Trade-offs of ETFs

- An Examination of Clean and Dirty Exchange Traded Funds in the Energy Sector

Hasselsjö, Mattias
Tang, Shaicoan

2016-06-06

Bachelor thesis in Financial Economics
The Department of Economics
Supervised by Charles Nadeau

UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Abstract

The aim of this thesis is to investigate if there is a difference in performance between clean and dirty exchange traded funds (ETFs) during the examination period January 2011–March 2016. Dirty ETFs are defined as ETFs that allocate in non-environmentally friendly industries such as oil or coal industries. Clean ETFs are defined as ETFs that allocate in alternative energy, for example wind or solar power industries. Two portfolios consisting of clean and dirty ETFs respectively are created using a matched pair approach controlling for size and age effects. By applying the Carhart (1997) four-factor model the market, book-to-market ratio and stock price momentum are also controlled for. In addition, the performance measures Sharpe ratio, Treynor ratio and Jensen's alpha are also employed and examined. The results suggest both that there are no statistically significant differences in performance between the clean and dirty ETF portfolios, and that the clean ETF portfolio does not perform worse than its counterpart. Different factors influence the two portfolios differently. For investors seeking ways to access opportunities in sustainable investing, the results could therefore be of much interest.

JEL classification: G11

Keywords: performance evaluation, exchange traded funds, sustainable investment, responsible investment

---

Acknowledgements
Primarily the authors would like to thank Senior Lecturer Charles Nadeau for supervising this thesis. Additionally, the econometric and statistical insights and help from Adam Farago, Joakim Ruist and Mattias Sundén have also been of great value and appreciated.

Authors' Contact Details
Mattias Hasselsjö (mattias.hasselsjo@gmail.com)
Shaicoan Tang (shaicoan.tang@gmail.com)

Table of Contents

Introduction

Purpose and Contribution

The purpose of this thesis is to contribute with an extended understanding of exchange traded funds (ETFs) relative to financial theory, and to create an understanding for clean ETFs from the perspective of an investor with a willingness to invest in sustainable financial instruments. In general, not much academic research has been done on sustainable, or “clean”, ETFs. This thesis will contribute to previous research with new results by investigating the risk-adjusted return of clean ETFs and dirty ETFs through the Sharpe ratio, Treynor ratio, Jensen’s alpha and the Carhart (1997) four-factor model. Furthermore, this research will be based on the latest data available. More specifically, based on data for the examination period January 2011–March 2016.

Background

The ongoing discussion and the increased interest in ETFs have got many investors from around the world interested in this financial instrument (Rosella and Pugliese, 2006). In 2013, 2300 billion dollars were invested in ETFs (Morningstar).

An ETF is essentially a portfolio of shares that can be bought or sold together as a unit. This financial instrument was introduced as recent as in 1993 and has its basis from mutual funds. ETFs enable investors to assemble a portfolio, often to a lower cost than mutual funds, covering a wide range of assets. Also, while mutual funds can only be traded at the end of the day, ETFs can be traded throughout the day (Bodie et al., 2014). These units are therefore traded on financial markets much like ordinary shares. One of the features that is attracting investors is the possibility to efficiently tailor the risk in a portfolio using ETFs (Gastineau, 2010).

As a result of the climate change, responsible investing has increased among investors since such investments may not only offer solutions to these kinds of environmental challenges, but also generate positive financial returns (UNPRI, 2012). United Nations Principles for Responsible Investment (UNPRI) defines responsible investments as “... an approach to investing that aims to incorporate environmental, social and governance (ESG) factors into investment decisions, to better manage risk and generate sustainable, long-term returns.”. Examples of investments that falls under the category environmental and social are those that generate positive environmental impacts along with positive financial returns, more precisely investments in such as renewable energy (UNPRI). Renewable energy includes energy generated by, for example, wind, water and sun.

Research Question

The main research question of this thesis is to examine whether there is a difference in the risk-adjusted return between a clean ETF portfolio and a dirty ETF portfolio or not. This, by using the risk-adjusted performance measures Sharpe ratio, Treynor ratio, Jensen's alpha and the Carhart (1997) four-factor model. The performance measures enable us to compare the performances in one way, while the usage of the four-factor model makes it possible for us to estimate the alpha values by using OLS as estimation method. Furthermore, it allows us to examine if the alphas are significantly different from zero and if there is a difference in the risk-adjusted return of the portfolios.

Assuming that the investment universe is smaller for clean ETFs than for dirty ETFs, the performance of clean ETFs is presumptively slightly worse.

In our analysis, "clean" ETFs are defined as ETFs that have holdings in mainly renewable energy companies. When it comes to the definition of "dirty" ETFs, these are defined as ETFs that have holdings in oil companies. In addition, it is worth mentioning that not all of the selected "dirty" ETFs contain solely holdings in oil companies but also in companies operating in for example mining or coal extraction. Importantly, the overall holdings are *dirty*, more specifically invested in companies with negative impact on the environment. This thesis strives to extend the understanding of ETFs relative to financial theory, and especially what conclusions can be drawn regarding clean ETFs in contrast to dirty ETFs. This study is, to the best of our knowledge, the first to investigate the relationship in terms of performance between dirty and clean ETFs as defined in this thesis.

Hypothesis

H0: no difference in risk-adjusted return between clean ETFs and dirty ETFs

H1: a difference in risk-adjusted return between clean ETFs and dirty ETFs

Delimitations

The results in this study are limited to the specific portfolios used, and during the examination period for which the data is available. They are not necessarily expected for all clean or dirty ETFs across all time periods. Nevertheless, the results in this study provide an insight in a sector that is relatively new and unexplored.

