codes: *** (<0.01); ** (<0.05); * (<0.1)*

With regard to the principal component regression results, the interpretation of the coefficients is less intuitive because each PC is the result of orthogonal transformation based on the original explanatory variables. No further transformation is conducted on the PCs besides judging the coefficients based on the above rotation matrix.

Lastly, an overview of all aforementioned models and the respective R squared is depicted in Table XX in order to compare the various models presented above.

Table XX: Model comparison of equal- and size-weighted domestic equity Smart Beta ETFs by category, Jan 2000 – Dec 2018

	Equal-weighted				Size-weighted			
Smart Beta category	Complete model	Macroeconomic model	Stepwise Regression model	PCR model	Complete model	Macroeconomic model	Stepwise Regression model	PCR model
Dividend	0.3470	0.3502	0.3538	0.3068	0.3547	0.3561	0.3599	0.3054
Earnings-Weighted	-0.0002	-0.0012	N/A	-0.0039	0.0105	0.0038	N/A	-0.0033
Equal-Weighted	0.2481	0.2498	0.2524	0.2080	0.3070	0.3072	0.3148	0.2759
Fundamentals	0.3216	0.3243	0.3244	0.2788	0.3183	0.3213	0.3285	0.2580
Growth	0.3157	0.3088	0.3115	0.2582	0.3062	0.2966	0.3052	0.2460
Low Volatility	0.0090	0.0129	0.0385	0.0111	-0.0087	-0.0017	0.0223	-0.0016
Value	0.3289	0.3319	0.3314	0.2862	0.3159	0.3193	0.3195	0.2736
Momentum	0.2348	0.2267	0.2007	0.1949	0.2582	0.2487	0.2266	0.2111
Multi-Factor	0.2768	0.2760	0.2493	0.2437	0.2594	0.2601	0.2270	0.2293
Non-traditional	0.0003	0.1007	0.0866	0.0809	0.1051	0.1093	0.0925	0.0846
Quality	0.0273	0.0100	0.0307	-0.0016	0.0224	0.0046	0.0258	-0.0011

7 Discussion

The mixture of passive and active management is exploiting a label to justify higher fees of SB ETFs. However, Smart Beta has become one of the fastest-growing segments of the asset management industry, as poor returns drive investors out of traditional, actively managed funds. This leads to higher competition, since more asset managers scramble to take advantage of the growth by launching their own Smart Beta ETFs. As mentioned before, as many as 97 ETFs emerged in 2018 alone, substantiating the rise in popularity of these funds. Ram (2017) concludes that some are concerned “that a price war could be looming”. According to Ben Johnson (2019), director of global ETF research at Morningstar, “the easiest way to differentiate one’s fund in this increasingly crowded landscape is to offer it at a lower price relative to incumbents.” Just recently, BlackRock Inc., the world’s largest asset manager, was cutting expenses on six SB ETFs as competition heats up making it difficult to win investors in a growing market (Willmer, 2016). As this trend might be followed by further participants, cutting fees is inevitable when the largest ETF providers want to dominate the passive investment market and new entrants want to enter the market. A similar trend is found in this analysis which likewise seems to have a strong influence on recent SB ETF (performance) developments. The next paragraph focuses on two particularly interesting SB strategies, namely dividend and multi-factor before summarizing the general development of all SB strategies. This serves as the basis of investor recommendations.

First, dividend SB ETF excess returns are examined due to their past popularity among investors. With long-term interest rates being so low, stocks may be the best way for investors to grow their income. Dividend ETFs forming part of the “Smart Beta” group try to enhance overall income with the use of screens designed to predict healthy pay-outs to shareholders in selected stocks. Thus, income seekers search for alternative to bonds, when investing in dividend-weighted Smart Beta funds. Dividend strategies were expected to experience erosion of its excess returns in recent years due to low interest environment since people are much likelier to invest in fixed income alike equities. Our findings support the sensitivity towards interest rate changes (only significant IR coefficient among all models). Contrary to last years, the relatively poor (recent) performance is likely decreasing future fund inflows. However, since interest rates seem to pick up again, the development of the SB dividend strategy seems particularly exciting and unpredictable. The earlier mentioned tax disadvantage of dividends will play another role in the development to eventually lower the attractiveness of a dividend SB strategy further.

After discussing the expansion of the CAPM by traditional as well as alternative factors in detail, the trend towards multi-factor strategy is not surprising. Investors increasingly gravitate to multi-factor approaches for the benefits of diversification and the potential for a "smoother ride" as well as the potential for lower tracking error compared to single factor strategies. It was expected that multi-factor strategy performs best as a result combined with increasing demand for multi-factor SB ETFs. The analysis showed that this strategy has the highest SR relative to its underlying benchmark (0.37 vs 0.06 and 0.28 vs 0.04 for equal- and size-weighted portfolio respectively). It remains questionable whether the high demand paired with the high fund-inflows remains a profitable long-term investment strategy as short-term performance seems to be deteriorating.

On a more general level, based on the development in the SB ETF market dividend and multi-factor strategies seem to remain the most popular SB strategies besides the established value and growth strategies, all of which experience statistically significant outperformance relative to their benchmark. Whereas the relative share of growth SB ETFs in the sample strongly declines since inception, both dividend and value reached their highest share during the financial crisis 2008 after which they declined similarly. In contrast, equal-weighted, fundamentals, low volatility and multi-factor strategies seem to be on the rise. The analysis suggests that all of these strategies outperform its benchmark, however only multi-factor exhibits statistically significant outperformance in both the equal- and size-weighted portfolio. Strategies with a smaller track record and fund variety such as earnings-weighted, low volatility, non-traditional and quality categories cannot be perfectly assessed yet. Based on the current analysis, earnings-weighted and non-traditional SB ETFs suggest the poorest performance and as a result are not recommendable investments. The analysis outlined several issues connected with a selection bias and the limited amount of observations though. Future research will shed more light on this question as these categories also grow in popularity and market size.

Before debating macroeconomic sensitivities with respect to the SB ETF strategies, however, an important discovery is that well established factors such as SMB and HML (Fama & French, 1993) are surprisingly insignificant throughout this study. Besides a mild significance of the

HML coefficient (growth and quality strategy for equally-weighted portfolio), the results suggest that other factors are in fact more relevant to explain SB ETF excess returns. Consequently, both SMB and HML are not considered as explanatory variables and dropped from the analysis. Interestingly, not even value, growth or multi-factor SB strategies exhibit high loadings on either HML or SMB as suggested earlier.

After assessing SB ETF exposure towards macroeconomic variables, the analysis suggests that out of eight macro variables only four (GDP, IPI, UR, CPI) are statistically significant in explaining SB ETF excess returns whereas the remaining four (IR, MS, OP, FDI) are not. Frankly, the macroeconomic model is only able to explain a fraction of the SB ETF excess returns for well-established SB ETF strategies such as value, growth, equal-weighted, momentum and multi-factor strategy. In these superior models, the model fit (R squared) ranges around 30%. In contrast, earnings-weighted, low volatility, non-traditional and quality are rather new SB strategies that are plagued by a short track record next to a very limited fund range. The induced selection bias seems to hinder the model construction. With sufficient data, the interpretation of the study might be extended to all SB ETF strategies. Meanwhile, macroeconomic sensitivity can be assessed in case of the more prevalent strategies.

Keeping the factor zoo in mind, instead of having a universe of theoretically infinite factors, we expected that our chosen macroeconomic factors can be reduced to maximum 3 factors while still providing most of the variance within the data set. Based on a fairly uncorrelated data set with respect to the explanatory variables it was difficult to determine the appropriate number of principal components. Nonetheless, three PCs are chosen to represent the macroeconomic variables. As hypothesised, the model fit decreases only slightly from the macroeconomic model to the PCR model. For example, the R squared of the dividend SB strategy declines from 35.02% (35.61%) to only 30.68% (30.54%) for the equally- and size-weighted portfolio respectively. Thus, reducing the dimensionality from eight to only three independent variables results in a model that performs relatively well. More importantly, significant PC coefficients are driven by GDP, UR and IPI in case of PC1 and IPI, IR and MS (PC3) whereas PC2 mildly significant at best due to the contribution of FDI, an explanatory variable that has been found to not significantly explain SB ETF excess returns earlier. Thus, neither FDI, OP nor IR seem to have a strong sensitivity towards various SB strategies. MS seems to be somewhat explaining excess returns besides not being a very significant coefficient. These macroeconomic variables might simply not be adequate indicators, alternatively, they might only affect SB strategy returns several months after changes in fiscal policy come into force. In contrast, changes in GDP, IPI are strongly positively related to SB ETF performance shortly afterwards (within one month). Similarly, to a lesser extent, changes in UR and CPI negatively affect most SB ETF excess returns. On a last note, the analysis did not allow to test for the direction of the UR coefficient depending on bullish or bearish periods since the data sample was not split into subsets due to feasibility constraints. All other dependencies between explanatory variables and SB ETF excess returns proved to be exactly what was hypothesised with the exception of MS. Instead of positively affecting returns, various SB strategies are negatively influenced after fiscal policy changes (in money supply) come into effect. This might be due to an imperfect lag of the MS variable in which SB ETF excess returns are positively affected only several months afterwards. Note that the significance levels were rather low though.

Another finding can be derived from the analysis when comparing the equal- vs. the size-weighted portfolio in regard to the macroeconomic model. Interestingly, all significant GDP and IPI coefficients are larger for the equal-weighted portfolio when compared with the same SB strategy. Thus, small SB ETF funds (measured in AUM) in many categories such as dividend, growth or value have a much higher sensitivity towards GDP and IPI changes that drive the coefficients upwards when each SB ETF irrespective of size receives equal weight. Similarly, return outperformance is higher for many of these SB ETF strategies. In the specific example of dividend, value and multi-factor strategy, for example, it seems attractive to invest into in smaller SB ETFs which are usually considered riskier due to the limited historical information. As one can see in Table X and XI though, the statistically significant outperformance increases by placing a higher weight on small funds. Surprisingly, SR and SoR is larger for the equal-weighted portfolio compared to the size-weighted portfolio in case of value and multi-factor strategy for example. Consequently, the risk-return trade-off seems to be superior when taking small SB ETFs into consideration since the higher SR signals that for the same level of risk, more return can be achieved. Caution is required, however, since the opposite effect is true for the growth strategy in which small funds deteriorate the risk-return trade-off so that investors cannot simply be advised to focus on small SB ETFs. Nonetheless, considering that multiple strategies experience this effect, this seemingly counterintuitive finding opens up new and important discussions.

Lastly, the analysis assessed a total of four models for both the equal- and size-weighted portfolio. The complete model can quickly be reduced to the macroeconomic model without losing explanatory power of the model. The model fits the data only better in case of stepwise regression. This does not come as a surprise since certain issues were mentioned with respect to stepwise regression. Consequently, these models should only be regarded complementary to the macroeconomic model. Similar coefficients and t-statistics mitigate many of the aforementioned concerns though. On a last note, the reduced principal component regression performed quite well compared to the macroeconomic model with eight independent variables. However, the paper mentions several times how research has often argued in favour of higher t-statistics in order to declare robust factors. Most of the significant factors in this study fail to be significant at a t-statistic threshold exceeding a value of three. Only a fraction holds under the stricter requirements with only five and seven significant coefficients in the equal and size-weighted portfolio respectively (remember that primarily CPI and UR coefficients exhibited a t-statistic >3). Summing up, SB ETF excess returns do in fact react to certain macroeconomic variables. Following the "factor zoo" criticism, these factors are not necessarily helpful in identifying the next successful investment strategy though. Instead, they are designed to understand the relationship between macroeconomic events and SB ETF returns. While it might be able to develop or extend a SB strategy based on GDP or IPI changes for example, this is not the purpose of the study. Rather, the exponential growth of equity factors is evaluated critically and supported by the findings that there might be an infinite number of factors if one digs deep enough. The proliferation of SB factors does not seem to be justified when looking at an investor's realised excess return. Only value and multi-factor strategies were able to outperform its benchmark at the 1% significance level (using conventional significance levels) in both the equal and size-weighted portfolio. Most factors do not indicate a good return on investment compared to traditional ETFs so that most equity factors are not recommendable from an investor's perspective. However, the relatively poor performance of many of these "smart" products does not hinder the current growth of SB ETFs. The desire to find the next shining "factor" leads many to invest in SB ETFs despite its controversy, a circle in which perhaps it is the fund managers who are profiting the most. However, decreasing fees will are lower the attractiveness and profitability of SB ETF managers which likely has a significant impact on the development of SB ETFs. In the short-term, at least, growth of SB ETFs is expected to persist.

8 Limitations

Throughout the study, several shortcomings need to be addressed. First of all, the sample of 327 SB ETFs does not represent the full SB ETF market since a majority of funds had to be removed. Furthermore, the individual classification of SB ETFs into factor strategies seems subjective at times when being cross-referenced with its prospectus. However, this poses an unavoidable problem due to the inconsistent factor construction found in recent literature. Logically, this makes it difficult to compare various SB strategies uniformly as an overlap among strategies is likely. To make matters worse, each SB ETF is assigned a tradable benchmark ETF based on its underlying index. This, again, involves a certain subjectivity and might lead to different results based on each researcher's choice. Next, measurement and reporting discrepancies are observed for the macroeconomic variables in which certain metrics are published mid-monthly versus end of a month. A total of eight macroeconomic factors are used due to their relativity in academia, however, these eight variables should be cross validated in future research in particular in the context of the uniformly applied lag length of one month. Although a uniform lag length of one was empirically verified through VAR the lag choice can be conducted more in depth. The model proposed by Sims (1980) is to build VAR models with no a priori exclusion restrictions. One has to be aware of the fact that this methodology however quickly exhausts the available degrees of freedom, because each variable has to appear in each equation with the same lag specification (Kunst & Neusser, 1986). Furthermore, the Cholesky ordering is highly arbitrary, since different orderings will produce different A matrices, which in turn will produce different impulse responses. The ordering of the variables cannot be determined with statistical methods but has to be specified by the analyst. Hence, varying the ordering of our variables results in significantly different results (Harris, 1997), which is why we do not build trustworthy assumptions on them. Furthermore, a major limitation of the system is its potential incompleteness. Although in real economic systems complex influences depend on each other, we work with a rather low-dimensional VAR system. All effects of omitted variables are assumed to be reflected in our eight chosen macroeconomic variables. Hence, if important variables are lacking in the system, this may lead to major distortions in the impulse responses and makes them worthless for structural interpretations.

With regard to the model construction, certain SB strategies proved to be poorly explained by the independent variables. This seems to be particular serious for US SB ETFs with low data availability. Hence, the macroeconomic analysis can neither be applied to all SB strategies in the case of the US market nor can it be extended to other markets outside of the United States at present. A cross-country analysis was regrettably infeasible at this point in time. In summary, the rather low model fit of certain SB strategies paired with above shortcomings implies that the results of this analysis should be interpreted more indicative with respect to macroeconomic changes. More research is necessary to validate the exact relationship among each SB strategy.

9 Future Research

As mentioned before, the research in the context of SB ETFs is fairly young so that several recommendations for future research apply. Besides cross-validating the existing findings in SB ETF papers, in particular the macroeconomic model needs to be redefined. For example, it would be desirable to split the sample period into subsets based on bearish or bullish periods. Certain macroeconomic variables such as UR tend to have very different effects on equity returns based on the business cycle of an economy. Additionally, it would be interesting to extend the study towards other markets once data availability of new SB ETF market(s) allows. Country-specific differences and a differentiation between global and local factors likely yields more robust and more comparable results because microeconomic frictions might have an unanticipated impact on the analysis. Next, cross-validating macroeconomic variables in terms of lag length is strongly recommended through a more detailed VAR analysis for example. Furthermore, by splitting the macroeconomic variables into their unexpected versus expected components rather than relying on realised macroeconomic changes, it is likelier to detect unprecedented relations. This approach is often found in reputable macroeconomic studies such as Brandt & Wang (2003) or Schwert (1981). Lastly, individual SB ETF holdings can and should be investigated to derive more specific conclusions about other macroeconomic relations (e.g. oil price sensitivity highly depends on industry) which in turn might yield information about how fund holdings relate to fund performance based on macroeconomic sensitivities.

10 Conclusion

Last decade has been characterized by a steep increase in demand for SB products, thereby opening up augmented research areas based on controversial views. Regarding the fact, that various authors have already analysed the performance of SB products, this paper aims to fill the void in academic literature when extending the analysis of SB ETFs performance with macroeconomic factors. It does so by answering the question if macroeconomic factors are able to explain the performance of SB ETFs in excess of their benchmarks. After providing an initial overview of the concept of "Smart Beta", an empirical research first assesses the US ETF market in order to assess risk-adjusted performance since inception and provide guidance as to what extent investors are better off replacing traditional ETFs with its SB counterparts. Capturing the concept of SB into a comprehensive definition proves to be a challenging process. Most of all, the class of SB ETFs differentiates itself from the omnipresent cap-weighting method, by creating ETFs that tilt towards one or multiple factors. This concept reminds of the widely recognized factor investing. As a matter of fact, SB ETFs tried to hit a niche market while floating between active and passive management strategies. This leads to the alleged advantage of lower management fees compared to active management at the first glance, but critics see it as a disadvantage since these offset the slight outperformance of SB ETFs (Malkiel, 2005).

Analysing a dataset of 327 US SB ETFs our results lead to the main conclusion that SB ETFs as an asset class should be treated with caution by investors based on the divergent fund performances. When taking risk into account, only two categories (multi- factor and value) are able to outperform their benchmark at a significance level of 0.01 for both equal-weighted and size-weighted portfolios. From a pure risk-perspective growth SB ETFs suggest the highest Sharpe and Sortino ratios, however, multi-factor and value SB strategies have a substantially higher risk-return trade-off compared to their "non-smart" benchmark. These SB strategies seem to perform best relatively to their benchmarks and seem to be a worthwhile investment for risk-averse investors. The increasing popularity of multi-factor SB ETFs can be explained by the empirical results; however, the outperformance seems to be reversing lately. Lastly, the equal-weighted portfolios exhibit more extreme and significant outperformance in contrast to the size-weighted portfolio for certain categories. Thus, the results suggest the attractiveness of smaller and more recently launched SB ETFs to deliver positive excess returns, however, higher management fees for smaller funds need to be considered.

Based on the analysis controversial opinions related to this research area find their right to exist and one should question the concept of Smart Beta. However, one needs to keep in mind, that competition within the Smart Beta ETFs seems to be increasing and consequently leads to a decline in management fees. Interestingly, excess returns seem to be decreasing as management fees decline. From that perspective, a slight decline in SB ETF outperformance is not necessarily bad for investors as long as the fees do not exceed the fund's alpha. Additionally, the analysis shows that particularly 2018 was the worst for stocks in 10 years and logically, return performances regarding SB ETFs took a large hit. Nonetheless, many SB categories fail to deliver returns in excess of traditional ETFs. In summary, the outlook for SB outperformance implies a negative trend as the excess returns of many SB strategies are retrogressive in recent years.

Furthermore, it is shown that both size and value factor were surprisingly insignificant in explaining SB ETF excess returns. In contrast to the hypothesis that many SB strategies are highly driven by inherent value or size tilts, it is shown that the reduced model consisting of only macroeconomic variables is superior to the complete model that includes SMB and HML variables.

To answer the research question "*Is the vast amount of equity return factors economically justifiable with respect to Smart Beta investing from an investor perspective and to what extent do macroeconomic forces explain Smart Beta performance?"* it can be concluded that only few equity factors do in fact provide investors with a statistically significant outperformance. To make matters worse, recent performance seems to be deteriorating. As a result, from an investor's perspective the variety of factor strategies do not deliver consistent outperformance. The factor zoo with respect to Smart Beta investing is not economically justifiable; it rather adds fuel to the ongoing debate surrounding the idea of "SB" where authors warn investors about the so-called "smart" aspect of these strategies while expanding the universe with even more definitions aiming to reveal their true face, such as "Smart Beta=dumb beta + smart marketing" (Montier, 2013). Others raise alarming voices about investor's appetite of SB ETFs to boost returns since it is reasonably likely that a "Smart Beta crash will be a consequence of the soaring popularity of factor-tilt strategies" (Arnott, Beck, Kalesnik, & West, 2016). With respect to the macroeconomic aspect of the analysis, four variables (GDP, IPI, UR, CPI) are somewhat robust predictors across all models and consequently suggest to at least partly explain SB ETF excess returns. Notably, GDP exhibits the highest sensitivity followed by IPI. The PCR confirms the high importance of GDP and IPI as the first PC is largely driven by these two variables. Additionally, the PC1 was found to be strongly significant in most models. The remaining macroeconomic forces (OP, FDI, MS, IR) performed quite badly on the other hand implying that changes in these macroeconomic forces seem to be scarcely related to SB ETF excess returns (at best).

Given the impressive history of SB strategies attracting enormous inflows, it remains interesting whether the SB ETF market will establish itself and potentially take over the traditional ETF market in terms of market value. In particular, the development of management fees will likely have an important impact on the future of SB ETFs. Until then, the factor zoo is expected to be flourishing.

Bibliography

2018 Investment Company Fact Book: A Review of Trends and Activities in the Investment Company Industry. (2018).

Adam, A. M., & Tweneboah, G. (2008). Macroeconomic factors and stock market movement: Evidence from Ghana. *Available at SSRN 1289842.*

Adrangi, B., Chatrath, A., & Raffiee, K. (1999). Inflation, output, and stock prices: Evidence from two major emerging markets. *Journal of Economics and Finance, 23*(3), 266-278.

Adrian, T., & Rosenberg, J. (2008). Stock returns and volatility: Pricing the short-run and long-run components of market risk. *The Journal of Finance*, *63*(6), 2997-3030.

Alam, M. D., & Uddin, G. (2009). Relationship between interest rate and stock price: empirical evidence from developed and developing countries. *International Journal of Business and Management (ISSN 1833-3850), 4*(3), 43-51.

Al-Mwalla, M., & Karasneh, M. (2011). Fama and French three factor model: Evidence from emerging market. *European Journal of Economics, Finance and Administrative Sciences*, *41*(14), 132-140.

Amenc, N., Goltz, F., & Lodh, A. (2016). Smart Beta is not monkey business. *The journal of index investing*, *6*(4), 12-29.

Amenc, N., Goltz, F., Le Sourd, V., & Lodh, A. (2015). Alternative Equity Beta Investing: A Survey. *An EDHEC-Risk Institute Publication*.

Amenc, N., Goltz, F., Sivasubramanian, S., & Lodh, A. (2015). Robustness of Smart Beta strategies. *The Journal of Index Investing*, *6*(1), 17.

Arnott, R. D., Beck, N., Kalesnik, V., & West, J. (2016). How Can'Smart Beta'Go Horribly Wrong?. *Available at SSRN 3040949*.

Arnott, R. D., Clements, M., & Kalesnik, V. (2017). Why Factor Tilts are Not Smart'Smart Beta'. *Available at SSRN 3040970*.

Arnott, R. D., Hsu, J., & Moore, P. (2005). Fundamental indexation. *Financial Analysts Journal*, *61*(2), 83-99.

Arnott, R. D., Hsu, J., Kalesnik, V., & Tindall, P. (2013). The Surprising Alpha from Malkiel's Monkey and Upside-Down Strategies. *Journal of Portfolio Management, 39*(4), 91-105.

Arnott, R., Kalesnik, V., Moghtader, P., & Scholl, C. (2010). Beyond cap weight. *Journal of Indexes*, *16*(1), 16-29.

Arshanapalli, B., Daniel Coggin, T., & Doukas, J. (1998). Multifactor asset pricing analysis of international value investment strategies. *Journal of Portfolio Management*, *24*(4), 10-23.

Asai, M., & Shiba, T. (1995). The Japanese stock market and the macroeconomy: an empirical investigation. *Financial Engineering and the Japanese Markets*, *2*(3), 259-267.

Asness, C. (2006). The value of fundamental indexing. *Institutional Investor International Edition*, *31*(8), 66.

Asness, C. S., Frazzini, A., & Pedersen, L. H. (2019). Quality minus junk. *Review of Accounting Studies*, *24*(1), 34-112.

AXA Investment Managers (2013). Australian Investors Embrace Smart Beta Strategies. From The Source, 11 April; available at https://www.adviservoice.com.au/2013/04/australian-investors-embrace-smart-beta-strategies/.

Baesel, J. B. (1974). On the assessment of risk: Some further considerations. *The Journal of Finance*, *29*(5), 1491-1494.

Bailey, R. E. (2005). *The economics of financial markets*. Cambridge University Press.

Bali, T. G., Brown, S. J., Murray, S., & Tang, Y. (2017). A lottery-demand-based explanation of the beta anomaly. *Journal of Financial and Quantitative Analysis*, *52*(6), 2369-2397.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of financial economics*, *9*(1), 3-18.

Basu, S. (1983). The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence. *Journal of financial economics*, *12*(1), 129-156.

Bernanke, B. S. (1986). Alternative explanations of the money-income correlation. In *Carnegie-rochester conference series on Public Policy* (Vol. 25, pp. 49-99). North-Holland.

Bhandari, L. C. (1988). Debt/equity ratio and expected common stock returns: Empirical evidence. *The Journal of Finance*, *43*(2), 507-528.

Bilson, C. M., Brailsford, T. J., & Hooper, V. J. (2001). Selecting macroeconomic variables as explanatory factors of emerging stock market returns. *Pacific-Basin Finance Journal*, *9*(4), 401-426.

Bird, R., Menzies, G., Dixon, P., & Rimmer, M. (2011). The economic costs of US stock mispricing. *Journal of Policy Modeling, 33*(4), 552-567.

Birz, G., & Lott Jr, J. R. (2011). The effect of macroeconomic news on stock returns: New evidence from newspaper coverage. *Journal of Banking & Finance, 35*(11), 2791-2800.

Black, F., Jensen, M. C., & Scholes, M. (1972). The capital asset pricing model: Some empirical tests. *Studies in the theory of capital markets*, *81*(3), 79-121.

Blanchard, O. J. (1981). Output, the stock market, and interest rates. *American Economic Review*, *71*(1), 132-143.

Blume, M. E. (1971). On the Assessment of Risk. *The Journal of Finance*, *26*(1), 1-10.

BMO Global Asset Management, 2014. Canadian ETF Outlook Mid-Year 2014. Montreal: Bank of Montreal, July.

Bodie, Z. (1976). Common stocks as a hedge against inflation. *The Journal of Finance, 31*(2), 459-470.

Bogle, J. (2015, January 27) quoted in J. Segal. Active Managers Losing Ground Can Thank John Bogle. *Institutional Investor.*

Boudoukh, J., & Richardson, M. (1993). Stock returns and inflation: A long-horizon perspective. *The American economic review, 83*(5), 1346-1355.

Boyadzhiev, D., Bryan, A., Choy, J., Johnson, B., & Venkataraman, A. (2017). A Global Guide to Strategic-Beta Exchange-Traded Products. *Morningstar Manager Research.* Retrieved from https://www.morningstar.com/content/dam/marketing/shared/research/foundational/824512.pdf

Boyd, J. H., Hu, J., & Jagannathan, R. (2005). The stock market's reaction to unemployment news: Why bad news is usually good for stocks. *The Journal of Finance, 60*(2), 649-672.

Brandt, M. W., & Wang, K. Q. (2003). Time-varying risk aversion and unexpected inflation. *Journal of Monetary Economics*, *50*(7), 1457-1498.

Bruce, I., & Levy, K. N. (2015). Smart Beta: too good to be true? *Journal of Financial Perspectives*, *3*(2).

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (2012). *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press.

Campbell, J. Y., Polk, C., & Vuolteenaho, T. (2009). Growth or glamour? Fundamentals and systematic risk in stock returns. *The Review of Financial Studies*, *23*(1), 305-344.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57-82.

Casson, R. J., & Farmer, L. D. (2014). Understanding and checking the assumptions of linear regression: a primer for medical researchers. *Clinical & experimental ophthalmology, 42(*6), 590-596.

Chan, L. K., & Lakonishok, J. (2004). Value and growth investing: Review and update. *Financial Analysts Journal, 60*(1), 71-86.

Chancharat, S., Valadkhani, A., & Havie, C. (2007). The influence of international stock markets and macroeconomic variables on the Thai stock market. *Applied Econometrics and International Development, 7*(1).

Chatterjee, S., & Hadi, A. S. (2012). Variable selection procedures. *Regression analysis by example. 5th edn: Wiley, 306.*

Chen, N. F. (1983). Some empirical tests of the theory of arbitrage pricing. *The Journal of Finance*, *38*(5), 1393-1414.

Chen, N. F., Roll, R., & Ross, S. A. (1986). Economic forces and the stock market. *Journal of Business*, *59*(3), 383-403.

Chen, S. S. (2007). Does monetary policy have asymmetric effects on stock returns?. *Journal of Money, Credit and Banking, 39*(2-3), 667-688.

Cho, D. C. (1984). On Testing the Arbitrage Pricing Theory: Inter-Battery Factor Analysis. *The Journal of Finance*, *39*(5), 1485-1502.

Cho, D. C., Eun, C. S., & Senbet, L. W. (1986). International arbitrage pricing theory: An empirical investigation. *The Journal of Finance, 41*(2), 313-329.

Chow, T. M., Hsu, J., Kalesnik, V., & Little, B. (2011). A survey of alternative equity index strategies. *Financial Analysts Journal, 67*(5), 37-57.

Chui, A. C., & Wei, K. J. (1998). Book-to-market, firm size, and the turn-of-the-year effect: Evidence from Pacific-Basin emerging markets. *Pacific-Basin finance journal*, *6*(3-4), 275-293.

Clare, A. D., & Thomas, S. H. (1994). Macroeconomic Factors, the APT and the UK Stockmarket. *Journal of Business Finance & Accounting*, *21*(3), 309-330.

Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of Finance, 66*(4), 1047-1108.

Cogent Research (2014). *The Evolution of Smart Beta ETFs: Smart Beta ETFs Poised for Growth Among Institutions.* Livonia, Mich.: Market Strategies International, January.

Cohen, P., West, S. G., & Aiken, L. S. (2014). *Applied multiple regression/correlation analysis for the behavioral sciences.* Psychology Press.

Cooley, T. F., & LeRoy, S. F. (1985). Atheoretical macroeconometrics: a critique. *Journal of Monetary Economics*, *16*(3), 283-308.

Cooper, R. V. (1974). Efficient capital markets and the quantity theory of money. *Journal of Finance,* 887-908.

Croux, C., Dhaene, G., & Hoorelbeke, D. (2004). Robust standard errors for robust estimators. *CES-Discussion paper series (DPS) 03.16*, 1-20.

Cuthbertson, K., & Nitzsche, D. (2005). *Quantitative financial economics: stocks, bonds and foreign exchange.* John Wiley & Sons.

Daniel, K., & Titman, S. (1997). Evidence on the characteristics of cross sectional variation in stock returns. *the Journal of Finance*, *52*(1), 1-33.

Davis, J. L., Fama, E. F., & French, K. R. (2000). Characteristics, covariances, and average returns: 1929 to 1997. *The Journal of Finance*, *55*(1), 389-406.

Day, T. E. (1984). Real stock returns and inflation. *The Journal of Finance, 39*(2), 493-502.

De Meyer, D. (2016). Can European investors outsmart the market? (Doctoral dissertation, GHENT UNIVERSITY).

Dhrymes, P. J., Friend, I., & Gultekin, N. B. (1984). A critical reexamination of the empirical evidence on the arbitrage pricing theory. *The Journal of Finance, 39*(2), 323-346.

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: Journal of the Econometric Society*, 1057-1072.

Doukas, J. A., Kim, C. F., & Pantzalis, C. (2010). Arbitrage risk and stock mispricing. *Journal of Financial and Quantitative Analysis, 45*(4), 907-934.

Dumas, B., & Solnik, B. (1995). The world price of foreign exchange risk. *The Journal of Finance, 50*(2), 445-479.

Dunteman, G. H. (1989). *Principal components analysis* (No. 69). Sage.

Durham, J. B. (2003). Monetary policy and stock price returns. *Financial Analysts Journal, 59*(4), 26-35.

Eraslan, V. (2013). Fama and french three-factor model: Evidence from Istanbul stock exchange. *Business and Economics Research Journal*, *4*(2), 11-22.

Evans, R., & Wilson, C. (2018, September 13). How ETFs Became the Market. *Bloomberg LP*. Retrieved from https://www.bloomberg.com/graphics/2018-growing-etf-market/?srnd=etfs.

Fama, E. F. (1965). Random walks in stock market prices. *Financial Analysts Journal*, *21*(5), 55-59.

Fama, E. F. (1991). Efficient capital markets: II. *The Journal of Finance*, *46*(5), 1575-1617.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. T*he Journal of Finance*, *47*(2), 427-465.

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

Fama, E. F., & French, K. R. (1996). The CAPM is wanted, dead or alive. *The Journal of Finance*, *51*(5), 1947-1958.

Fama, E. F., & French, K. R. (1998). Value versus growth: The international evidence. *The Journal of Finance*, *53*(6), 1975-1999.

Fama, E. F., & French, K. R. (2004). The capital asset pricing model: Theory and evidence. *Journal of Economic Perspectives*, *18*(3), 25-46.

Fama, E. F., & French, K. R. (2012). Size, value, and momentum in international stock returns. *Journal of Financial Economics*, *105*(3), 457-472.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, *81*(3), 607-636.

Fama, E. F., & Schwert, G. W. (1977). Asset returns and inflation. *Journal of Financial Economics*, *5*(2), 115-146.

Fama, E. F., Fisher, L., Jensen, M. C., & Roll, R. (1969). The adjustment of stock prices to new information. *International Economic Review*, *10*(1), 1-21.

Fedorova, E. A., & Pankratov, K. A. (2010). Influence of macroeconomic factors on the Russian stock market. *Studies on Russian Economic Development, 21*(2), 165-168.

Ferson, W. E., & Harvey, C. R. (1994). Sources of risk and expected returns in global equity markets. *Journal of Banking & Finance, 18*(4), 775-803.

Ferson, W. E., Kandel, S., & Stambaugh, R. F. (1987). Tests of asset pricing with time-varying expected risk premiums and market betas. *The Journal of Finance, 42*(2), 201-220.

Fisher, G. S. (2013). Dividend Investing: A Value Tilt in Disguise?. *Journal of Financial Planning*, 52.

Flannery, M. J., & Protopapadakis, A. A. (2002). Macroeconomic factors do influence aggregate stock returns. *The Review of Financial Studies, 15*(3), 751-782.

Foye, J., Mramor, D., & Pahor, M. (2013). A Respecified Fama French Three-Factor Model for the New European Union Member States. *Journal of International Financial Management & Accounting*, *24*(1), 3-25.

FTSE Russell (2017a). Trends and outlook for Smart Beta. *Insights*. Retrieved from https://www.ftserussell.com/sites/default/files/trends-and-outlook-for-smart-beta-final-1.pdf

FTSE Russell (2017b). The anatomy of Smart Beta. *Insights* (May 2017). Retrieved from https://www.ftserussell.com/research/anatomy-smart-beta

FTSE Russell (2018). Five-year trends and outlook for Smart Beta. *Insights* (June 2018). Retrieved from https://www.ftserussell.com/sites/default/files/research/five-year_trends_and_outlook_for_smart_beta_final_rev.pdf

Gastineau, G. L. (2001). An introduction to exchange-traded funds (ETFs). *Journal of Portfolio Management and Economics*, *27*(3), 88-96.

Geske, R., & Roll, R. (1983). The fiscal and monetary linkage between stock returns and inflation. *The Journal of Finance, 38*(1), 1-33.

Giannone, D., Reichlin, L., & Sala, L. (2004). Monetary policy in real time. *NBER macroeconomics annual*, *19*, 161-200.

Glushkov, D. (2016). How Smart Are Smart Beta Exchange-Traded Funds? Analysis of Relative Performance and Factor Exposure. *Journal of Investment Consulting*, *17*(1), 50-74.

Granger, C. W., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics, 2*(2), 111-120.

Green, J., Hand, J. R., & Zhang, X. F. (2014). The remarkable multidimensionality in the cross-section of expected US stock returns. *Available at SSRN, 2262374.*

Griffin, J. M. (2002). Are the Fama and French factors global or country specific?. *The Review of Financial Studies*, *15*(3), 783-803.

Grinold, R. C. (1992). Are benchmark portfolios efficient? *Journal of Portfolio Management*, *19*(1), 34-40.

Growth versus value investing - Fidelity. (2016). Retrieved from https://www.fidelity.com/learning-center/investment-products/mutual-funds/2-schools-growth-vs-value

Gruber, M. J. (1996). Another puzzle: The growth in actively managed mutual funds. *The Journal of Finance, 51*(3), 783-810.

Gultekin, N. B. (1983). Stock market returns and inflation: evidence from other countries. *The Journal of Finance, 38*(1), 49-65.

Hamilton, J. D., & Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics, 64*(1-2), 307-333.

Harris, D. (1997). Principal components analysis of cointegrated time series. *Econometric Theory, 13*(4), 529-557.

Harvey, C. R., Liu, Y., & Zhu, H. (2016). … and the cross-section of expected returns. *The Review of Financial Studies, 29*(1), 5-68.

Hatemi-j, A. (2003). A new method to choose optimal lag order in stable and unstable VAR models. *Applied Economics Letters*, *10*(3), 135-137.

Haugen, R. A., & Baker, N. L. (1991). The efficient market inefficiency of capitalization–weighted stock portfolios. *The Journal of Portfolio Management*, *17*(3), 35-40.

Haugen, R. A., & Baker, N. L. (1993). Interpreting the evidence on risk and expected return: Comment. *Journal of Portfolio Management*, *19*(3), 36-43.

Haugen, R. A., & Baker, N. L. (1996). Commonality in the determinants of expected stock returns. *Journal of Financial Economics*, *41*(3), 401-439.

Henry, O. T., Olekalns, N., & Thong, J. (2004). Do stock market returns predict changes to output? Evidence from a nonlinear panel data model. *Empirical Economics, 29*(3), 527-540.

Homa, K. E., & Jaffee, D. M. (1971). The supply of money and common stock prices. *The Journal of Finance, 26*(5), 1045-1066.

Hondroyiannis, G., & Papapetrou, E. (2001). Macroeconomic influences on the stock market. *Journal of Economics and Finance*, *25*(1), 33-49.

Hong, H. (1977). Inflation and the market value of the firm: Theory and tests. *The Journal of Finance, 32*(4), 1031-1048.

Hsu, J. (2014). Value Investing: Smart Beta versus Style Indexes. *The Journal of Index Investing*, *5*(1), 121-126.

Hsu, J. C. (2004). Cap-weighted portfolios are sub-optimal portfolios. *Journal of Investment Management*, *4*(3).

Hsu, J. C. (2014). Value investing: Smart Beta versus Style Indexes. *Journal of Indexes, Forthcoming, 4*, 121-133.

Hsu, J., & Kalesnik, V. (2014). Finding Smart Beta in the factor zoo. *Research Affiliates (July).*

Hsu, J., Kalesnik, V., & Viswanathan, V. (2015). A framework for assessing factors and implementing Smart Beta strategies. *The Journal of Index Investing, 6*(1), 89.

Humpe, A., & Macmillan, P. (2009). Can macroeconomic variables explain long-term stock market movements? A comparison of the US and Japan. *Applied Financial Economics, 19*(2), 111-119.

Hurvich, C. M., & Tsai, C. L. (1989). Regression and time series model selection in small samples. *Biometrika*, *76*(2), 297-307.

Isidore, C. (2018, December 31). 2018 was the worst for stocks in 10 years. *CNN Business*[New York]. Retrieved from https://edition.cnn.com/2018/12/31/investing/dow-stock-market-today/index.html

Jacobs, B. I. (2015). Is Smart Beta state of the art? *Journal of Portfolio Management, 41*(4), 1-3.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning: with applications in R*. New York: springer.

Jegadeesh, N. (1992). Does market risk really explain the size effect?. *Journal of Financial and Quantitative Analysis*, *27*(3), 337-351.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance, 48*(1), 65-91.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *The Journal of Finance*, *23*(2), 389-416.

Jensen, M. C. (1969). Risk, the pricing of capital assets, and the evaluation of investment portfolios. *Journal of Business*, *42*(2), 167-247.

Jensen, M. C. (1978). Some anomalous evidence regarding market efficiency. *Journal of Financial Economics*, *6*(2/3), 95-101.

Johnson, B. (2019, March 13). Strategic beta ETF market still expanding. Retrieved from http://www.morningstar.co.uk/uk/news/182188/strategic-beta-etf-market-still-expanding.aspx

Jolliffe, I. (2011). Principal component analysis. In *International encyclopedia of statistical science* (pp. 1094-1096). Springer, Berlin, Heidelberg.

Jones, S. L., & Netter, J. M. (2008). Efficient capital markets. *The Concise Encyclopedia of Economic*, 15.

Kahn, R. N., & Lemmon, M. (2016). The asset manager's dilemma: How Smart Beta is disrupting the investment management industry. *Financial Analysts Journal*, *72*(1), 15-20.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31-36.