Section Description

The remainder of this thesis is organized as follows. The next section presents previous studies of ETFs and mutual funds where some apply similar methods. Thereafter, explanations of the methods employed are presented. Then, the data and methodology used will be presented. Finally, the results, analysis and conclusion will conclude the thesis.

Literature Review

The literature review will focus on studies that examine ETFs, using the Sharpe ratio, Treynor ratio, Jensen's alpha and the Carhart (1997) four-factor model.

Initial research comparing the pre-tax and after-tax return on the largest ETF (the SPDR) with the return on the largest equity index fund (the Vanguard Index 500 fund) was done by Poterba and Shoven (2002). What they found was that both the pre- and after-tax return of the equity index is larger than the pre- and after-tax return of the ETF during the years 1994–2000. But what is important to emphasize, is that the difference in return is very small. This modest difference suggests that the returns of the ETF and the equity index fund pre- and after- tax are quite similar.

In a study by Carhart (1997), the persistence of mutual fund performance was studied by employing a four-factor model. This model is an extension of, and has its basis in, the Capital Asset Pricing Model (CAPM) and Fama-French three factor model to which a fourth momentum factor is added. The Carhart (1997) four-factor model controls for market, size, book-to-market ratio and stock price momentum. The results of the study suggest that funds with high returns one year have persistently higher expected returns the following year, but not in years thereafter.

Another research by Kreander et al. (2005) used a matched pair analysis and the Sharpe ratio, Treynor ratio and Jensen's alpha to evaluate the performance of European ethical and non-ethical funds. Their results suggest that there is no significant difference between the performance of ethical and non-ethical funds.

In 2006, Gallagher and Segara examine the performance and trading characteristics of ETFs in Australia. Their results show that ETFs do not perfectly follow the performance of the benchmark due to market frictions in the short-run, but their findings do suggest that investors with a long-term horizon can be able to achieve investment returns that are similar to the benchmark returns. In an additional investigation by Harper et al. (2006) risk and return performance of foreign markets ETFs and closed-end country funds are compared. The study makes use of performance proxies, such as risk-adjusted returns and mean returns. Also, the study utilizes performance measures such as Sharpe ratio and Jensen's alpha. The research shows that, on average, ETFs have higher risk-adjusted returns than closed-end funds. Furthermore, the article suggests that a passively managed portfolio of ETFs may serve as a viable option to an actively managed portfolio of closed-end funds in terms of achieving better risk-adjusted returns. Additionally, other studies have shown that despite their different features, ETFs are substitutes to mutual funds, albeit not perfect ones. This has been shown by Agapova (2010) using empirical investigation of the substitutability.

In an empirical analysis of ETFs, Buetow and Henderson (2012) conclude that the majority of ETFs traded on U.S exchanges track the returns of their benchmark indices closely. The ETFs that tend to have large tracking errors, and do not track their benchmark indices closely, are those who invest in indices with less liquid assets. Also, the volatility of ETFs has been examined in a study by Kadapakkam et al. (2013). The market efficiencies of ETFs and size-based portfolios were investigated and evidence for that volatility spills over from ETFs of larger firms to those of smaller firms was presented.

Research made by Ivanov (2013) show that oil ETFs have a close relationship in terms of price with the underlying asset price and futures price. The study uses intradaily data to study how closely gold-, silver- and oil ETFs follow their underlying asset. The research concludes that the ETFs follow the underlying asset closely, and that the oil market predominantly has price discovery in the futures market.

A recent study of performance characteristics of ETFs made by Khan et al. (2015) have shown that emerging markets ETFs have a higher tracking error to their respective index than developed market ETFs. However, the emerging markets ETFs examined in the study had higher risk-adjusted returns than their counterparts.

By following the methodologies used in Carhart (1997) and Kreander et al. (2005) this study contributes to the previous research done on ETFs. This, by examining the performance of clean ETFs and dirty ETFs by using a multi-factor model and different performance measures.

Theory Review

This thesis is based on theoretical models within financial economics. Relevant models within return, volatility and risk are used in the thesis. The Sharpe ratio, Treynor ratio, Jensen's alpha and the Carhart (1997) four-factor model will be used to examine the differences in performance.

Sharpe Ratio

The Sharpe ratio is used to measure and compare the level of risk-adjusted return in a portfolio. A reward-to-total-volatility ratio that measures the performance of stocks and funds. The ratio is computed by dividing the average portfolio excess return over the sample period by the standard deviation of returns, also known as total risk, of that period (Bodie et al., 2014). The higher the ratio, the better the portfolio performs relative to the risk taken (Grable and Chatterjee, 2014). In our research the *ex post* version of the Sharpe ratio expressed below will be used (Hodges et al., 1997).

$$\frac{\overline{r_p} - \overline{r_f}}{\sigma_p} \tag{1}$$

$\overline{r_p} = mean\ return\ on\ the\ portfolio$

$\overline{r_f} = risk\ free\ rate$

$\sigma_p = sample\ standard\ deviation\ of\ returns$

Standard Deviation ($\boldsymbol{\sigma}$)

The standard deviation is a measurement of the discrepancy of a value and its mean. In financial terms it is a measure of total risk and describes how large the difference of the expected return deviates from its mean. The higher the deviation, the more spread apart are the values (Bodie et al., 2014).