Kolluri, B., & Wahab, M. (2008). Stock returns and expected inflation: evidence from an asymmetric test specification. *Review of Quantitative Finance and Accounting, 30*(4), 371-395.

Kothari, S. P., Shanken, J., & Sloan, R. G. (1995). Another look at the cross-section of expected stock returns. *The Journal of Finance*, *50*(1), 185-224.

Kunst, R., & Neusser, K. (1986). A forecasting comparison of some VAR techniques. *International Journal of Forecasting*, *2*(4), 447-456.

Kwag, S., & Whi, S. (2006). Value investing and the business cycle. *Journal of Financial Planning, 19*(1), 64.

Lakonishok, J., & Shapiro, A. C. (1986). Systematic risk, total risk and size as determinants of stock market returns. *Journal of Banking & Finance, 10*(1), 115-132.

Leamer, E. E. (1985, January). Vector autoregressions for causal inference?. In *Carnegie-rochester conference series on Public Policy* (Vol. 22, pp. 255-304). North-Holland.

Levi, Y., & Welch, I. (2014). Long Term Capital Budgeting. Working Paper (March 29).

Li, Y. D., İşcan, T. B., & Xu, K. (2010). The impact of monetary policy shocks on stock prices: Evidence from Canada and the United States. *Journal of International Money and Finance*, *29*(5), 876-896.

Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The Journal of Finance, 20*(4), 587-615.

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), 13-37.

Lütkepohl, H. (1993). *Introduction to Multiple Time Series Analysis, Second Edition.* Berlin: Springer-Verlag.

Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of Economic Perspectives*, *17*(1), 59-82.

Malkiel, B. G. (2005). Reflections on the efficient market hypothesis: 30 years later. *Financial Review*, *40*(1), 1-9.

Malkiel, B. G. (2014). Is Smart Beta really smart?. *The Journal of Portfolio Management*, *40*(5), 127-134.

Malkiel, B. G., & Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance, 25*(2), 383-417.

Malkiel, B. G., & Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance, 25*(2), 383-417.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77-91.

McQueen, G., & Roley, V. V. (1993). Stock prices, news, and business conditions. *The Review of Financial Studies*, *6*(3), 683-707.

Miller, M., & Modigliani, F. (1961). Dividend policy, growth, and the valuation of shares.

Mirza, N., & Shabbir, G. (2005). The death of CAPM: A critical review. *Lahore Journal of Economics*, *10*(2), 35-54.

Moerman, G. A. (2005). How domestic is the Fama and French three-factor model? An application to the Euro area. *ERIM Report Series Reference No. ERS-2005-035-F&A.*

Montgomery, D. C., Peck, E. A., Vining, G. G., & Kowalski, S. M. (2001). *Student solutions manual to accompany Introduction to linear regression analysis.* Wiley.

Montier, J. (2013). No Silver Bullets in Investing. *White Paper december*.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica: Journal of the Econometric Society*, 768-783.

Muhammad, S. D., Hussain, A., Ali, A., & Jalil, M. A. (2009). Impact of macroeconomics variables on stock prices: Empirical evidence in case of KSE. *Available at SSRN 1683357.*

Myers, R. H. (1990). *Classical and modern regression with applications* (Vol. 2). Belmont, CA: Duxbury press.

Newey, W. K., & West, K. D. (1986). A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix.

Newlands, C., & Marriage, M. (2016, October 24). 99% of actively managed US equity funds underperform. Retrieved from https://www.ft.com/content/e139d940-977d-11e6-a1dc-bdf38d484582

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, *108*(1), 1-28.

O'Shaughnessy, J., Meredith, C., Bartone, S., Fairchild, T., O'Shaughnessy, P., Stanhope, E., & Zhu, M. (2012). OSAM research: dividend yield vs. dividend growth. Retrieved from O'Shaughnessy Asset Management website: https://www.osam.com/pdfs/research/_34_OSAM_research_DivYield-vs-DivGrowth_Sep12.pdf

Osborne, J., & Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical assessment, research & evaluation, 8*(2), 1-9.

Öztuna, D., Elhan, A. H., & Tüccar, E. (2006). Investigation of four different normality tests in terms of type 1 error rate and power under different distributions. *Turkish Journal of Medical Sciences, 36*(3), 171-176.

Park, J., & Ratti, R. A. (2008). Oil price shocks and stock markets in the US and 13 European countries. *Energy economics, 30*(5), 2587-2608.

Pearce, D. K., & Roley, V. V. (1984). Stock prices and economic news.

Pensions & Investments (2013). Low-Volatility Strategies Are Gaining Speed in Asia, by Douglas Appell. *Pensions & Investments* 41(1,24 June):1.

Philips, C. B., Bennyhoff, D. G., Kinniry Jr, F. M., Schlanger, T., & Chin, P. (2015). An evaluation of Smart Beta and other rules-based active strategies. *Vanguard white paper*.

Plyakha, Y., Uppal, R., & Vilkov, G. (2015). Why Do Equal-Weighted Portfolios Outperform Value-Weighted Portfolios?. In *Working paper, SSRN 2724535*.

Poon, S., & Taylor, S. J. (1991). Macroeconomic factors and the UK stock market. *Journal of Business Finance & Accounting*, *18*(5), 619-636.

Prather, L., & Bertin, W. J. (1999). Market efficiency, discount-rate changes, and stock returns: A long-term perspective. *Journal of Economics and Finance, 23*(1), 56-63.

Ram, A. (2017, February 5). Signs that Smart Beta price war has started. *Financial Times*. Retrieved from https://www.ft.com/content/7c649480-e4a1-11e6-9645-c9357a75844a

Reinganum, M. R. (1981). A new empirical perspective on the CAPM. *Journal of Financial and Quantitative Analysis, 16*(4), 439-462.

Roenfeldt, R. L., Griepentrog, G. L., & Pflaum, C. C. (1978). Further evidence on the stationarity of beta coefficients. *Journal of Financial and Quantitative Analysis*, *13*(1), 117-121.

Rogalski, R. J., & Vinso, J. D. (1977). Stock returns, money supply and the direction of causality. *The Journal of Finance, 32*(4), 1017-1030.

Rosenberg, B., Reid, K., & Lanstein, R. (1985). Persuasive evidence of market inefficiency. *The Journal of Portfolio Management*, *11*(3), 9-16.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory, 13*(3), 341-360.

Royston, J. P. (1982). An extension of Shapiro and Wilk's W test for normality to large samples. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 31*(2), 115-124.

Russell Investments (2014). Smart Beta: A Deeper Look at Asset Owner Perceptions. Findings from Russell Indexes Global Smart Beta Survey (April). Retrieved from http://russell.com/documents/indexes/research/smart-beta-survey.pdf

Schwert, G. W. (1981). The adjustment of stock prices to information about inflation. *The Journal of Finance*, *36*(1), 15-29.

Sehgal, S., & Balakrishnan, A. (2013). Robustness of Fama-French three factor model: Further evidence for Indian stock market. *Vision*, *17*(2), 119-127.

Shanken, J. (1992). On the estimation of beta-pricing models. *The Review of Financial Studies*, *5*(1), 1-33.

Shanken, J., & Weinstein, M. (1990). Macroeconomic variables and asset pricing: Further results.

Sharma, R., & Mehta, K. (2013). Fama and French: Three factor model. *SCMS Journal of Indian Management*, *10*(2), 90-105.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance, 19*(3), 425-442.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business, 39*(1), 119-138.

Sharpe, W. F. (1982). Factors in New York stock exchange security returns, 1931–1979. *The Journal of Portfolio Management*, *8*(4), 5-19.

Sharpe, W. F. (1994). The sharpe ratio. *Journal of Portfolio Management*, *21*(1), 49-58.

Sharpe, W. F., & Litterman, R. (2014). Past, present, and future financial thinking. *Financial Analysts Journal*, *70*(6), 16-22.

Shiller, R. J. (2003). From efficient markets theory to behavioral finance. *Journal of Economic Perspectives*, *17*(1), 83-104.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: journal of the Econometric Society*, 1-48.

Singh, T., Mehta, S., & Varsha, M. S. (2011). Macroeconomic factors and stock returns: Evidence from Taiwan. *Journal of Economics and International Finance, 3*(4), 217-227.

Sortino, F. A., & Forsey, H. J. (1996). On the use and misuse of downside risk. *Journal of Portfolio Management, 22*(2), 35.

Sortino, F. A., & Price, L. N. (1994). Performance measurement in a downside risk framework. *The Journal of Investing, 3*(3), 59-64.

Sortino, F. A., & Van Der Meer, R. (1991). Downside risk. *Journal of Portfolio Management, 17*(4), 27.

Spence Johnson (2014). *Smart Beta Funds: A Growing Alternative to Active Approaches.* Investment Product Market Intelligence, 2014. London: Spence Johnson, April.

Steward, M. (2014). Smart Beta: Smart Investing or Smart Trading. *Investment and Pensions Europe*.

Stock, J. H., & Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American statistical association*, *97*(460), 1167-1179.

Taneja, Y. P. (2010). Revisiting fama french three-factor model in indian stock market. *The Journal of Business Perspective*, *14*(4), 267-274.

Tangjitprom, N. (2011). Macroeconomic factors of emerging stock market: the evidence from Thailand. *International Journal of Financial Research, 3*(2), 105-114.

Tangjitprom, N. (2012). The review of macroeconomic factors and stock returns. *International Business Research, 5*(8), 107.

Thompson, B. (1995). Stepwise regression and stepwise discriminant analysis need not apply here: A guidelines editorial.

Tobin, J. (1958). Liquidity preference as behavior towards risk. *The Review of Economic Studies*, *25*(2), 65-86.

Trzcinka, C. (1986). On the number of factors in the arbitrage pricing model. *The Journal of Finance, 41*(2), 347-368.

Ung, D., & Luk, P. (2016). What Is in Your Smart Beta Portfolio? A Fundamental and Macroeconomic Analysis. *The Journal of Index Investing*, *7*(1), 49-77.

Watson, M. W. (2004). Comment on Giannone, Reichlin, and Sala. *NBER Macroeconomics Annual*, *2004,* 216-221.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48*(4), 817-838.

Williams, M. N., Grajales, C. A. G., & Kurkiewicz, D. (2013). Assumptions of multiple regression: Correcting two misconceptions.

Willmer, S. (2016, December 19). BlackRock Cuts Fees on ETFs as Price War Moves to Smart Beta. *Bloomberg BusinessWeek.* Retrieved from https://www.bloomberg.com/news/articles/2016-12-19/blackrock-cuts-fees-on-etfs-as-price-war-moves-to-smart-beta

Yogo, M. (2006). A consumption-based explanation of expected stock returns. *The Journal of Finance, 61*(2), 539-580.

Appendix I: Linearity Assumption

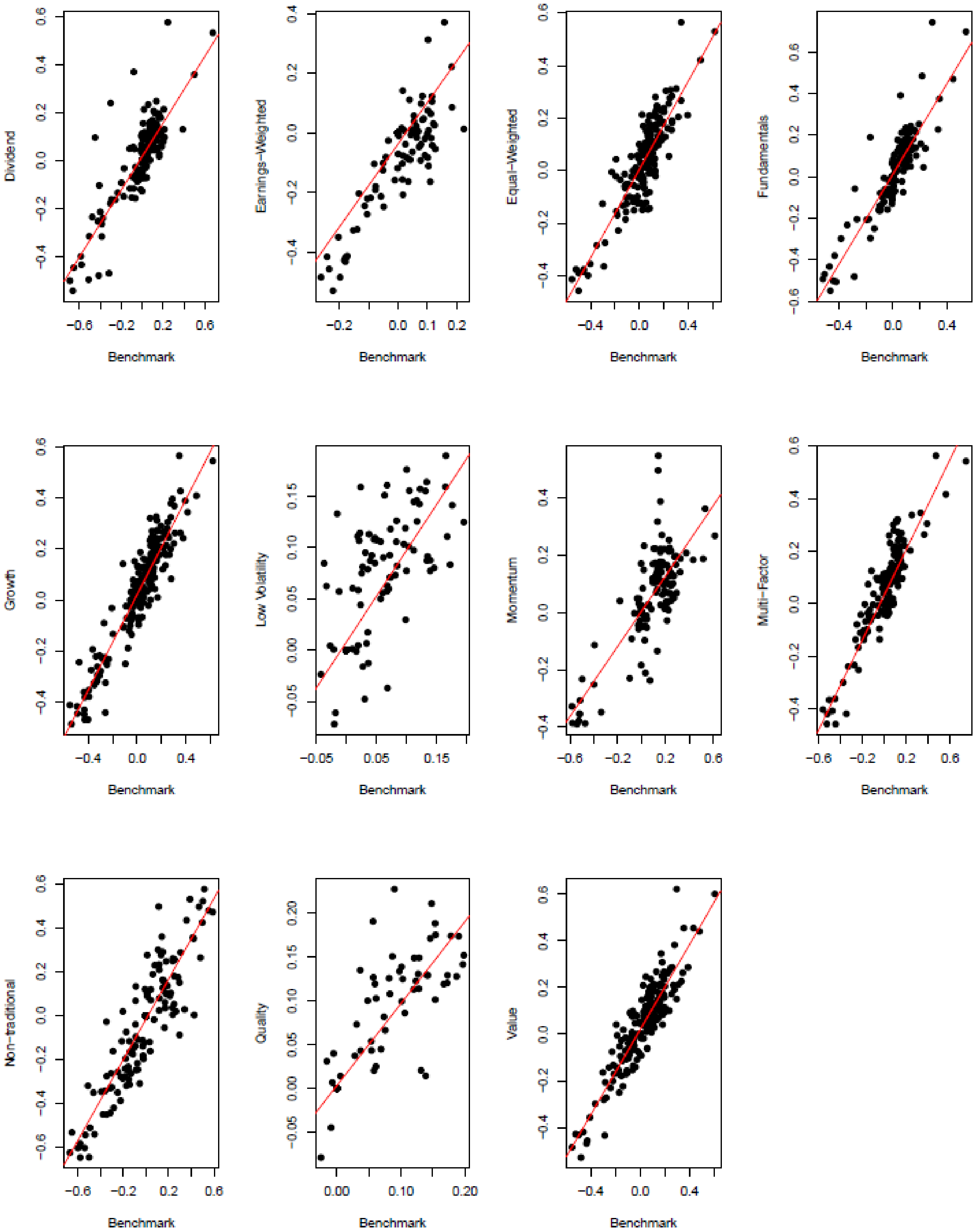


Figure: Linearity Check for equally-weighted SB ETFs per category with respect to declared benchmark

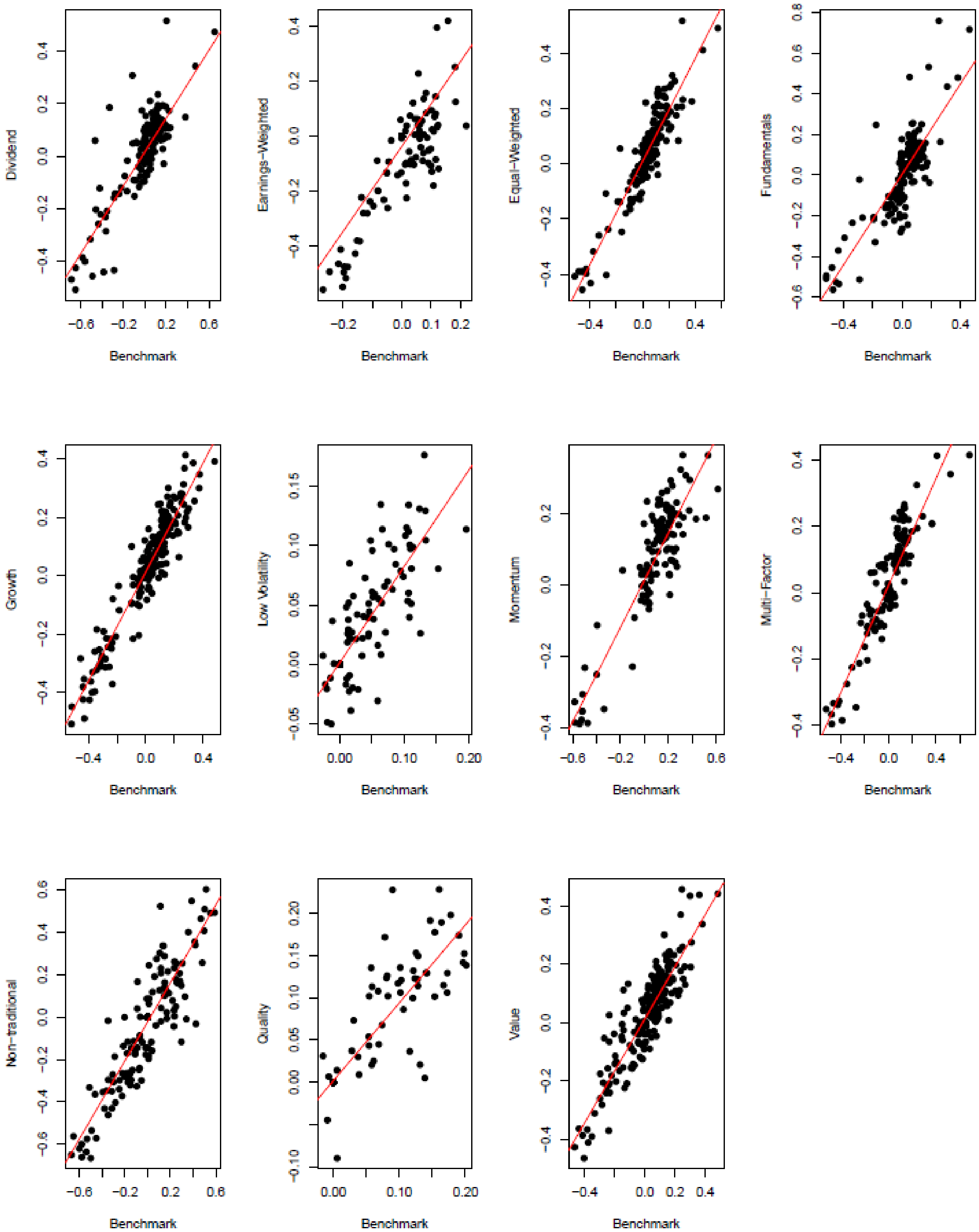


Figure: Linearity Check for size-weighted SB ETFs per category with respect to declared benchmark

Appendix II: Stationarity & Autocorrelation Assumption

	Variable	*Augmented Dickey Fuller test statistic (equal-weighted)*	*Augmented Dickey Fuller test statistic (size-weighted)*
SB ETF Category (dependent variable)	*Dividend*	*-3.352982 (***)*	*-3.240953 (***)*
	Earnings-Weighted	*-3.054778 (***)*	*-2.967624 (***)*
	Equal-Weighted	*-3.331092 (***)*	*-3.031834 (***)*
	Fundamentals	*-3.475782 (***)*	*-3.615877 (***)*
	Growth	*-3.193456 (***)*	*-3.064685 (***)*
	Low volatility	*-2.379082 (**)*	*-2.793156 (***)*
	Value	*-3.3012 (**)*	*-3.17868 (***)*
	Momentum	*-3.501904 (***)*	*-2.848691 (***)*
	Multi-Factor	*-3.069884 (***)*	*-2.998784 (***)*
	Non-traditional	*-3.453642 (***)*	*-3.637755 (***)*
	Quality	*-3.287546 (***)*	*-3.299504 (***)*
Macroeconomic Variables	*Gross Domestic Product (GDP)*	*-2.80993 (***)*	
	Industrial Production Index (IPI)	*-7.360732 (***)*	
	Unemployment Rate (UR)	*-7.742175 (***)*	
	Interest Rate (IR)	*-11.22309 (***)*	
	Money Supply (MS)	*-4.384812 (***)*	
	Consumer Price Index (CPI)	*-11.71027 (***)*	
	Oil Price (OP)	*-8.499673 (***)*	
	Foreign Direct Investment (FDI)	*-10.60664 (***)*	

*Dickey Fuller test to check for stationarity for monthly SB excess returns and macroeconomic variables. Note that significance level is indicated with *** (<0.01); ** (<0.05); * (<0.1) respectively. Under the null hypothesis the time-series has a unit root. Rejecting the null hypothesis (t-statistic < critical value) implies no unit root and hence that the time-series is a stationary process.*