$$\sigma = \sqrt{Var} = \sqrt{\frac{\Sigma(x_{it} - \bar{x})^2}{N}} \tag{2}$$

$var = the\ variance$

$x_{it} = the\ return\ at\ time\ t$

$\bar{x} = the\ average\ arithmetic\ return$

$N = the\ number\ of\ observations$

Treynor Ratio

Similar to the Sharpe ratio, the Treynor ratio is also a measurement of excess return per unit of risk. The difference is that the Treynor ratio uses systematic risk instead of total risk. Systematic risk is a non-diversifiable risk that is attributable to risk sources that affect the whole market (Bodie et al., 2014).

$$\frac{\overline{r_p} - \overline{r_f}}{\beta_p} \tag{3}$$

$\overline{r_p} = mean\ return\ on\ the\ portfolio$

$\overline{r_f} = risk\ free\ rate$

$\beta_p = systematic\ risk$

Jensen's Alpha

Given the portfolio's beta and the average market return, the Jensen's Alpha is the average return on the portfolio over and above that predicted by the CAPM. A portfolio is undervalued and outperforms the market when its alpha is positive. Thus, a negative alpha indicates that a portfolio is overvalued and underperforms the market (Bodie et al., 2014).

$$\alpha_p = \overline{r_p} - [\overline{r_f} + \beta_p(\overline{r_M} - \overline{r_f})] \tag{4}$$

$\overline{r_p} = mean\ return\ on\ the\ portfolio$

$\overline{r_f} = risk\ free\ rate$

$\beta_p = systematic\ risk$

$\overline{r_M} = expected\ market\ return$

Carhart Four-Factor Model

The Carhart four-factor model was introduced by Carhart (1997). It provides an extension to CAPM and the Fama-French three factor model, which are frequently used when examining mutual fund performance, by adding a fourth factor representing momentum. This fourth factor aims at capturing the momentum anomaly as studied by Jegadeesh and Titman (1993). Thus, the model controls for market, size, book-to-market ratio and stock price momentum and is in theory more explanatory in explaining the risk-adjusted return.

$$R_{it} - R_{ft} = \alpha_i + \beta_{1i}(R_{mt} - R_{ft}) + \beta_{2i}SMB_t + \beta_{3i}HML_t + \beta_{4i}Mom_t + \varepsilon_{it} \quad (5)$$

R_{it} = return on the individual portfolio at time t
R_{ft} = the risk-free rate at time t
α_i = four-factor alpha i.e. the risk-adjusted return for portfolio i
$R_{mt} – R_{ft}$ = excess return of the market at time t
SMB_t = the difference in return between a small cap portfolio and a large cap portfolio at time t
HML_t = the difference in return between a high book-to-market portfolio and a low book-to-market portfolio at time t
Mom_t = the difference in return of portfolios consisting of past winners and past losers at time t
ε_{it} = error term for portfolio i at time t.

The first factor ($R_{mt} – R_{ft}$), known as the Market factor, is the excess return, which has its basis from the CAPM-model developed by Sharpe (1964), Lintner (1965) and Mossin (1966). The Market factor captures the systematic, non-diversifiable, risk. β_{1i} measures the extent to which the portfolio returns mimics the market return. In other words, it indicates how sensitive the portfolio is to market movements. A portfolio with a beta-value above 1 indicates that the portfolio has an above-average sensitivity to market swings. While a portfolio with a beta-value below 1 indicates that the portfolio has a below-average sensitivity to market swings (Bodie et al. 2014).

The following two factors, SMB_t and HML_t were developed by Fama and French and may be used as proxies for yet-unknown more-fundamental variables that may capture sensitivity to risk factors in the market. SMB_t measures the size effects on small firms versus large firms. More specifically, it is the difference in return between a small cap portfolio and a large cap portfolio (Bodie et al. 2014). Previous studies by

Banz (1981) and Van Dijk (2011) show that small cap shares tend to demonstrate both higher risk and higher returns. β_{2i} measures the sensitivity towards the size effects.

HML_t is the difference in return between a high book-to-market portfolio and a low book-to-market portfolio. Book-to-market is a ratio that compares the book value (the value of a firms' assets at the time they entered the balance sheet) to its market value. Previous studies have shown that high book-to-market ratios have resulted in excess risk-adjusted returns over longer periods (Bodie et al. 2014). Value-oriented portfolios tend to have high book-to-market ratios value while growth-oriented portfolios tend to have low book-to-market ratios. Value-oriented portfolios tend to have more investments in traditional value sectors more specifically, sectors that tend to have higher environmental risk (Bauer et al. 2005). β_{3i} measures the sensitivity towards book-to-market effects.

The fourth factor Mom_t was added by Carhart (1997) which in turn is based on the difference in return between the past years "winners" in the portfolio and the past years "losers" i.e. the best and worst in terms of performance. Jegadeesh and Titman (1993) found that profit opportunities could be offered by portfolios of the best performing stocks in the recent past. Hendricks et al. (1993), Goetzmann and Ibbotson (1994) and Wermers (1996) all found that there is persistence in mutual fund performance over short-term horizons, in which they concluded that such persistence could be related to investment strategies. Wermers (1996) also concluded that the short-term persistence may be a result of following momentum strategies. A momentum strategy is when an investor buys the past winner-stocks and sells the past loser-stocks. Thus, considering the Mom_t-factor as an investment strategy, the β_{4i} would pose as a measure of the sensitivity of the portfolio while following the momentum strategy. A coefficient value of 1 would indicate a perfect relationship. The opposite of this strategy is a contrarian strategy where the investor buys stocks that performs poorly, and sells the stocks that perform well.