SB ETF Category	*Durbin-Watson test statistic*	*p-value*	*Breusch-Pagan test statistic*	*p-value*
Dividend	*0.64171*	*2.2e-16*	*5.5394*	*0.01859*
Earnings-Weighted	*0.49807*	*2.2e-16*	*9.0033*	*0.002695*
Equal-Weighted	*0.93943*	*2.318e-16*	*2.4167*	*0.12*
Fundamentals	*0.97944*	*2.696e-15*	*0.77559*	*0.3785*
Growth	*1.4808*	*3.071e-05*	*0.072677*	*0.7875*
Low volatility	*1.3178*	*7.934e-08*	*8.4919*	*0.003567*
Value	*1.3112*	*5.931e-08*	*0.71343*	*0.3983*
Momentum	*1.1759*	*1.266e-10*	*2.4999*	*0.1139*
Multi-Factor	*0.67042*	*2.2e-16*	*2.5188*	*0.1125*
Non-traditional	*1.4915*	*4.36e-05*	*3.3707*	*0.06637*
Quality	*1.6129*	*0.001347*	*28.489*	*9.425e-08*

*Significance Codes: *** (<0.01); ** (<0.05); * (<0.1)*

Table Durbin Watson & Breusch Pagan test to control for heteroscedasticity and autocorrelation for Smart Beta equally-weighted portfolios (Simple Regression)

SB ETF Category	Durbin-Watson test statistic	p-value	Breusch-Pagan test statistic	p-value
Dividend	0.64422	2.2e-16	7.6009	0.005834
Earnings-Weighted	0.48438	2.2e-16	12.452	0.0004176
Equal-Weighted	1.5723	0.0004643	2.3166	0.128
Fundamentals	0.60516	2.2e-16	0.7922	0.3734
Growth	0.15121	2.2e-16	0.20729	0.6489
Low volatility	1.6927	0.008318	44.886	2.088e-11
Value	0.94355	2.969e-16	1.6046	0.2052
Momentum	0.97017	1.529e-15	9.0931	0.002566
Multi-Factor	0.72871	2.2e-16	6.4421	0.01114
Non-traditional	1.399	1.86e-06	4.1834	0.04082
Quality	1.6504	0.003305	36.549	1.489e-09

*Significance Codes: *** (<0.01); ** (<0.05); * (<0.1)*

Table Durban Watson & Breusch Pagan test to control for heteroscedasticity and autocorrelation for SB size-weighted portfolios (Simple Regression)

SB ETF Category	Durbin-Watson test statistic	p-value	Breusch-Pagan test statistic	p-value
Dividend	0.54757	2.2e-16	12.695	0.1228
Earnings-Weighted	0.18334	2.2e-16	11.093	0.1965
Equal-Weighted	0.4247	2.2e-16	9.7492	0.2831
Fundamentals	0.53072	2.2e-16	11.769	0.1618
Growth	0.51078	2.2e-16	10.066	0.2604
Low volatility	0.30379	2.2e-16	11.964	0.1528
Value	0.52214	2.2e-16	7.4705	0.4868
Momentum	0.67755	2.2e-16	9.667	0.2892
Multi-Factor	0.49932	2.2e-16	13.185	0.1056
Non-traditional	0.3091	2.2e-16	26.514	0.0008572
Quality	0.27919	2.2e-16	8.5487	0.3818

*Significance Codes: *** (>0.01); ** (0.05); * (<0.1)*

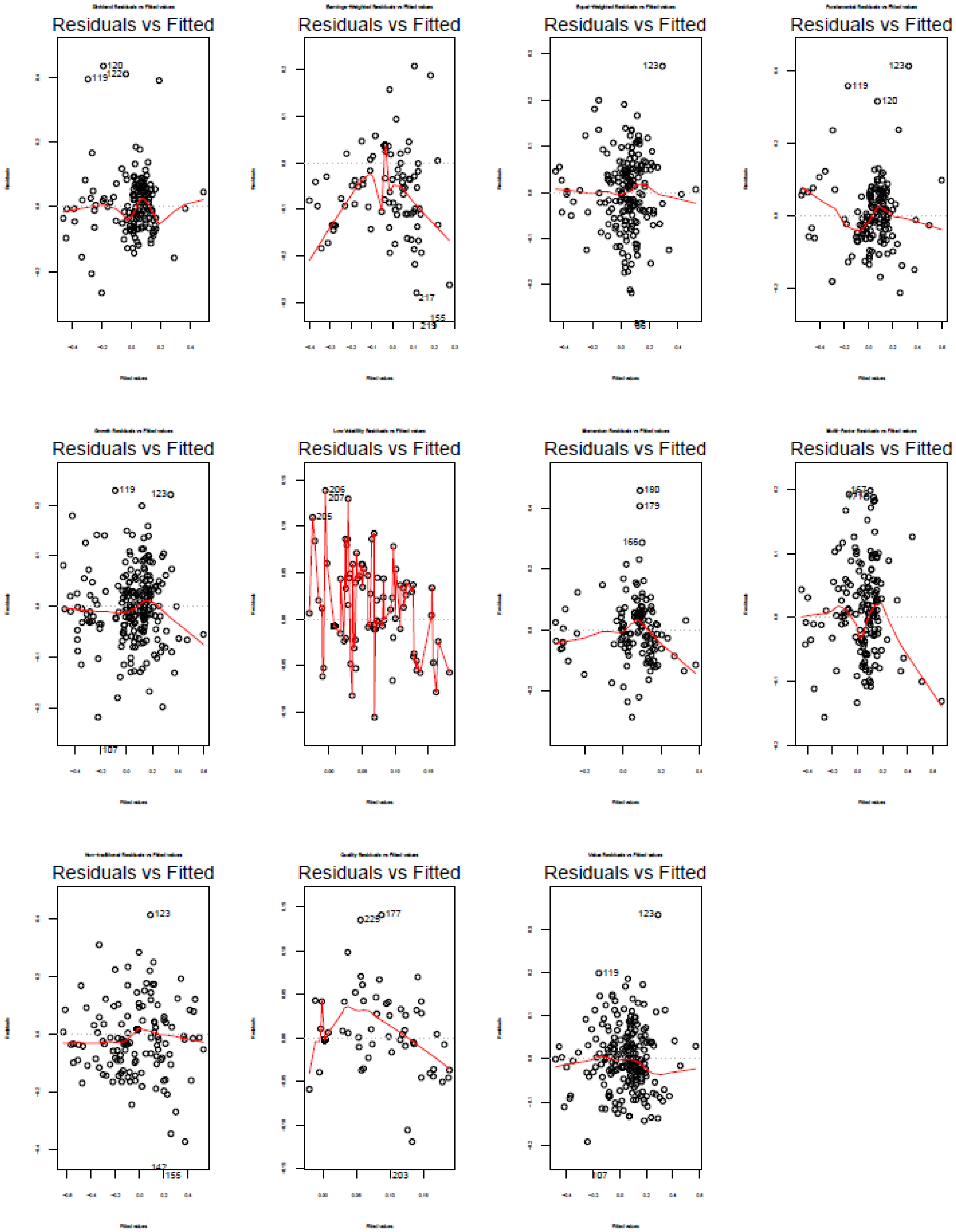
Table: Durban Watson & Breusch Pagan test to control for heteroscedasticity and autocorrelation for SB equally-weighted portfolios (Macroeconomic Model)

SB ETF Category	*Durbin-Watson test statistic*	*p-value*	*Breusch-Pagan test statistic*	*p-value*
Dividend	*0.56569*	*2.2e-16*	*13.381*	*0.0994*
Earnings-Weighted	*0.17778*	*2.2e-16*	*12.86*	*0.1168*
Equal-Weighted	*0.54565*	*2.2e-16*	*10.44*	*0.2355*
Fundamentals	*0.53656*	*2.2e-16*	*12.281*	*0.1391*
Growth	*0.50800*	*2.2e-16*	*8.4825*	*0.3878*
Low volatility	*0.29753*	*2.2e-16*	*10.112*	*0.2572*
Value	*0.48048*	*2.2e-16*	*5.0237*	*0.7550*
Momentum	*0.51702*	*2.2e-16*	*9.7587*	*0.2824*
Multi-Factor	*0.46325*	*2.2e-16*	*12.517*	*0.1296*
Non-traditional	*0.32505*	*2.2e-16*	*26.485*	*0.0009*
Quality	*0.27747*	*2.2e-16*	*8.6745*	*0.3705*

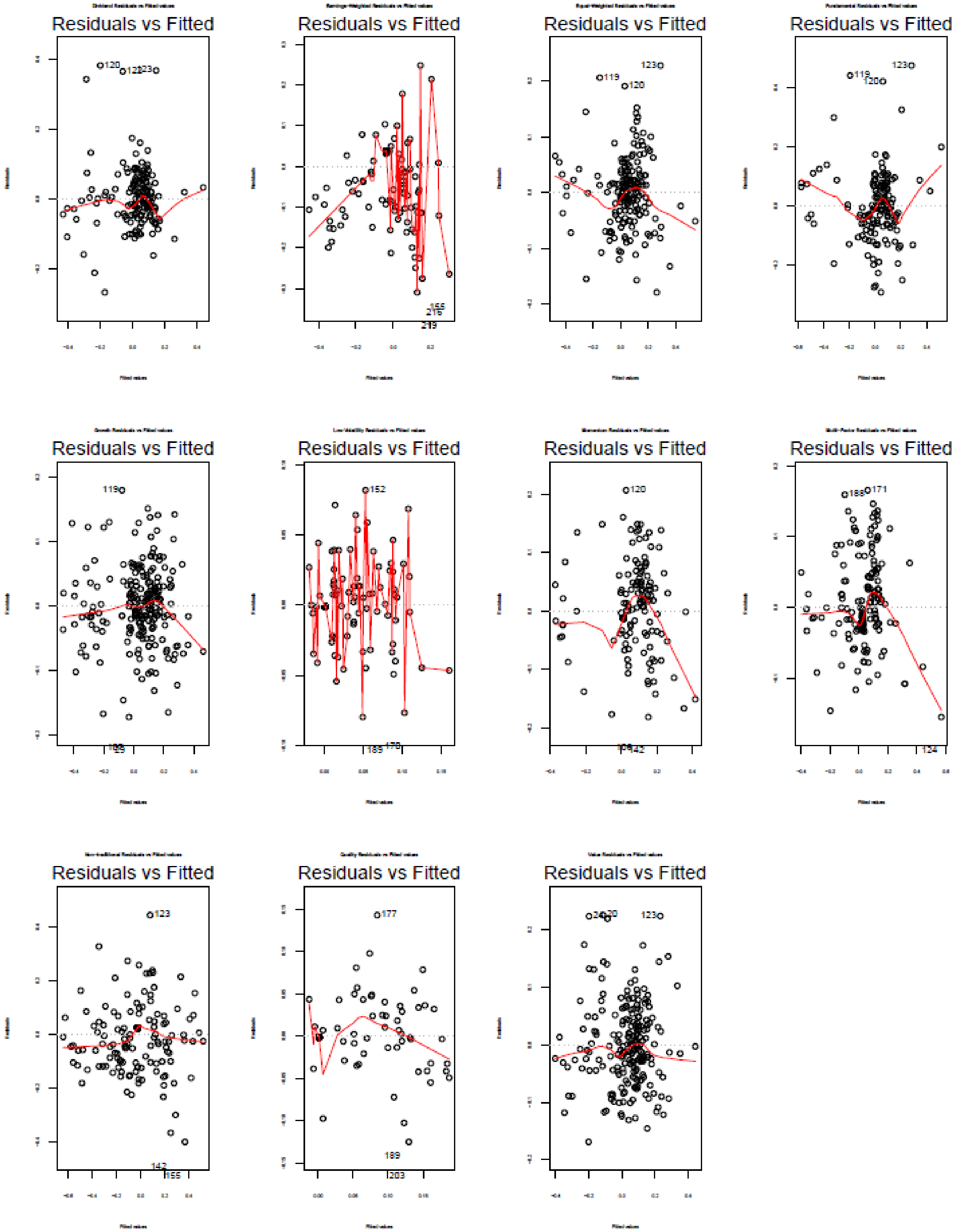
*Significance Codes: *** (>0.01); ** (0.05); * (<0.1)*

Table: Durban Watson & Breusch Pagan test to control for heteroscedasticity and autocorrelation for SB size-weighted portfolios (Macroeconomic Model)

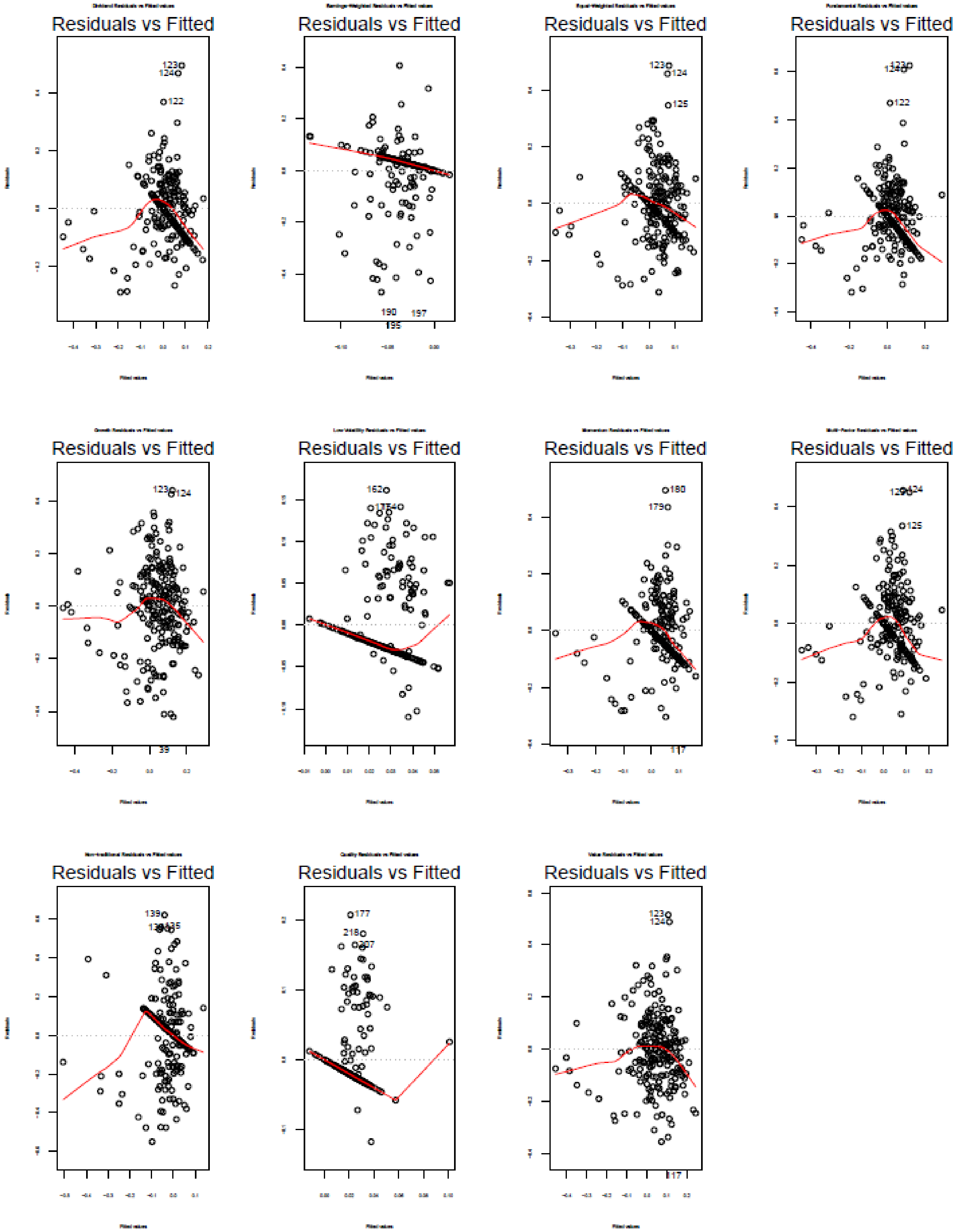
Appendix III: Homoscedasticity Assumption



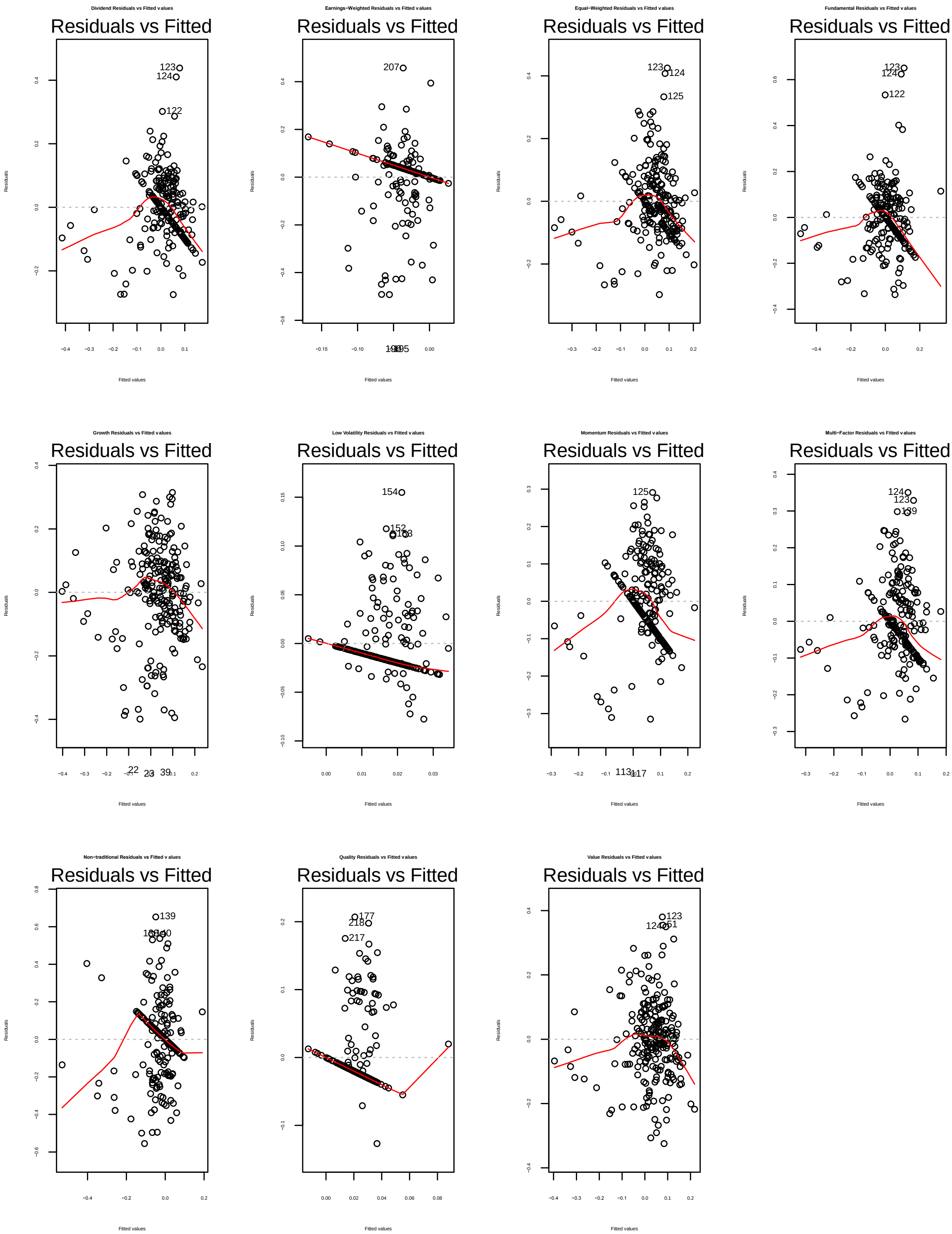
Residuals vs Fitted plots to check for Homoscedasticity for each equally-weighted SB ETF (Simple Regression)



Residuals vs Fitted plot to check for Homoscedasticity for each size-weighted SB ETF (Simple Regression)

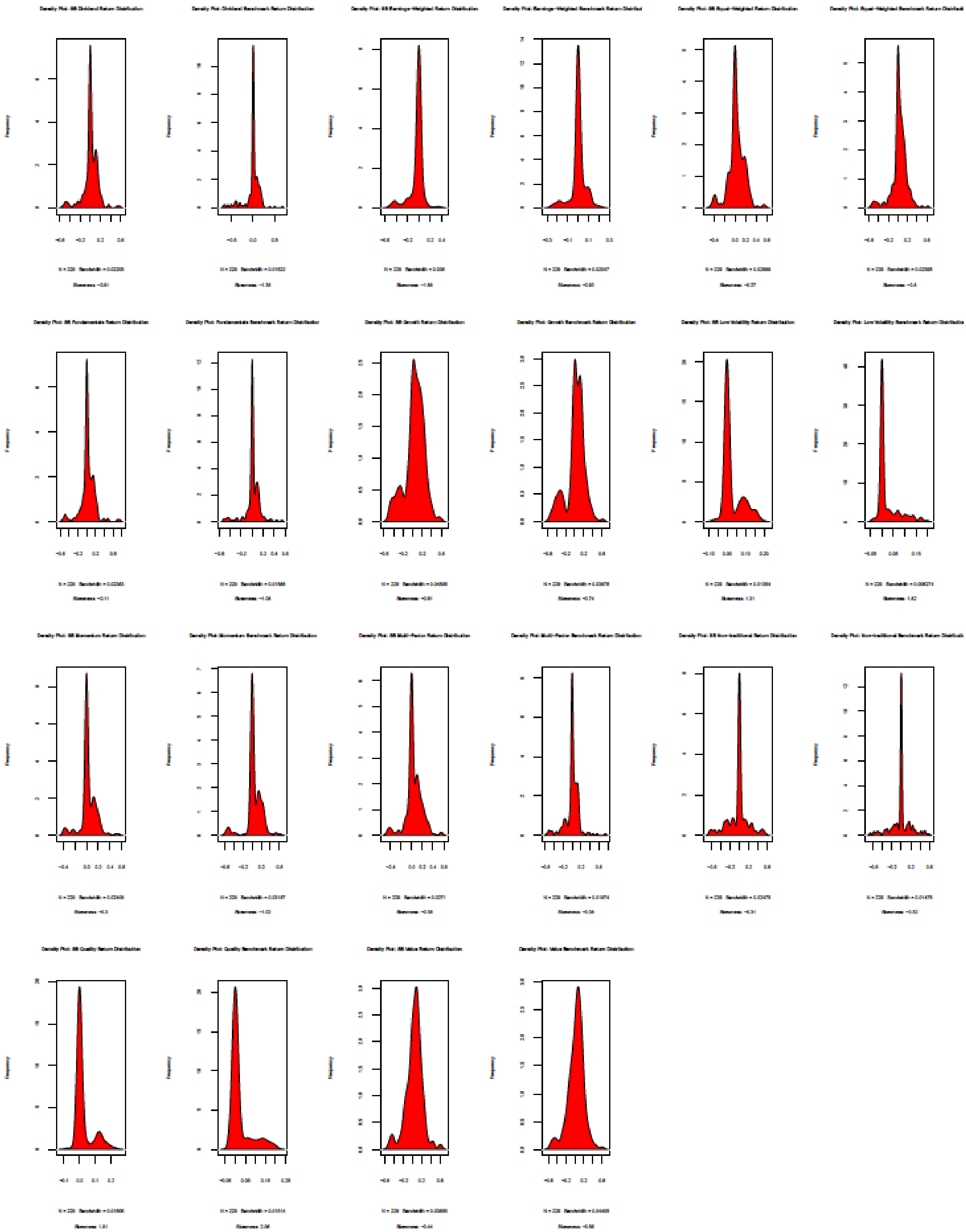


Residuals vs Fitted Plot for each equally-weighted SB ETF category Multiple Regression (Macroeconomic Model)

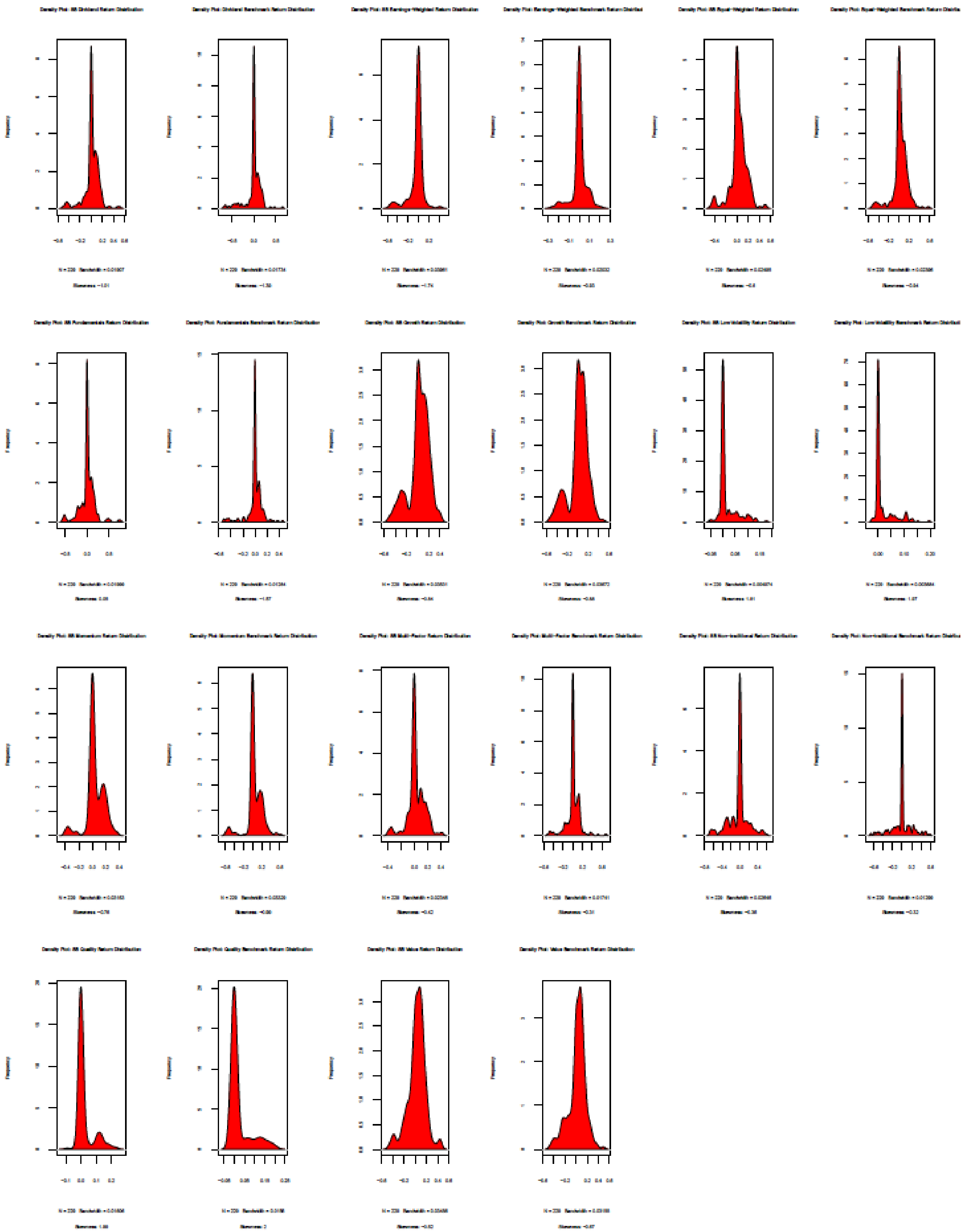


Residuals vs Fitted Plot for each size-weighted SB ETF category Multiple Regression (Macroeconomic Model)

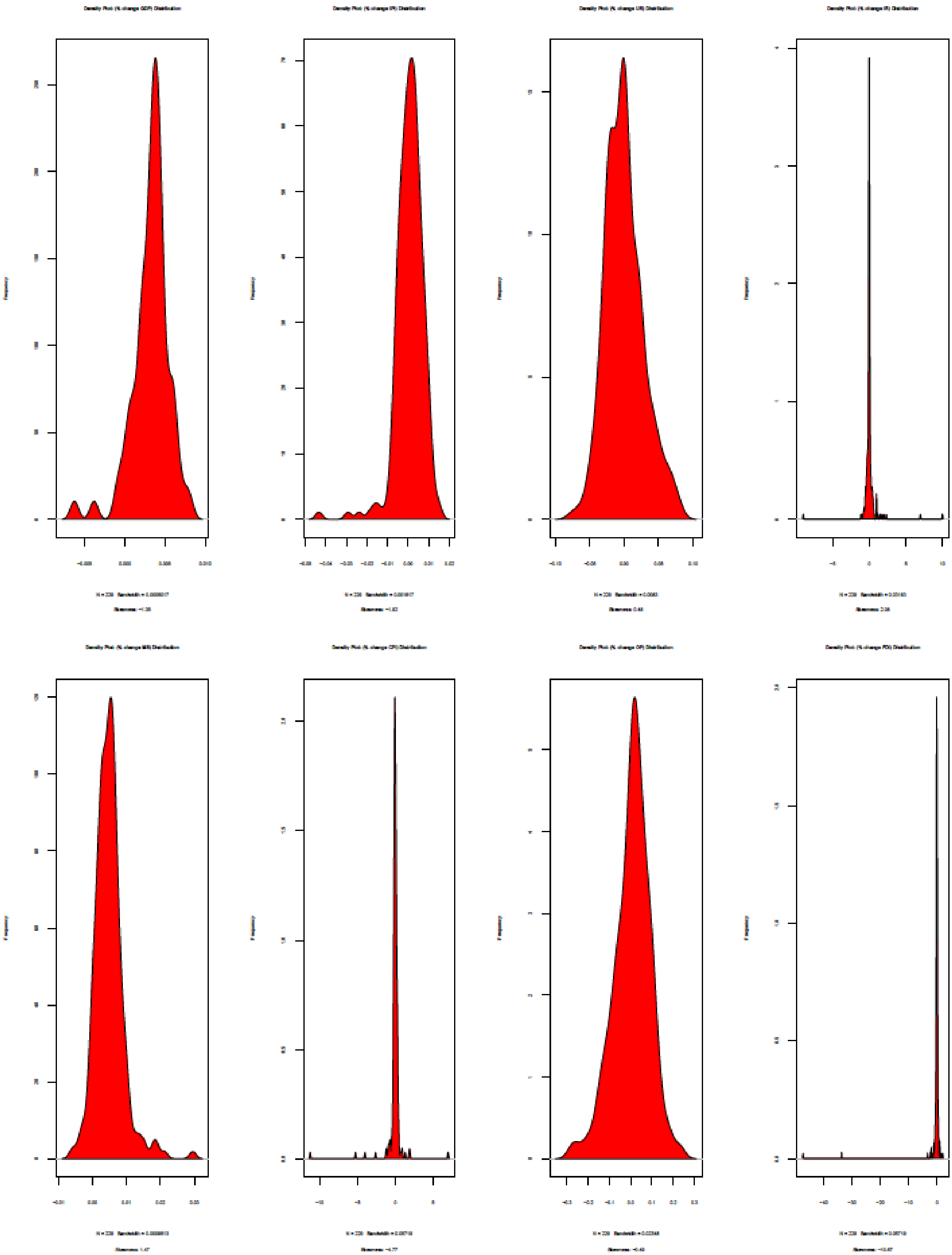
Appendix IV: Normality Assumption



Density plots for equally-weighted SB ETFs per category and declared benchmark



Density plots for size-weighted SB ETFs per category and declared benchmark

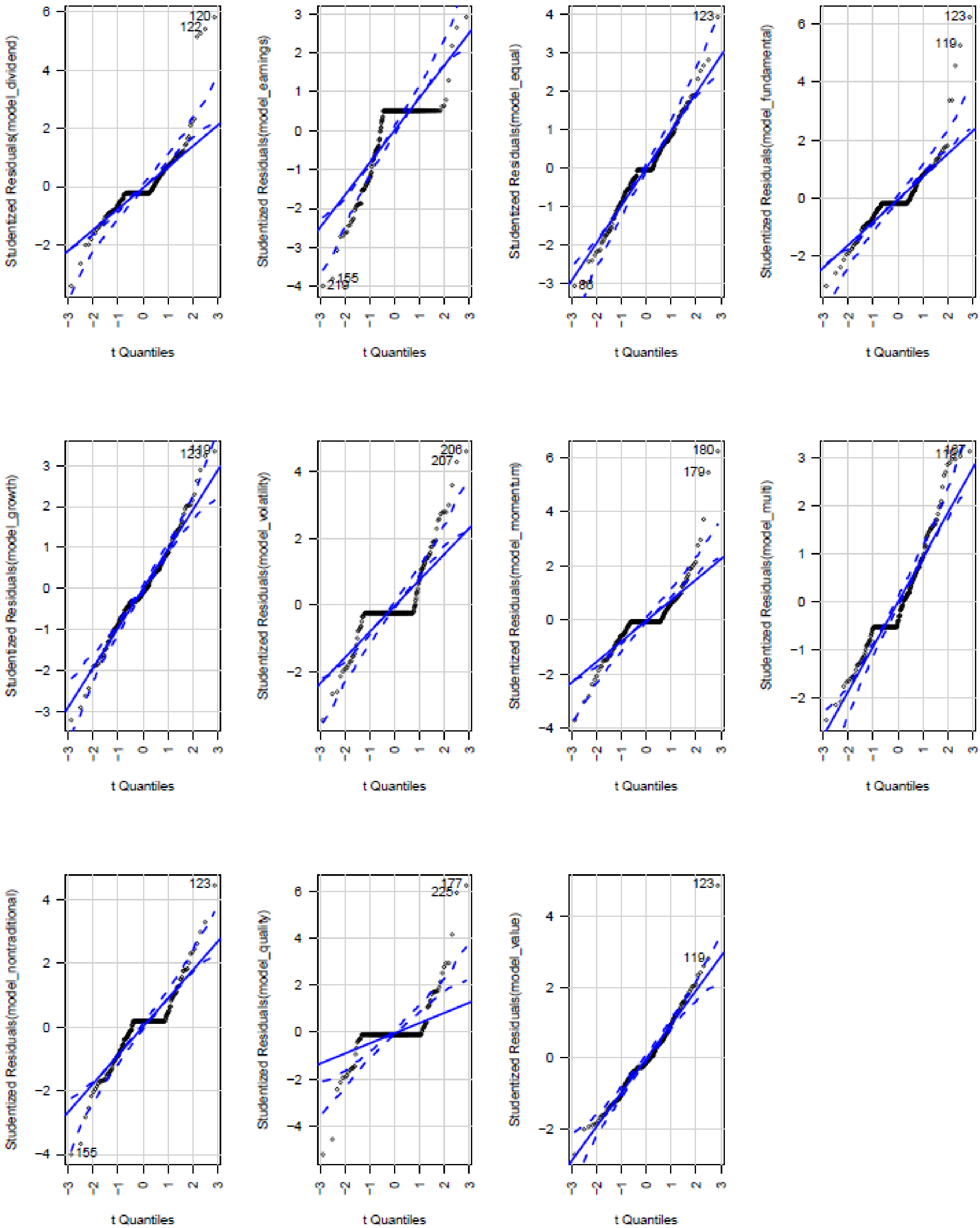


Density Plots for explanatory macroeconomic variables

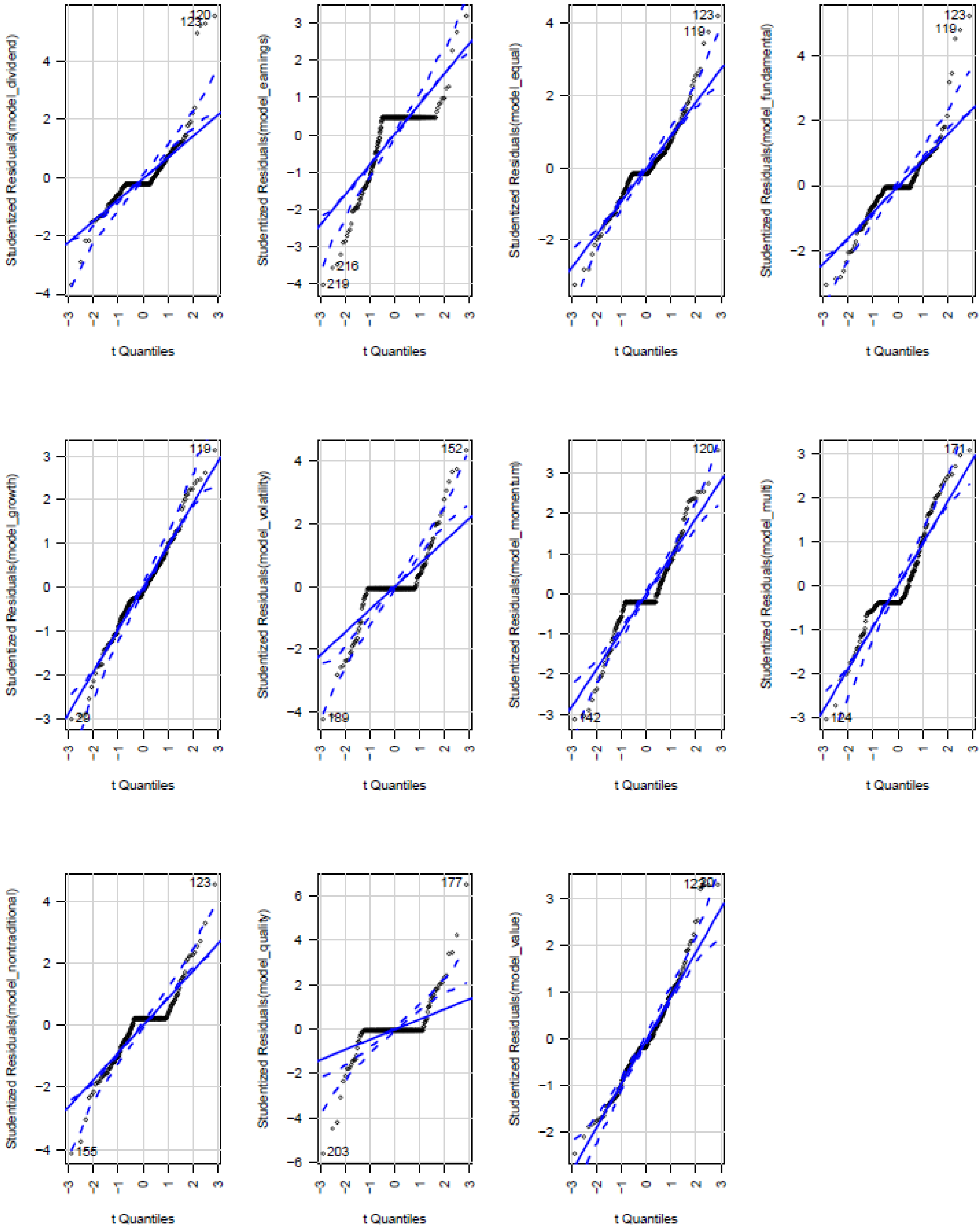
SB ETF Category	W statistic (equal-weighted)	p-value	W statistic (size-weighted)	p-value
Dividend	0.85321 (***)	5.687e-14	0.84064 (***)	1.328e-14
Earnings-Weighted	0.63578 (***)	2.2e-16	0.65076 (***)	2.2e-16
Equal-Weighted	0.93773 (***)	2.706e-08	0.91582 (***)	4.236e-10
Fundamentals	0.85387 (***)	6.156e-14	0.85901 (***)	1.147e-13
Growth	0.95712 (***)	2.401e-06	0.93616 (***)	1.953e-08
Low volatility	0.69613 (***)	2.2e-16	0.68015 (***)	2.2e-16
Value	0.95613 (***)	1.866e-06	0.95947 (***)	4.443e-06
Momentum	0.84358 (***)	1.851e-14	0.84953 (***)	3.684e-14
Multi-Factor	0.90139 (***)	3.959e-11	0.90285 (***)	4.975e-11
Non-traditional	0.89287 (***)	1.084e-11	0.89035 (***)	7.495e-12
Quality	0.56188 (***)	2.2e-16	0.5557 (***)	2.2e-16