Last but not least there is a four-factor alpha (α_i) and an error term (ε_{it}) left in the model. The first is the risk-adjusted return, a portfolio with a positive alpha can be interpreted as outperforming the market. Conversely, a portfolio with a negative alpha can be interpreted as underperforming the market. The second is the error term.

Data

The data's structure is time series with daily frequency during the examination period January 2011 to March 2016 and the number of observations is 1319. The dataset was collected from the databases: Bloomberg and Kenneth R. French Data Library. ETF Database and Yahoo! Finance were used as complements.

The dataset used in this study contains missing values. However, the missing values are limited in numbers, and unable to significantly impact the results. None of the ETFs examined in this thesis can be considered an outlier in the sense that it creates a significant skewness of the results.

In a study by Brown et al. (1992) survivorship bias in performance studies are examined. Their findings show that such bias may arise when the collected data, to evaluate performance, only consists of active funds which in turn can favour the final result. To create a data set free of survivorship bias, all known - both active and inactive - funds over the sample period must be included.

The dataset consists of 7 clean ETFs and 7 dirty ETFs with at least five years of daily data respectively. In order to control for their size and age, two equally weighted portfolios were created by using a matched pair approach introduced by Mallin et al. (1995). When matching the portfolios this thesis aims at eliminating differences in size (AUM) and percentage allocated in U.S as much as possible. This resulted in a difference of 16.1 million US dollars in AUM and 19.15 percentage points in percentage allocated in U.S between the two portfolios. By reading the prospectus it can be ensured that the ETFs are "clean" and "dirty" according to the definitions applied in this thesis.

Table 1
Descriptive statistics of the portfolios and the benchmark index

Portfolios	*Obs.*	*Avg. return*	*Std. Dev.*	*Min*	*Max*	*# of ETFs*	*Avg. age*	*Size*
Clean ETF portfolio	1319	-0.072 %	1.968 %	-8.996 %	8.923 %	7	9	671.48
Dirty ETF portfolio	1319	-0.103 %	1.697 %	-6.984 %	6.439 %	7	9	687.58
Difference portfolio	1319	0.031 %	1.534 %	-4.642 %	7.019 %			
S&P500 Index	1319	-0.003 %	0.985 %	-6.669 %	4.723 %			

Notes: Presented are descriptive statistics for each portfolio and the benchmark index. Obs. is the number of observations. Avg. return is average daily return. Std. Dev. is the standard deviation. Min and Max are minimum and maximum return. # of ETFs is the amount of ETFs included in the portfolios. Avg. age is the average age of each portfolio. Size is the total asset under management in millions of US dollars. The difference portfolio is created by subtracting the daily returns of the dirty ETF portfolio from the daily returns of the clean ETF portfolio.

Methodology

ETF Portfolio Construction

To control for the age and size and to be able to compare the performance of the clean ETF portfolio and the dirty ETF portfolio, a matched-pair approach was applied and two equally weighted portfolios were formed. Also, to improve the comparability between the performances of the two portfolios a difference portfolio was constructed. The difference portfolio was constructed by subtracting the daily risk-adjusted returns of the dirty ETF portfolio from the daily risk-adjusted returns of the clean ETF portfolio.

Econometric Model Specification

Using risk-adjusted performance measures such as Sharpe ratio, Treynor ratio, Jensen's alpha, along with the Carhart (1997) four-factor model, a comparison is made between the clean ETF portfolio and the dirty ETF portfolio. In terms of risk measurements, the standard deviation and the beta are observed for the Sharpe ratio and Treynor ratio. Thereafter, the results are compared for the two groups in order to determine risk characteristics.

The two portfolios' returns are constructed by first collecting daily returns for each individual ETF and then by calculating average returns for each portfolio. Since the analysis is made for two constructed portfolios and not for single fund returns, the log of the returns has not been calculated.

The proxy for the risk-free rate (R_{ft}) used in this study is the daily treasury bill rates collected from the U.S Department of the Treasury. The excess return ($R_{it} – R_{ft}$) of the portfolios is obtained by subtracting the daily treasury bill rates from the average daily return of the portfolios.

The following is a model specification of the Carhart (1997) four-factor model:

$$R_{it} - R_{ft} = \alpha_i + \beta_{1i}(R_{mt} - R_{ft}) + \beta_{2i}SMB_t + \beta_{3i}HML_t + \beta_{4i}Mom_t + \varepsilon_{it} \quad (5)$$

R_{it} = return on the individual portfolio at time t

R_{ft} = the risk-free rate at time t

α_i = four-factor alpha i.e. the risk-adjusted return for portfolio i

$R_{mt} - R_{ft}$ = excess return of the market at time t

SMB_t = the difference in return between a small cap portfolio and a large cap portfolio at time t

HML_t = the difference in return between a high book-to-market portfolio and a low book-to-market portfolio at time t

Mom_t = the difference in return of portfolios consisting of past winners and past losers at time t

ε_{it} = error term for portfolio i at time t.

By regressing the Carhart (1997) four-factor model estimated beta- and alpha values are obtained. The alphas are the risk-adjusted return while the betas are the dependent variables' sensitivity against the specific factor, ceteris paribus.