*Significance Codes: *** (<0.01); ** (<0.05); * (<0.1)*

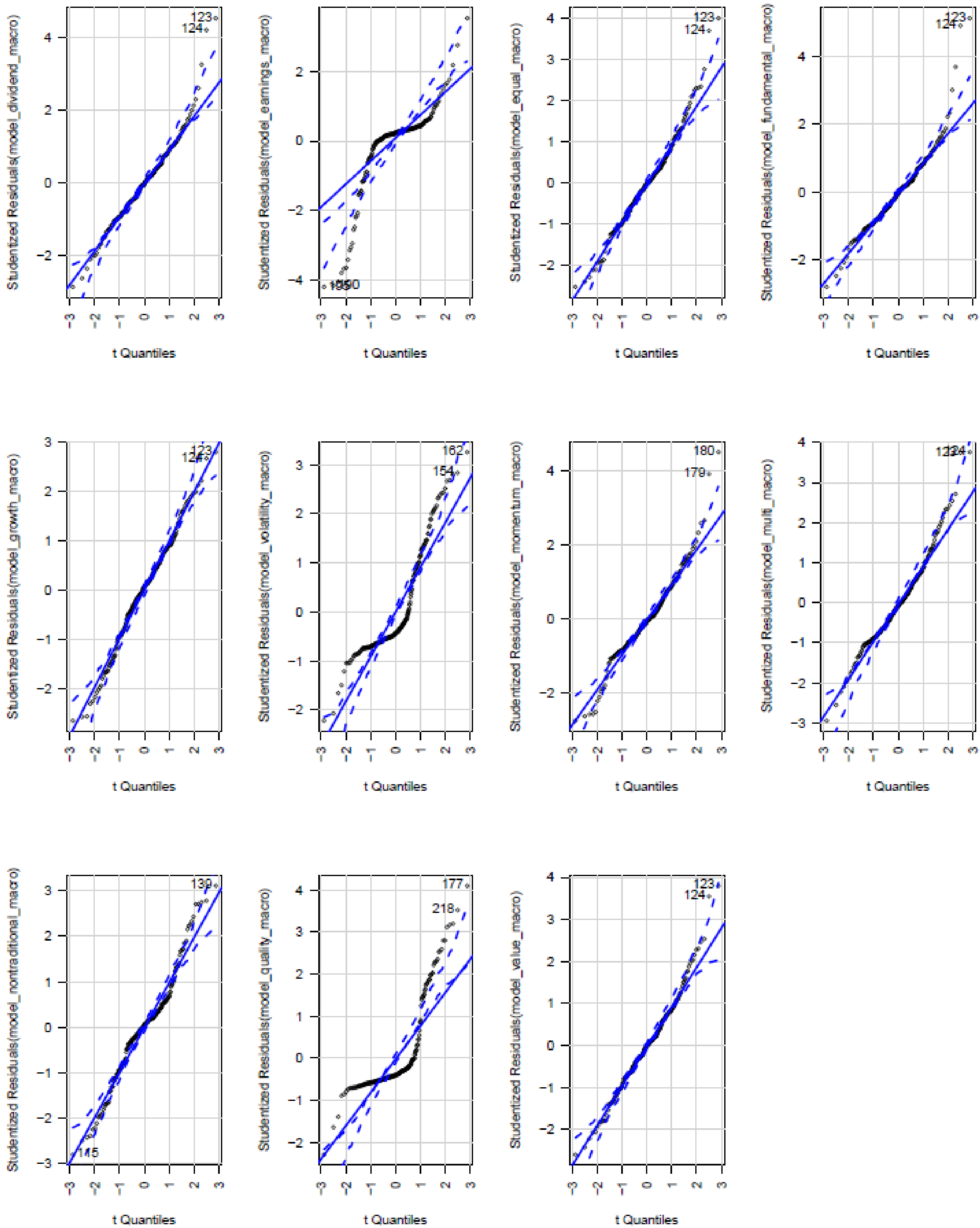
Shapiro-Wilk test to test stationarity where null hypothesis of this test is that population is normally distributed



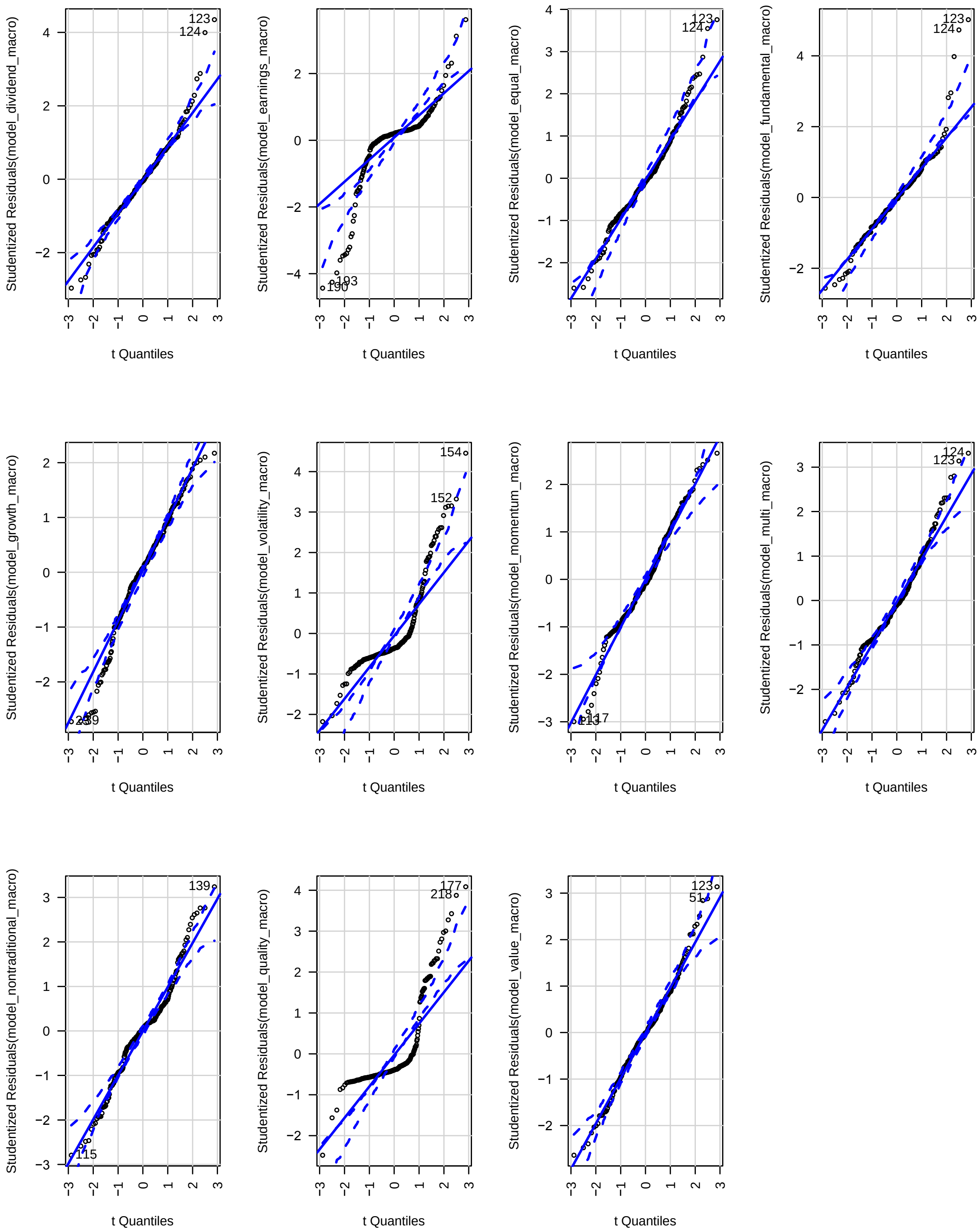
Q-Q Plots for equally-weighted SB ETF
(Simple Regressions)



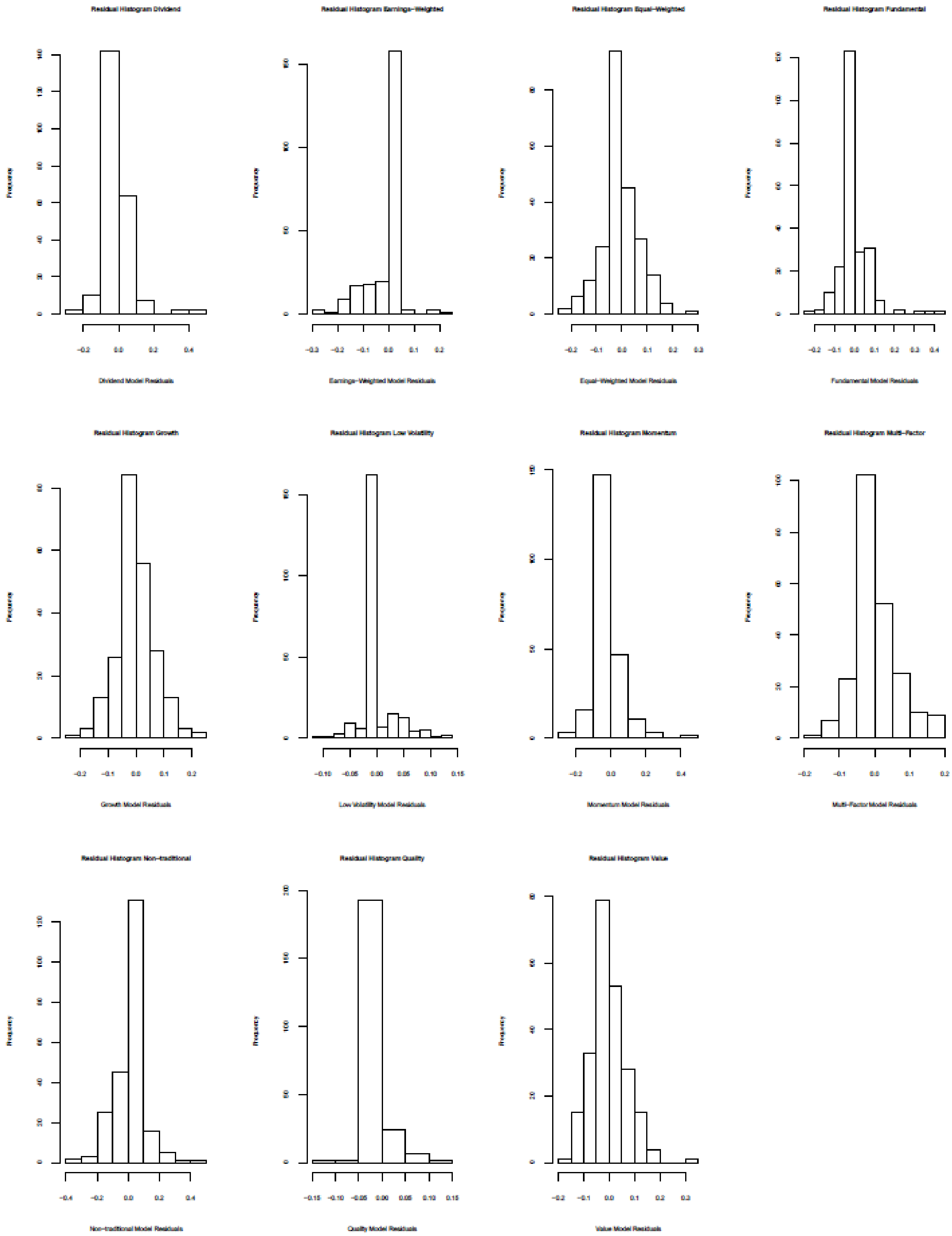
Q-Q Plots for size-weighted SB ETF
(Simple Regressions)



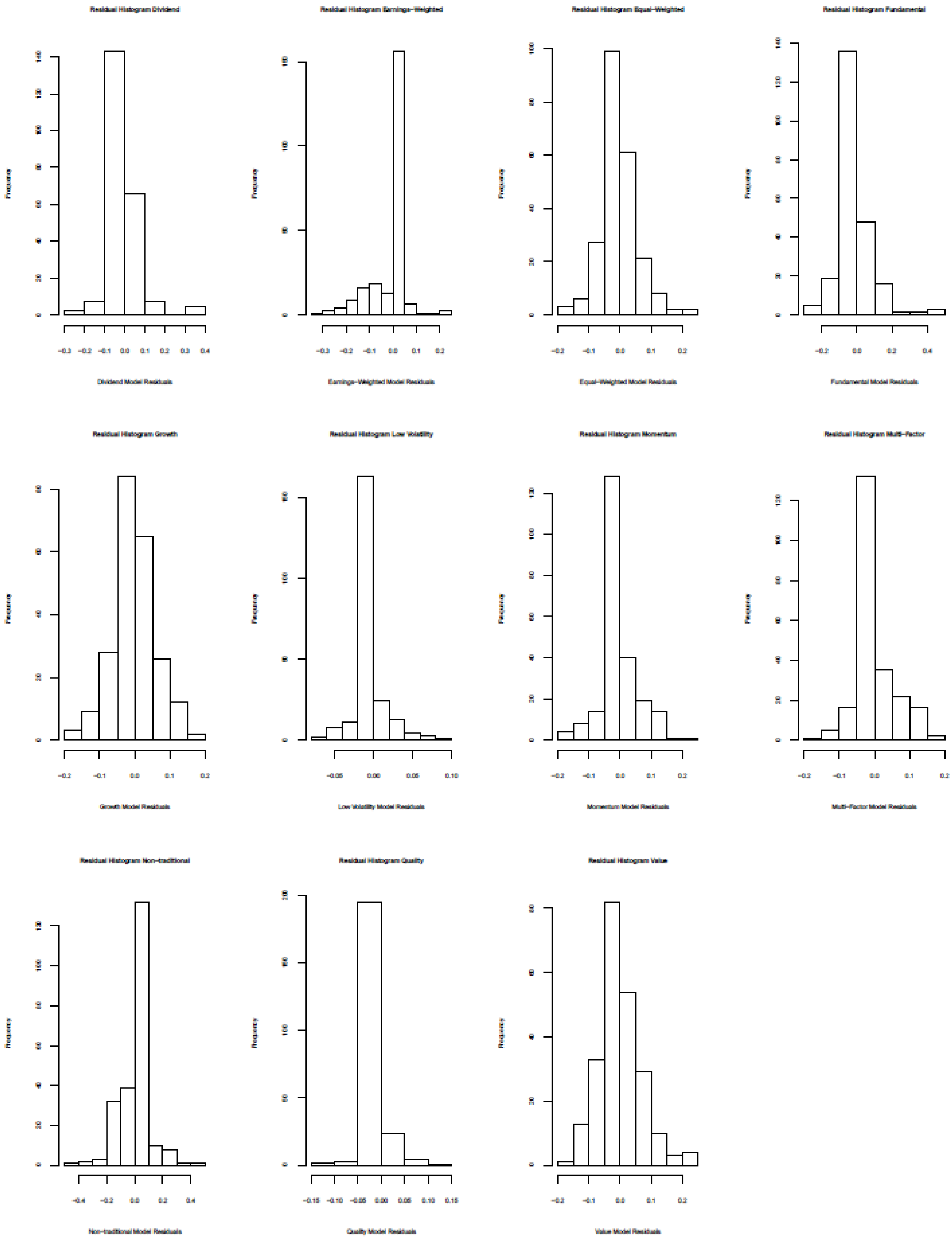
Q-Q Plot for each equally-weighted SB ETF category (Macroeconomic Model)



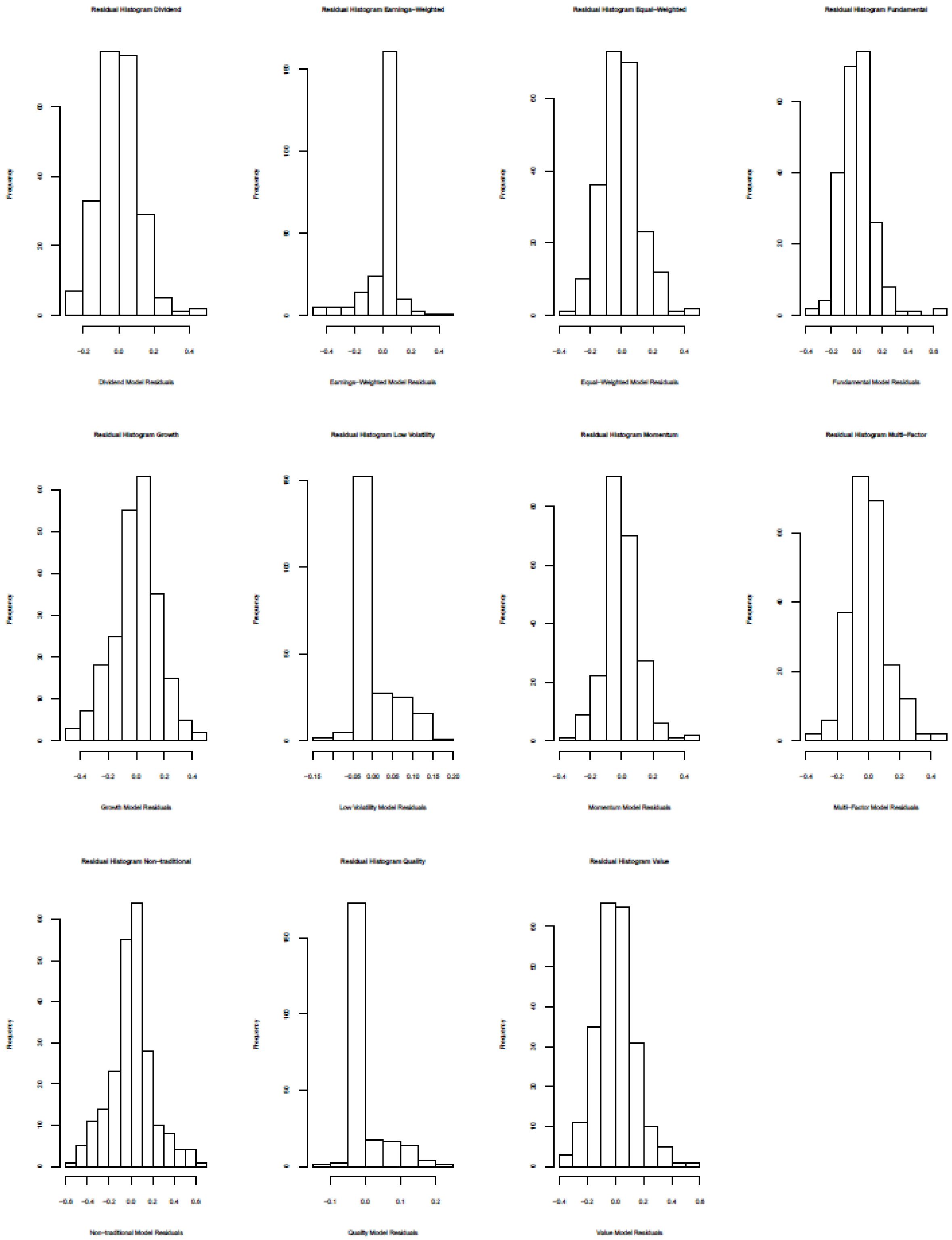
Q-Q Plot for each equally-weighted SB ETF category (Macroeconomic Model)



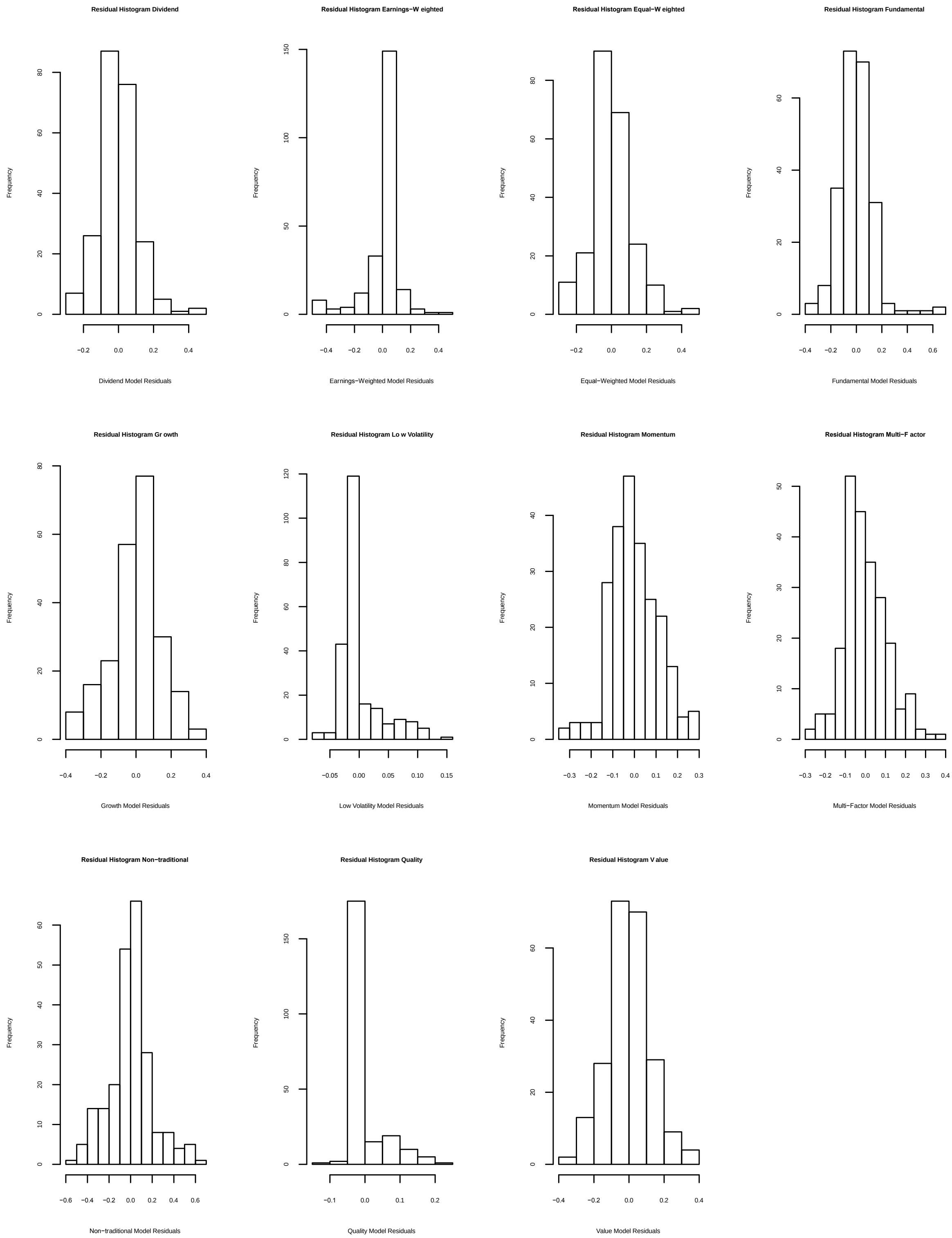
Histogram of Residuals for each equally-weighted SB ETF
(Simple Regressions)



Histogram of Residuals for each size-weighted SB ETF
(Simple Regressions)



Residual Histogram for Multiple Regressions for each equally-weighted SB ETF category (Macroeconomic Model)

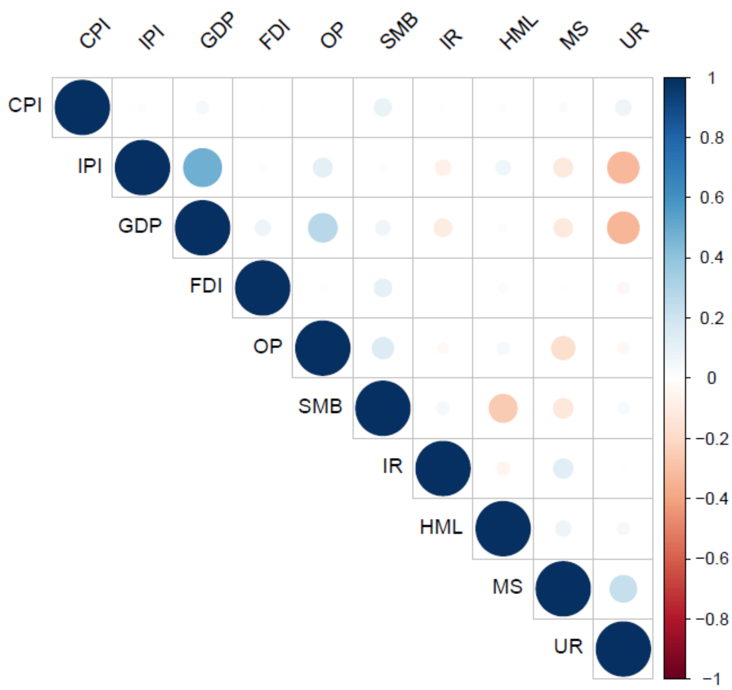


Residual Histogram for Multiple Regressions for each size-weighted SB ETF category (Macroeconomic Model)

Appendix V: Multicollinearity Assumption

	IPI	*MS*	*IR*	*OP*	*UR*	*CPI*	*FDI*	*GDP*	*SMB*	*HML*
IPI	*1*	*-0.12*	*-0.08*	*0.11*	*-0.33*	*0.02*	*0.11*	*0.48*	*0.02*	*0.07*
MS	*-0.12*	*1*	*0.12*	*-0.18*	*0.24*	*0.02*	*0.01*	*-0.11*	*-0.13*	*0.08*
IR	*-0.08*	*0.12*	*1*	*-0.03*	*0.01*	*0.01*	*-0.06*	*-0.10*	*0.05*	*-0.05*
OP	*0.11*	*-0.18*	*-0.03*	*1*	*-0.04*	*-0.01*	*0.01*	*0.27*	*0.15*	*0.05*
UR	*-0.33*	*0.24*	*0.01*	*-0.04*	*1*	*0.08*	*-0.04*	*-0.33*	*0.04*	*0.04*
CPI	*0.02*	*0.02*	*0.01*	*-0.01*	*0.08*	*1*	*-0.27*	*0.05*	*0.09*	*-0.01*
FDI	*0.11*	*0.01*	*-0.06*	*0.01*	*-0.04*	*-0.27*	*1*	*0.05*	*-0.01*	*0.14*
GDP	*0.48*	*-0.11*	*-0.10*	*0.27*	*-0.33*	*0.05*	*0.05*	*1*	*0.07*	*0.02*
SMB	*0.02*	*-0.13*	*0.05*	*0.15*	*0.04*	*0.09*	*-0.01*	*0.07*	*1*	*-0.26*
HML	*0.07*	*0.08*	*-0.05*	*0.05*	*0.04*	*-0.01*	*0.14*	*0.02*	*-0.26*	*1*

Correlation Matrix for independent variables in Multiple Regression

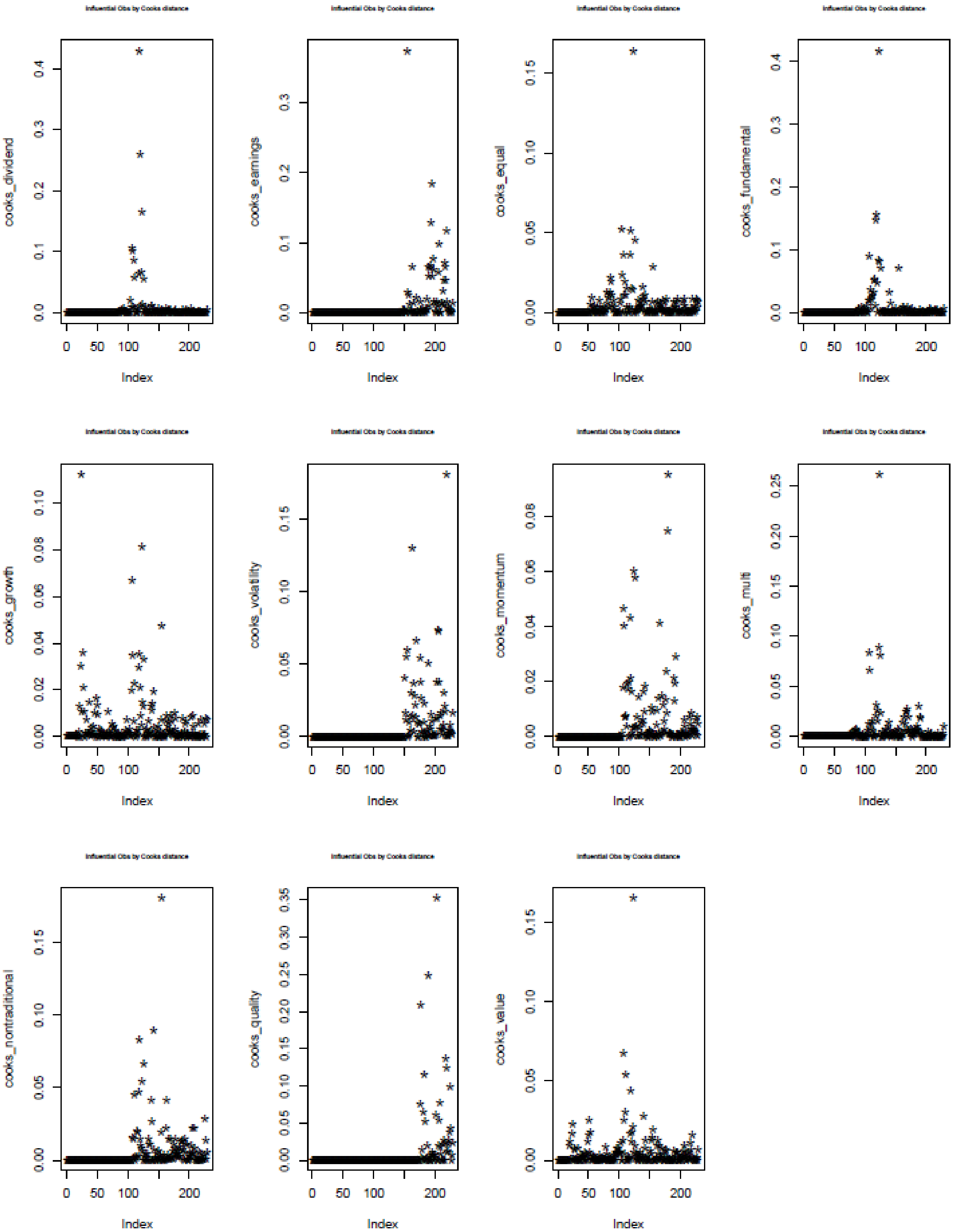


Correlation plot for independent variables in Multiple Regression

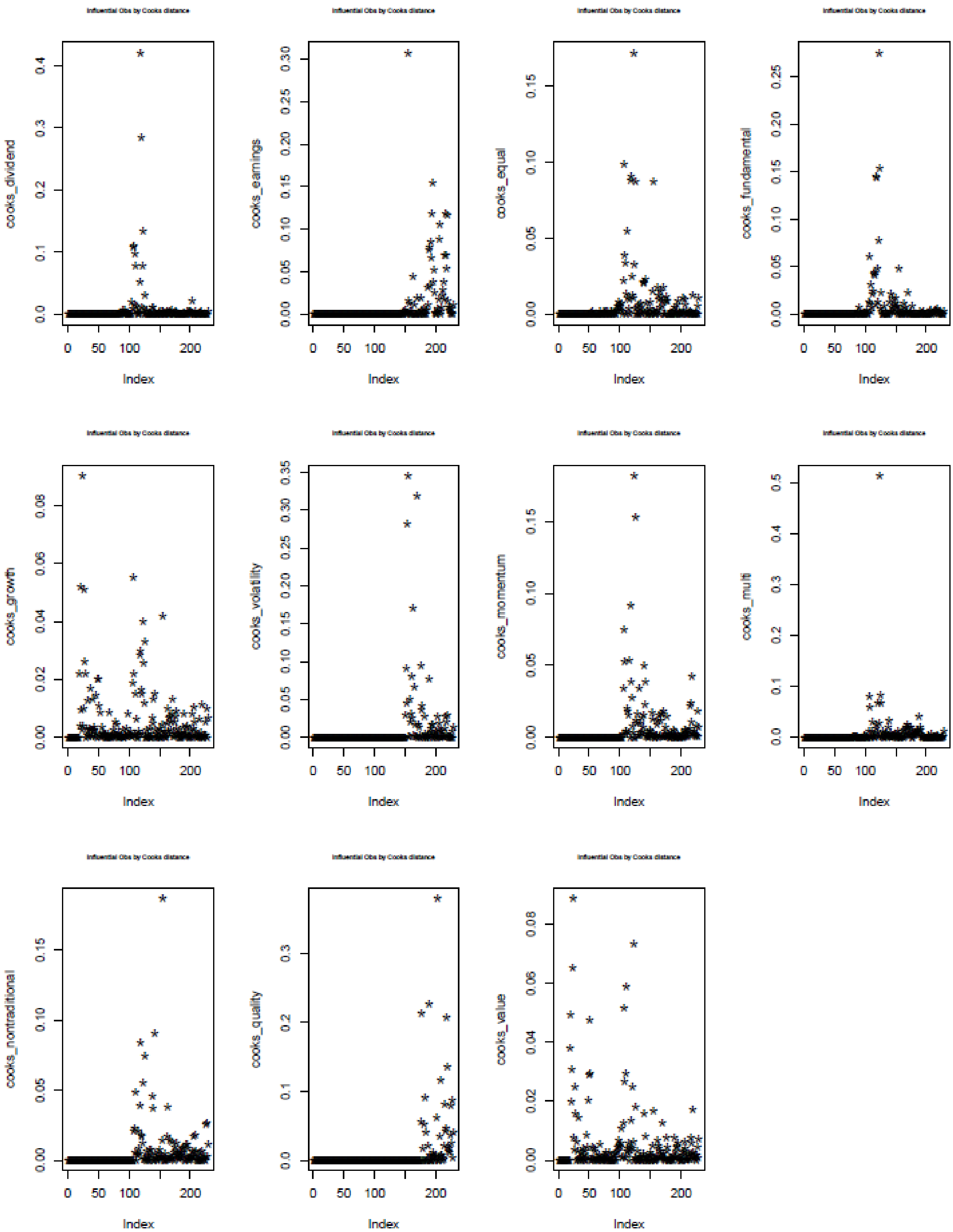
Dependent Variable	Independent Variable									
Variable	*GDP*	*IPI*	*UR*	*IR*	*MS*	*CPI*	*OP*	*FDI*	*SMB*	*HML*
SB ETF Excess Return (Equally-weighted; macroeconomic model)	*1.48*	*1.38*	*1.27*	*1.03*	*1.12*	*1.10*	*1.13*	*1.10*		
SB ETF Excess Return (Equally-weighted; complete model	*1.49*	*1.39*	*1.31*	*1.04*	*1.16*	*1.10*	*1.13*	*1.13*	*1.11*	*1.18*
SB ETF Excess Return (Size-weighted; macroeconomic model)	*1.48*	*1.38*	*1.27*	*1.03*	*1.12*	*1.10*	*1.13*	*1.10*		
SB ETF Excess Return (Size-weighted; complete model	*1.49*	*1.39*	*1.31*	*1.04*	*1.16*	*1.10*	*1.13*	*1.13*	*1.11*	*1.18*

Variance Inflation Factor for independent variables for each Multiple Regression

Appendix VI: Outliers

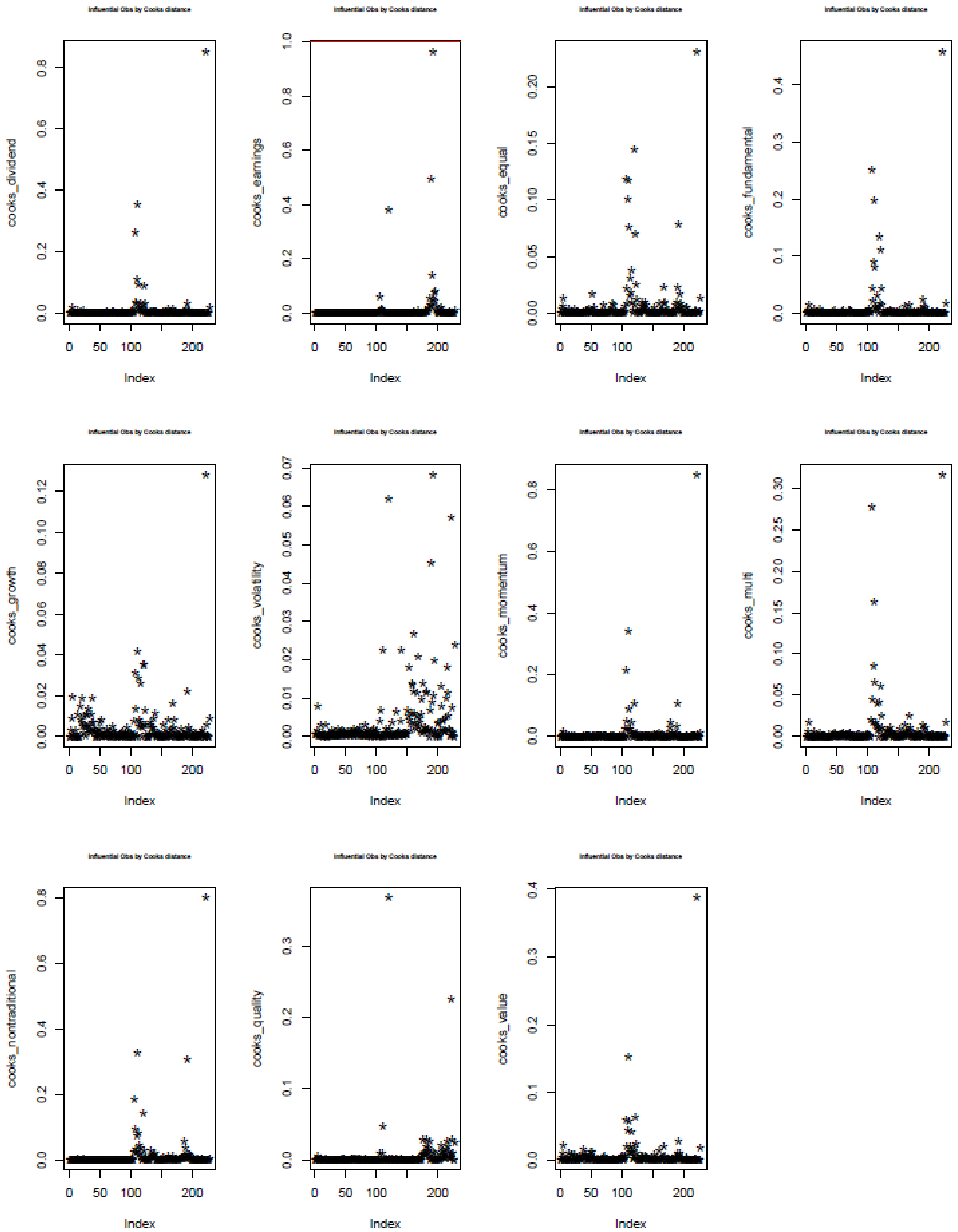


Cook's distance for equally-weighted SB ETF (Simple Regression)



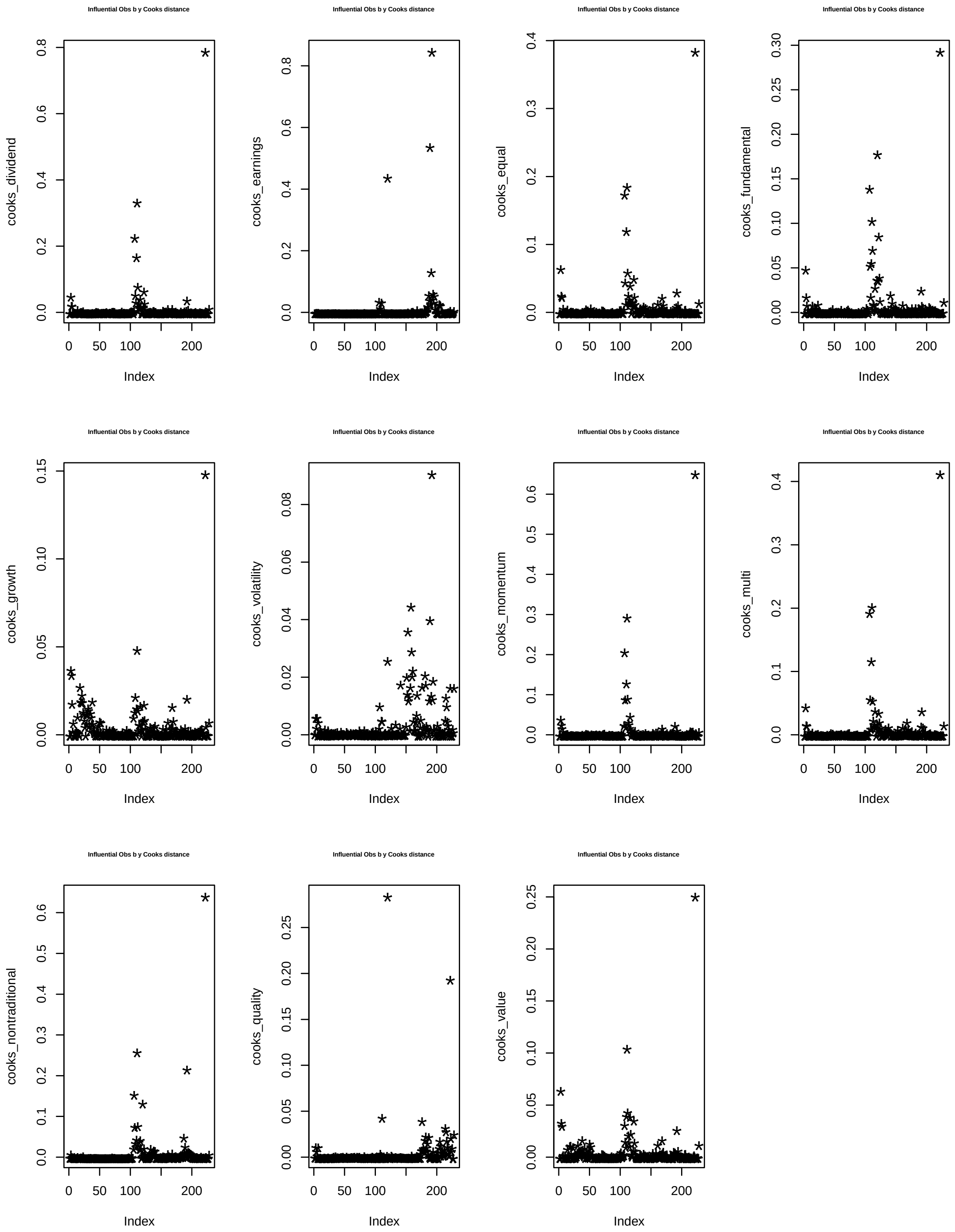
Cook's distance for size-weighted SB ETF