Factor Portfolio Construction

The value-weighted return of all US firms listed on the NYSE, AMEX or NASDAQ - subtracted from the one-month T-Bill - is the excess return for the market ($R_m - R_f$). The SMB and HML factors are constructed by first forming six value-weighted portfolios that are ranked on size and book-to-market of all NYSE, AMEX and NASDAQ stocks. These portfolios are the intersections of two portfolios ranked on size (small or big) and three portfolios ranked on the book-to-market ratio (value, neutral or growth). To define whether the stocks are small or big at year *t*, the size median for the NYSE market equity is used. Further, to define whether the stocks are growth-, neutral- or value oriented, the 30th and 70th percentiles are used.

Figure 1
Six portfolios formed on size and book-to-market,
Kenneth R. French Data Library (2016)

	Median ME	
70th BE/ME percentile	Small Value	Big Value
	Small Neutral	Big Neutral
30th BE/ME percentile	Small Growth	Big Growth

The SMB factor is constructed by subtracting the average return on the three big portfolios from the average return on the three small portfolios.

$$SMB = \frac{1}{3}\,(Small\ Value + Small\ Neutral + Small\ Growth) - \frac{1}{3}\,(Big\ Value + Big\ Neutral + Big\ Growth) \qquad (6)$$

The HML factor is constructed by subtracting the average return on the two growth portfolios from the average return on the two value portfolios.

$$HML = \frac{1}{2}\,(Small\ Value + Big\ Value) - \frac{1}{2}(\ Small\ Growth + Big\ Growth) \qquad (7)$$

The Mom factor is constructed by using a six value-weighted portfolio ranked on size and prior returns of all NYSE, AMEX and NASDAQ stocks. These portfolios are the intersections of two portfolios ranked on size (small or big) and three portfolios ranked on prior returns. To define whether the stocks are small or big, the daily size median for NYSE equity is used. Further, to define whether the stocks are growth-, neutral- or value oriented, the 30^{th} and 70^{th} NYSE percentiles are used.

Figure 2

Six portfolios formed on size and momentum,
Kenneth R. French Data Library (2016)

	Median ME	
70th prior (2-12) percentile	Small Up	Big Up
	Small Medium	Big Medium
30th prior (2-12) percentile	Small Down	Big Down

The Mom factor is constructed by subtracting the average return on the two low prior return portfolios from the average return on the two high prior return portfolios.

$$Mom = \frac{1}{2}\ (Small\ Up + Big\ Up) - \frac{1}{2}(Small\ Down + Big\ Down) \tag{8}$$

Result and Analysis

The result and analysis section is divided into two parts. First, the results estimated by using OLS as estimation method on the four-factor model (Carhart, 1997) for the clean and dirty portfolio, their difference portfolio and the benchmark index are shown. Secondly, the calculated performance measures, average returns, standard deviations and betas for each portfolio and the benchmark index are presented. Both the first and the second part contribute to answering the research question, which is if there is a difference in risk-adjusted return between the clean ETF portfolio and the dirty ETF portfolio.

Carhart Four-Factor Model

This section and table 2 present the result given by the Carhart (1997) four-factor model applied on our clean ETF portfolio, dirty ETF portfolio, difference portfolio and the benchmark index.

Table 2
Carhart (1997) four-factor model

Variables	*Clean ETF portfolio*	*Dirty ETF portfolio*	*Difference portfolio*	*S&P500 Index*
Four-factor alpha	-0.124*	-0.127*	0.004	-0.052*
	(0.031)	(0.026)	(0.039)	(0.003)
Market	1.272*	0.876*	0.396*	0.995*
	(0.035)	(0.037)	(0.042)	(0.003)
SMB	0.612*	0.079	0.533*	-0.137*
	(0.067)	(0.069)	(0.086)	(0.005)
HML	-0.125	0.291*	-0.416*	0.007
	(0.078)	(0.090)	(0.109)	(0.007)
Mom	-0.260*	-0.645*	0.385*	0.000
	(0.048)	(0.056)	(0.070)	(0.005)
Observations	1319	1319	1319	1319

Notes: Presented are estimates for each portfolio and the benchmark index during the examination period January 2011–March 2016. The estimates are the results from regressing daily excess return on the daily factor returns. Newey-West standard errors, correcting for serial correlation and heteroscedasticity, are in parentheses. The difference portfolio is created by subtracting the daily returns of the dirty ETF portfolio from the daily returns of the clean ETF portfolio. The OLS estimation was made using the following equation:

$$R_{it} - R_{ft} = \alpha_i + \beta_{1i}(R_{mt} - R_{ft}) + \beta_{2i}SMB_t + \beta_{3i}HML_t + \beta_{4i}Mom_t + \varepsilon_{it} \quad (5)$$

Where R_{it} is the portfolio return, R_{ft} is the risk-free rate, $(R_{mt} - R_{ft})$ is the excess return of the market, SMB_t captures the size effects, HML_t captures book-to-market effects, Mom_t captures the momentum effects

* Significant at the 1 % level.

The estimated four-factor alphas are statistically and economically significant, and negative, for the two portfolios and the benchmark index. This indicates that none of the two portfolios nor the benchmark generated excess return during the examination period. When interpreting the alpha of the difference portfolio it shows a statistically insignificant alpha, meaning that there is no significant difference in performance between the portfolios. As shown, none of the two portfolios outperformed the benchmark index during the examination period and this is in turn in line with the work of Gallagher D.R and Segara R (2006) where the results showed that ETFs in the Australian market do not perfectly follow the performance of the benchmark. As concluded in that study, this may be due to market frictions in the short-run. However, their study does suggest that investors with a long-term horizon can be able to achieve investment returns that are similar to the benchmark returns.

As for the Market factors, these are also statistically and economically significant for both portfolios and the benchmark index. As shown in the table, the clean portfolio
(1.272) is slightly more exposed to the Market factor compared to its counterpart (0.876). The level of exposure to this factor indicates how volatile the portfolios are to the market. In other words, the clean portfolio is more volatile to the market than its counterpart. The statistically significant market beta of the difference portfolio indicates that there is a significant difference in exposure to the market between the two portfolios. One possible explanation could be due to the fact that ETFs in general are quite new, and that their investment universe might be smaller compared to other financial instruments. Also, when looking within the ETF universe, the ones with holdings within a sector with negative environmental impact (i.e. dirty ETFs) tend to be older and larger ETFs than those with positive environmental impacts (i.e. clean ETFs). Thus, ETFs with holdings within a sector with negative environmental impact tend to be more stable through time i.e. less volatile to market movements.

In contrast to the two previously analyzed factors, the SMB factors are only statistically and economically significant for the clean ETF portfolio and the benchmark index. However, the magnitude and the signs of these factors differ. While the SMB factor of the clean portfolio has a positive sign, the SMB factor of the benchmark index has a negative sign. The higher magnitude and the positive SMB factor for the clean portfolio indicate that this portfolio has more exposure to, and more investments in, small cap investments. The benchmark index with its negative SMB factor indicates that it obviously has more exposure to large caps. The significant SMB factor of the difference portfolio indicates that there is a significant difference in exposure to small caps between the two portfolios. These results may be explained with a similar interpretation as for the Market factor above, saying that the universe of clean ETFs tends to be smaller than that of dirty ETFs, but more specifically the fact that the clean portfolio has more exposure to small caps. According to Banz (1981) and Van Dijk (2011), small cap shares tend to demonstrate both higher risk and higher returns.

When analyzing the HML factor, it is only statistically and economically significant for the dirty portfolio. Its positive sign indicates that the dirty portfolio is more value-oriented, or less growth-oriented. Likewise, it also suggests that the dirty portfolio has a high book-to-market ratio. When interpreting the HML factor of the difference portfolio, a statistically significant difference in terms of exposure to the factor is shown. A possible explanation to this may be due to the fact that the portfolios in this study consist of a mixture between international and domestic ETFs, and that the two portfolios have investments in quite different sectors. Also, as explained in the Theory Review section, the result points towards the conclusion that there is a difference in exposure to renewable energy companies versus oil companies.

The final factor analyzed in this section is the Mom factor. The factor is negative and statistically and economically significant for both portfolios but not for the benchmark index. A Mom factor equal to 1 indicates that the portfolio follows the momentum strategy perfectly. Thus, the negative signs in the results indicate that the portfolios do not follow the momentum strategy, but rather follow the contrarian strategy. In other words, the portfolios contain more of contrarian stocks and less of momentum stocks. The statistically significant Mom factor of the difference portfolio implies that there is a significant difference in how well the portfolios follow the momentum strategy. As presented, both portfolios have negative Mom factors, the dirty portfolio has a more negative Mom factor (-0.645 compared to -0.260) which can be interpreted as it is following the contrarian strategy more than what its counterpart does. The notion that neither portfolio follows the momentum strategy contradicts with the findings of Buetow and Henderson (2012). Their study concluded that the majority of ETFs traded on U.S exchanges track the return of their benchmark indices closely. However, the ETFs included in the portfolios as assembled in this thesis are not strictly U.S allocated. This may, at least partially, explain the inconsistent result. Furthermore, another plausible explanation is the fact that the oil price during the examination period demonstrated a negative trend. This negative trend, along with a similar negative development of the coal price, is a likely cause to the negative Mom factor for the dirty portfolio. A continued discussion regarding the oil- and coal price follows in the next section.

Performance Measures

This section aims at presenting the different performance measures used to evaluate the performance for both the clean and the dirty portfolio, along with the benchmark. Results for the various measures are displayed in table 3.

Table 3
Portfolio Charachteristics

Portfolio	*Average Return*	*Standard Deviation*	*Beta*	*Sharpe Ratio*	*Treynor Ratio*	*Jensen's Alpha*
Clean ETF	-0.072 %	1.968 %	1.412	-3.635 %	-5.007 %	-4.599 %
Dirty ETF	-0.103 %	1.697 %	0.956	-6.039 %	-11.646 %	-10.240 %
S&P500 Index	-0.003 %	0.985 %	1.000	-0.259 %	-0.255 %	0.000 %

Notes: Presented are performance measures for each portfolio and the benchmark index during the examination period January 2011–March 2016. The average returns are average daily returns.

Using an arithmetical average for the daily returns, a comparison can be made between the two portfolios and the benchmark. As can be seen in the table, none of the portfolios nor the benchmark generated positive daily returns on average over the examination period. Both the clean and the dirty portfolio performed worse than the market proxy over the examination period. Interesting to note is that the dirty portfolio averaged the most negative returns of the portfolios. Furthermore, the negative returns for both portfolios are considered economically significant, since the magnitudes -0.072% and -0.103% accumulates into annual average effects for the examination period of -17.956% and -25.728% (Wooldridge, 2014). In comparison to the economically insignificant annual average of -0.638% for the benchmark index, the numbers strongly indicate that the examination period on average was a negative period for energy ETFs. A test where the single best performing ETF and the single worst performing ETF are removed from each portfolio before the calculation generates daily magnitudes of -0.071% for the clean portfolio and -0.100% for the dirty portfolio. These magnitudes accumulate into annual averages of -17.838 and -24.997% respectively, which are still economically significant magnitudes.