(Simple Regression)



Cook's Distance Plot for each equally-weighted SB ETF category

(Macroeconomic Model)



Cook's Distance Plot for each size-weighted SB ETF category

(Macroeconomic Model)

Appendix VII: Largest SB ETFs

Name	Ticker	Strategy	Benchmark	Expense Ratio	Market Value (in $bil)
Vanguard Value ETF	VTV	Value	iShares MSCI EAFE ETF	0,0005	46,25
iShares Russell 1000 Growth ETF	IWF	Growth	iShares Russell 1000 Growth ETF	0,002	41,78
iShares Russell 1000 Value ETF	IWD	Value	iShares Core S&P 500 ETF	0,002	38,28
Vanguard Growth ETF	VUG	Growth	iShares MSCI EAFE Growth ETF	0,0005	36,79
Vanguard Dividend Appreciation ETF	VIG	Dividend	Invesco High Yield Equity Dividend Achievers™ ETF	0,0008	32,40
iShares Edge MSCI Min Vol USA ETF	USMV	Low Volatility	iShares Edge MSCI Min Vol USA ETF	0,0015	23,05
Vanguard High Dividend Yield ETF	VYM	Dividend	Invesco High Yield Equity Dividend Achievers™ ETF	0,0008	22,55
iShares S&P 500 Growth ETF	IVW	Growth	iShares Russell 1000 Growth ETF	0,0018	21,98
SPDR S&P Dividend ETF	SDY	Dividend	Invesco High Yield Equity Dividend Achievers™ ETF	0,0035	18,06
iShares S&P 500 Value ETF	IVE	Value	iShares Core S&P 500 ETF	0,0018	15,30
Invesco S&P 500® Equal Weight ETF	RSP	Equal	iShares Core S&P 500 ETF	0,002	15,23
Vanguard Small Cap Value ETF	VBR	Value	iShares S&P SmallCap 600 Value ETF	0,0007	13,36
iShares Russell Midcap Value ETF	IWS	Value	iShares Core S&P 500 ETF	0,0025	11,28
iShares Edge MSCI Min Vol EAFE ETF	EFAV	Low Volatility	iShares Edge MSCI Min Vol EAFE ETF	0,002	10,74
iShares Russell Midcap Growth ETF	IWP	Growth	iShares Russell 1000 Growth ETF	0,0025	9,95
Invesco S&P 500® Low Volatility ETF	SPLV	Low Volatility	SPLV Invesco S&P 500 Low Volatility ETF	0,0025	9,52
iShares Russell 2000 Value ETF	IWN	Value	iShares Core S&P 500 ETF	0,0024	9,51
iShares Edge MSCI USA Quality Factor ETF	QUAL	Quality	Invesco S&P 500 Quality ETF	0,0015	9,47
Alerian MLP ETF	AMLP	Value	iShares MSCI EAFE ETF	0,0085	9,30

Largest 20 SB ETFs among sample

Appendix VIII: Number of dead SB ETFs per factor, between 2000 and 2018

Smart Beta category	Dividend	Earnings-Weighted	Equal-Weighted	Fundamentals	Growth	Low Volatility	Momentum	Multi-Factor	Non-traditional commodity	Quality	Value	Total
2000	0	0	0	0	0	0	0	0	0	0	0	0
2001	0	0	0	0	0	0	0	0	0	0	0	0
2002	0	0	0	0	0	0	0	0	0	0	0	0
2003	0	0	0	0	0	0	0	0	0	0	0	0
2004	0	0	0	0	0	0	0	0	0	0	0	0
2005	0	0	0	0	0	0	0	0	0	0	0	0
2006	0	0	0	0	0	0	0	0	0	0	0	0
2007	0	0	0	0	0	0	0	0	0	0	0	0
2008	0	0	0	0	0	0	0	0	0	0	0	0
2009	0	0	0	0	0	0	0	0	0	0	0	0
2010	0	0	0	0	0	0	0	0	0	0	0	0
2011	0	0	0	0	0	0	0	0	0	0	0	0
2012	0	0	0	0	0	0	0	0	0	0	0	0
2013	0	0	1	0	0	0	0	0	0	0	0	1
2014	0	0	0	0	0	0	0	0	0	0	1	1
2015	0	0	0	0	0	0	0	0	0	0	0	0
2016	1	2	1	0	0	2	0	2	0	0	0	8
2017	1	0	2	2	1	0	0	1	0	0	1	8
2018	0	0	0	2	0	0	0	7	0	0	0	9

Number of dead SB ETFs per category, between 2000 and 2018

Appendix IX: VAR (Optimal Lag Length)

Selection-order criteria
Sample: 13 - 227 Number of obs = 215

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	1445.96				2.1e-16	-13.3764	-13.3257	-13.251
1	1634.41	376.89	64	0.000	6.7e-17*	-14.534*	-14.0779*	-13.4053*
2	1698.07	127.32	64	0.000	6.8e-17	-14.5309	-13.6694	-12.3987
3	1759.73	123.33	64	0.000	7.0e-17	-14.5092	-13.2423	-11.3737
4	1804.03	88.597	64	0.023	8.5e-17	-14.3259	-12.6536	-10.1871
5	1851.57	95.075	64	0.007	1.0e-16	-14.1727	-12.0951	-9.03056
6	1898.34	93.544	64	0.009	1.2e-16	-14.0125	-11.5294	-7.86695
7	1960.57	124.46	64	0.000	1.3e-16	-13.996	-11.1075	-6.84713
8	2001.58	82.008	64	0.064	1.7e-16	-13.7821	-10.4882	-5.62986
9	2060.59	118.03	64	0.000	1.9e-16	-13.7357	-10.0364	-4.58014
10	2138.04	154.9	64	0.000	1.8e-16	-13.8609	-9.75617	-3.7019
11	2174	71.925	64	0.232	2.6e-16	-13.6	-9.08995	-2.43774
12	2240.03	132.05*	64	0.000	2.9e-16	-13.6189	-8.70338	-1.45321

Endogenous: IPI MS IR OP UR CPI FDI GDP
Exogenous: _cons

Optimal Leg Length

Augmented Dickey-Fuller test for unit root Number of obs = 225

	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
		Interpolated Dickey-Fuller		
Z(t)	-7.404	-3.468	-2.882	-2.572

MacKinnon approximate p-value for Z(t) = 0.0000

Test for Stationarity

Eigenvalue stability condition

Eigenvalue	Modulus
.8147205	.81472
.2495151 + .1955893*i*	.317038
.2495151 - .1955893*i*	.317038
-.2724785 + .02249533*i*	.273406
-.2724785 - .02249533*i*	.273406
-.03596799 + .03423632*i*	.049657
-.03596799 - .03423632*i*	.049657
.04709805	.047098

All the eigenvalues lie inside the unit circle.
VAR satisfies stability condition.

Test for Stability

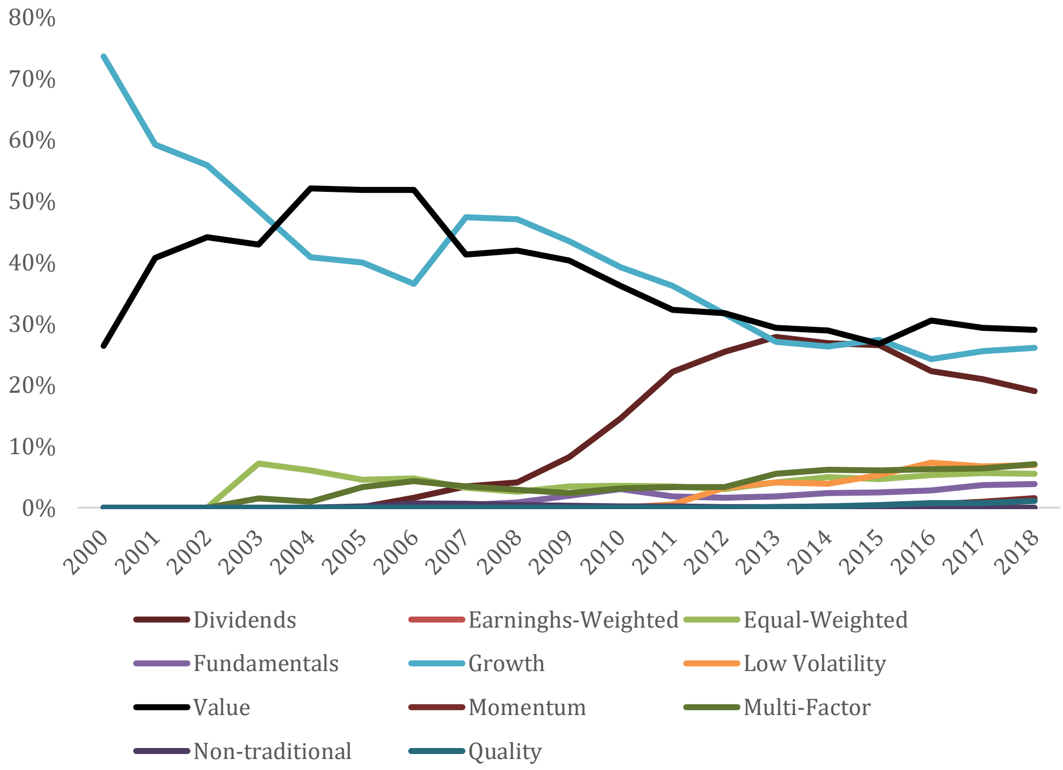
Lagrange-multiplier test

lag	chi2	df	Prob > chi2
1	108.4560	64	0.00044
2	94.3819	64	0.00806
3	116.2438	64	0.00007
4	56.6485	64	0.73136

H0: no autocorrelation at lag order

Test for No-Autocorrelation in Residuals (Lagrange-Multiplier Test)

Appendix X: Relative weight of active SB ETF’s per category (measured in AUM) between 2000 and 2018



Relative Weight of active SB ETFs per category (measured in AUM), between 2000 and 2018

Appendix XI: Results Multi-Factor Models (Complete Model)

Smart Beta category	Intercept	GDP	IPI	UR	IR	MS	CPI	OP	FDI	SMB	HML	R^2
Dividend	-0.04	18.89 (*)	6.05 (***)	-0.8334 (**)	-0.01	-2.26	-0.01 (**)	-0.01	0.00	0.01	-0.00	0.3470
Earnings-Weighted	-0.05	3.44	1.24	0.5009	0.00	0.97	0.0079	0.12	0.00	0.01	0.01	-0.0002
Equal-Weighted	-0.02	18.36 (**)	4.35 (**)	-0.8063 (**)	-0.01	-1.25	-0.01 (*)	0.07	0.00 (*)	0.00	-0.01	0.2481
Fundamentals	-0.02	18.17 (*)	7.01 (***)	-0.7812 (*)	-0.01	-3.82	-0.02 (***)	0.02	0.00	0.00	-0.00	0.3216
Growth	-0.02	26.07 (***)	5.05 (**)	-1.7296 (***)	-0.01	-5.05 (*)	-0.02 (***)	0.03	-0.00	-0.00	-0.00 (*)	0.3157
Low Volatility	0.03 (*)	-0.45	0.28	-0.4284 (**)	-0.00	0.52	0.00	-0.00	-0.00	-0.00	-0.00	0.0090
Value	-0.03	0.29 (***)	0.50 (**)	-1.1948 (***)	-0.01	-3.79	-0.01 (**)	-0.01	0.00	0.00	-0.00	0.3289
Momentum	-0.00	0.13	3.88 (***)	-0.9721 (**)	-0.01 (**)	-1.81	-0.01	0.01	-0.00	0.00	-0.01	0.2348
Multi-Factor	0.00	14.15	5.57 (***)	-1.0926 (***)	-0.01	-2.20	-0.02 (***)	0.02	0.00	0.00	-0.01	0.2768
Non-traditional	-0.09	17.71	4.34	-0.5091	-0.02	-0.16	-0.01	0.14	0.00	0.00	0.01	0.0003
Quality	0.03 (*)	1.10	-0.23	-0.3492 (**)	-0.00	-0.80	-0.00	-0.01	-0.00 (*)	-0.00	-0.00 (*)	0.0273

*Significance Codes: *** (<0.01); ** (<0.05); * (<0.1)*

Multiple Regression of equal-weighted domestic equity Smart Beta ETFs vs. 10 factors by category, Jan 2000 – Dec 2018 (Complete Model)

Appendix XII: Results Multi-Factor Models (Complete Model)

Smart Beta category	Intercept	GDP	IPI	UR	IR	MS	CPI	OP	FDI	SMB	HML	R^2
Dividend	-0.03	17.07 (*)	5.50 (***)	-0.90 (**)	-0.01 (***)	-2.12	-0.01 (***)	-0.03	0.01	0.01	-0.01	0.3547
Earnings-Weighted	-0.06	3.11	1.37	0.52	0.01	1.70	0.01	0.16	0.01	0.01	0.01	0.0105
Equal-Weighted	-0.01	17.99 (*)	4.89 (**)	-1.05 (***)	-0.01	-1.57	-0.01 (***)	0.01	0.01	0.01	-0.01	0.3070
Fundamentals	-0.04	22.24 (**)	6.80 (**)	-0.37	-0.01	-6.57 (*)	-0.03 (***)	-0.01	0.01	0.01	-0.01	0.3183
Growth	-0.02	23.17 (***)	3.95 (**)	-1.69 (***)	-0.01	-4.59 (**)	-0.01 (***)	0.01	-0.00	-0.00	-0.01	0.3062
Low Volatility	0.02	-0.60	0.19	-0.26 (*)	-0.00	0.52	0.00	-0.01	0.00	0.00	-0.00	-0.0087
Value	-0.03	23.87 (***)	3.73 (**)	-1.07 (***)	-0.01	-2.80	-0.01 (*)	-0.04	0.00	0.00	0.00	0.3159
Momentum	0.01	12.02	3.80 (***)	-1.18 (***)	-0.01	-1.66	-0.01 (***)	0.02	-0.00	-0.00	-0.01	0.2582
Multi-Factor	-0.01	12.63	4.48 (***)	-0.92 (***)	-0.01 (**)	-1.54	-0.01 (**)	0.02	0.00	0.00	-0.00	0.2594
Non-traditional	-0.10	19.42	4.22	-0.51	-0.02	-0.78	-0.02	0.14	-0.00	0.00	0.00	0.1051
Quality	0.02 (*)	1.11	-0.20	-0.34 (**)	-0.00		-0.82	0.02	-0.00	-0.00	-0.00	0.0224

*Significance Codes: *** (<0.01); ** (<0.05); * (<0.1)*

Multiple Regression of size-weighted domestic equity Smart Beta ETFs vs. 10 factors by category, Jan 2000 – Dec 2018
(Complete Model)

Appendix XIII: Results Step-wise Multi-Factor Models

Smart Beta category	Intercept	GDP	IPI	UR	IR	MS	CPI	OP	FDI	R^2
Dividend	-0.05	19.11 (*)	6.21 (***)	-0.85 (**)	-0.0114 (*)		-0.02 (**)			0.3538
Earnings-Weighted										
Equal-Weighted	-0.03	19.96 (**)	4.61 (**)	-0.73 (**)			-0.01 (**)			0.2524
Fundamentals	-0.04	18.60 (*)	7.15 (***)	-0.84 (**)	-0.01 (*)		-0.02 (***)			0.3244
Growth	-0.05	26.49 (***)	5.37 (**)	-1.74 (***)			-0.02 (***)			0.3115
Low Volatility	0.03 (**)			-0.39 (**)						0.0385
Value	-0.05	26.63 (***)	5.20 (**)	-1.27 (***)			-0.01 (**)			0.3314
Momentum	0.03 (*)		5.86 (***)	-1.17 (**)	-0.01 (**)					0.2007
Multi-Factor	0.04 (*)		7.76 (***)	-1.31 (**)	-0.01 (*)		-0.02 (***)			0.2493
Non-traditional	-0.12	28.71 (*)								0.0866
Quality	0.02 (*)			-0.35 (*)					-0.01 (*)	0.0307

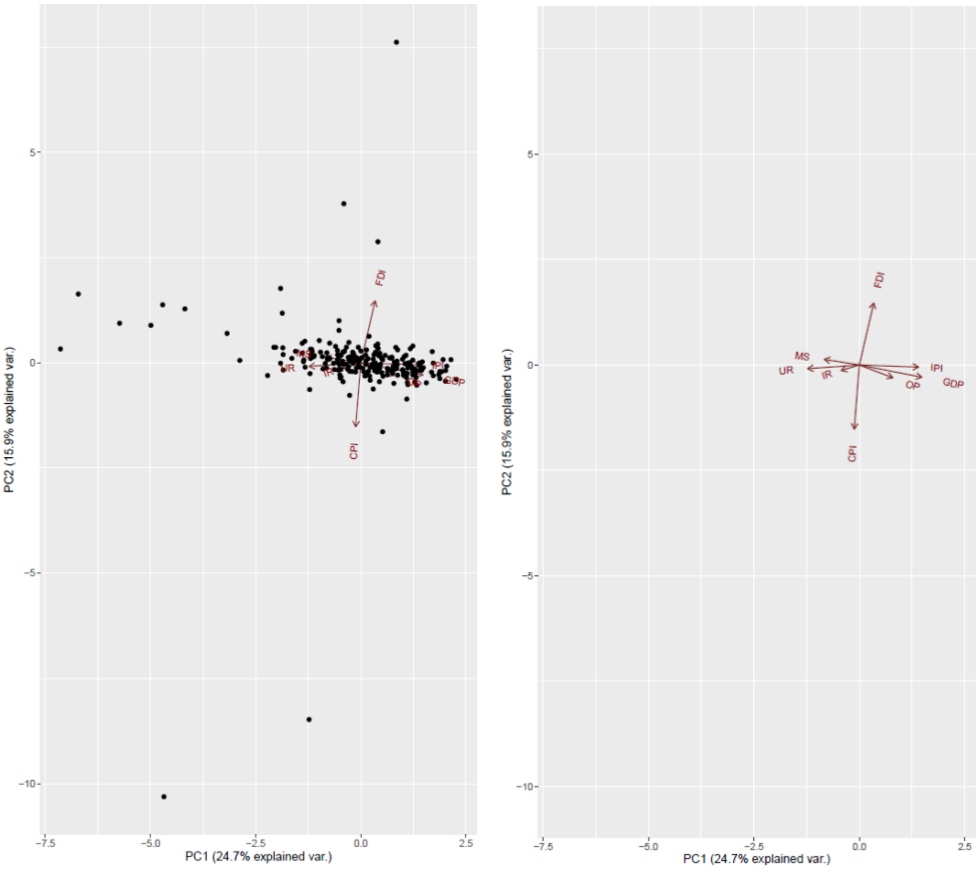
Equal-weighted domestic equity Smart Beta ETFs vs. 8 macroeconomic factors by category, Jan 2000 – Dec 2018
(Step-wise Model)

Smart Beta category	Intercept	GDP	IPI	UR	IR	MS	CPI	OP	FDI	R^2
Dividend	-0.04	16.92 (*)	5.70 (***)	-0.91 (**)	-0.01 (*)		-0.01 (**)			0.3599
Earnings-Weighted										
Equal-Weighted	-0.02	18.14 (*)	5.06 (***)	-1.02 (***)	-0-01 (*)		-0.01 (***)			0.3148
Fundamentals	-0.04	23.43 (**)	7.14 (***)			-6.99 (*)	-0.03 (***)			0.3285
Growth	-0.03	22.96 (***)	4.19 (**)	-1.57 (***)		-3.80 (*)	-0.02 (***)			0.3052
Low Volatility	0.02 (*)			-0.22 (**)						0.0223
Value	-0.05	24.23 (***)	3.91 (**)	-1.13 (***)			-0.01 (*)			0.3195
Momentum	0.04 (**)		5.64 (***)	-1.33 (***)	-0.01 (**)		-0.01 (***)			0.2266
Multi-Factor	0.02		6.55 (***)	-1.10 (**)			-0.01 (**)			0.2270
Non-traditional	-0.13	30.15 (*)								0.0925
Quality	0.02 (*)			-0.34 (*)					-0.00 (*)	0.0258

*Significance Codes: *** (<0.01); ** (<0.05); * (<0.1)*

Size-weighted domestic equity Smart Beta ETFs vs. 8 macroeconomic factors by category, Jan 2000 – Dec 2018
(Step-wise Model)

Appendix XIV: Results Principal Component Analysis



Principal Component Analysis Plot of the first two Principal Components including and excluding data points