One plausible explanation for the large and negative average returns for the dirty ETFs is that some of the companies that the ETFs track have experienced negative growth during the examination period, along with the ETFs having exposure to sectors with negative growth. As previously mentioned the coal sector, which is influencing a number of the ETFs in the dirty portfolio, can serve as an example. According to the MVIS Global Coal Index, annualized returns for the largest companies in the coal sector over a five-year period until now are approximately -27.11% (Morningstar). Moreover, crude oil prices have dropped during the examination period with approximately -58.02% (Reuters). Explanations for the negative returns in the clean portfolio are likely linked to the negative price developments in the fossil energy sector. Presumably, a likely cause for the negative returns, and thus downward pressure in the sector, is the drop in oil prices as previously mentioned. Lower oil prices are likely to make private consumers and businesses that are looking for a good time to transition into sustainable energy to wait, since oil prices are getting lower, and since a transition to clean energy is likely to incur some initial fixed cost. Thus, the option of transitioning into alternative energy might not have seemed like such a good investment considering the alternative during the examination period.

In terms of standard deviations and betas, the clean portfolio is noteworthy. The standard deviation is higher than the benchmark, as is the beta value. Compared to the dirty portfolio, the standard deviation is moderately higher, and the beta is higher. It is a reasonable conclusion from observing the standard deviations that the clean portfolio is the most volatile during the examination period of this study. The standard deviation is noticeably higher than the benchmark. The beta is higher for the clean portfolio compared both with the dirty portfolio and the benchmark. This translates into that movements in the markets have larger implications for the clean portfolio than for the dirty. Presumptively, investors that are volatility-averse might shy away from a portfolio demonstrating such aggressive betas (Bodie et al., 2014).

Moreover, it is interesting to note that while the standard deviation is slightly higher for the clean portfolio than for the dirty portfolio, the differences in terms of beta are larger. For the dirty portfolio, with a beta closer to 1, this implies that over the period measured in this study the returns for the dirty portfolio are more market-reverting (Bodie et al., 2014). Such a value is in line with the research by Buetow and Henderson (2012) showing that the majority of ETFs traded on U.S exchanges track the returns to their benchmark indices closely. The notion that ETFs that are tracking oil prices follow their underlying asset closely gives more weight to this possibility (Ivanov, 2013). A possible explanation to the differences could be the fact that many of the companies in the oil sector are much larger in size than companies in the renewable energy sector. This translates into the fact that oil companies, due to their size, are more likely than their counterparts to drive the market in a certain direction. Hence, the beta close to 1 could at least partially be a result of this relationship between the market and the large oil companies. To mention one company as a relevant example, Exxon Mobil Corporation (XOM) has been the largest company in the world in terms of market capitalization and is an energy company engaged in the production and exploration of crude oil and natural gas (Reuters). Also, given that the clean energy sector is younger and smaller, higher volatility compared to the larger, “dirty”, energy sector is reasonable to assume.

Further explanations to the high volatility for the clean portfolio are not obvious. As described in the methodology section, this study has aimed from the beginning at eliminating differences in size and allocation. Even though some minor differences in size and allocation are present between the portfolios, it is not evident that they explain the differences in volatility. However, in line with the previously discussed SMB factor, the clean portfolio is likely more exposed to small cap shares than its counterpart. This notion provides an additional explanation, since small cap shares have been shown to demonstrate both higher risk and higher returns in some studies (Banz, 1981; Van Dijk, 2011). Thus, the clean portfolio seems to perform in accordance with previously studied small cap portfolios.

When comparing Sharpe ratios for the two portfolios, the clean portfolio performed better than the dirty portfolio. With a Sharpe ratio of -3.635% it performed noticeably better than its dirty counterpart, for which the Sharpe ratio was -6.039%. These results are in line with the previously discussed results showing that the clean portfolio demonstrated a higher average return. An explanation for the results is likely the fact that the oil price declined during the end of the examination period, along with negative development in the coal sector. As showed in the research by Ivanov (2013) oil ETFs tend to closely track the performance of the underlying asset. Most of the ETFs included in the dirty portfolio are heavily allocated in oil or oil extraction related activities. In this case, since the oil price declined, the return on ETFs with a heavy allocation in the oil sector is likely to also be negatively affected. When it comes to the denominator of the Sharpe ratio, namely the standard deviation, the dirty portfolio has a slightly smaller standard deviation than the clean portfolio. However, since the average returns, the numerators, are quite different, this translates into Sharpe ratios that are distinctively different. Due to the negative ratios, investors whose primary objective is to seek a good Sharpe ratio are likely to dismiss both portfolios.

Subsequently, the Treynor ratios for the both portfolios give similar indications since the Treynor ratio measures the excess return per unit of systematic risk. The clean portfolio has a Treynor ratio of -5.007% whereas the dirty portfolio has -11.646%. Treynor ratio for the benchmark is -0.255%. Following this result, the Treynor ratios give indications of relatively high risk and low reward compared with the benchmark. Again, this suggests that the dirty portfolio has a slightly worse performance in comparison with the clean portfolio.

Jensen's alpha is -4.599% for the clean portfolio and -10.240% for the dirty portfolio. This result points towards the notion that both the clean portfolio and the dirty portfolio have lower returns than that predicted by the CAPM, given the average market return and the portfolios' betas. Given this information, neither of the portfolios as assembled in this thesis would be likely to attract an investor who is seeking primarily to allocate capital where the alpha is highest. Such an "alpha betting" strategy would most likely make investors consider other options than the portfolios as assembled in this thesis (Bodie et al., 2014). This is in line with the previously discussed performance measures which all point at the same direction.

Conclusions

The main objective of this study is to examine if there is a difference in performance between clean and dirty ETFs using a matched pair approach to control for the effects of size and age. To investigate the primary objective of this study, performance measures such as the Sharpe ratio, Treynor ratio and Jensen's alpha and the Carhart (1997) four-factor model are used. The dataset consists of daily return data during the examination period January 2011 to March 2016.

The results in this study are limited to the specific portfolios used, and during the examination period for which the data is available. They are not necessarily expected for all clean or dirty ETFs across all time periods. The results suggest that clean ETFs, as part of the portfolio assembled for this study, in terms of performance are not worse off compared to dirty ETFs. For investors seeking ways to access opportunities in sustainable investing, the results in this study could therefore be of much interest.

While the different measures used in this study point towards the conclusion that the clean ETFs are not worse than the dirty in many regards, the price to pay for investors seems to be higher volatility. As previously discussed, investors face higher volatility through the clean portfolio in this study. Not surprisingly, the individual ETF demonstrating the highest volatility in this study was found in the clean portfolio. Nevertheless, as previously discussed, alpha betting investors and investors seeking high Sharpe ratios are likely to dismiss both portfolios.

Since this study has assembled the portfolios using criteria such as "clean" and "dirty", a suggestion for further research would be to explore the heterogeneity of ETFs in general, and the heterogeneity of clean energy ETFs in particular. Especially since there is a vast number of different types and with different focuses. An example of how such research could be made is to utilize subgroups for which different research hypotheses are tested. Secondly, a division of the research into different, smaller, time cycles could also be made in order to pinpoint and analyze differences more accurately. Such a division might also help link the differences in performance to specific macroeconomic events.

Answering the research question raised in this thesis, there are no statistically significant differences in performance between clean and dirty ETFs according to the Carhart (1997) four-factor model. However, there are some differences in performance as reported by the Sharpe ratio, Treynor ratio and Jensen's alpha. Because of this ambiguity, investors pursuing different strategies may reach different conclusions as to which portfolio is the most preferable. Different factors influenced the two portfolios differently. As previously shown, the portfolio of clean ETFs did not perform worse than the portfolio of dirty ETFs during the examination period used in this thesis.

References

Agapova, A. (2010). Conventional mutual index funds versus exchange-traded funds. *Journal of Financial Markets*, 14 (2), pp. 323-343.

Banz, R. (1981). The Relationship between Return and Market Value of Common Stocks. *Journal of Financial Economics*, 9, pp. 3-18.

Bauer, R., Koedijk, K. and Otten, R. (2005). International Evidence on ethical mutual fund performance and investment style. *Journal of Banking & Finance*, 29, pp. 1751-1767.

Bodie Z., Kane A. and Marcus A.J. (2014). *Investments*. Tenth Edition. Berkshire: McGraw-Hill.

Brown, S.J., Goetzmann, N.W., Ibbotson, R.G. and Ross, S.A. (1992). Survivorship bias in performance studies. *Review of Financial Studies*, 5, pp. 553-580.

Buetow, G.W., Henderson, B.J. (2012). An Empirical Analysis of Exchange-Traded Funds. *The Journal of Portfolio Management*, 38 (4), pp. 112-127.

Gallagher, D.R and Segara, R. (2006). The performance and trading characteristics of exchange-traded funds. *Journal of Investment Strategy*, 1 (2), pp.49-60.

Gastineau, G.L (2010) *Exchange Traded Funds Manual*. Second Edition. Wiley Finance.

Goetzmann, N. W. and Ibbotson, G. R. (1994). Do winners repeat? Pattern in mutual fund performance. *Journal of Portfolio Management,* 20, pp. 9-18.

Grable, J.E. and Chatterjee, S. (2014). The Sharpe Ratio and Negative Excess Returns: The Problem and Solution, *Journal of Financial Service Professionals*, pp.12-13.

Harper, J.T., Madura, J. and Schnusenberg, O. (2006). Performance comparison between exchange-traded funds and closed-end country funds. *Journal of International Financial Markets, Institutions and Money*, 16 (2), pp.104-122.

Hendricks, D., Patel, J. and Zeckhauser, R. (1993). Hot Hands in Mutual Funds: Short-Run Persistence of Relative Performance, 1974-1988. *The Journal of Finance*, 48 (1), pp. 93-130.

Hodges, C.W., Taylor, W R.L. and Yoder, J.A. (1997). Stocks, Bonds, the Sharpe Ratio and the Investment Horizon, *Financial Analysts Journal*, 53 (6), pp.74-80.

Ivanov, S.I. (2013) The influence of ETFs on the price discovery of gold, silver and oil. *Journal of Economics and Finance,* 37, pp. 453–462

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

Kadapakkam, P.R., Krause, T. and Tse, Y. (2013). Exchange traded funds, size-based portfolios, and market efficiency. *Review of Quantitative Finance and Accounting*, 45 (1), pp. 89-110.

Khan, A.P, Bacha, O.I and Mashi, A.M.M. (2015). Performance and Trading Characteristics of Exchange Traded Funds: Developed vs Emerging Markets. *Capital Markets Review*, 23, pp. 40-64.

Kreander, N., Gray, R.H., Power, D.M. and Sinclair, C.D. (2005). Evaluating the Performance of Ethical and Non-ethical Funds: A Matched Pair Analysis. *Journal of Business Finance & Accounting*, 32, pp. 1465-1493.

Mallin, C.A., Saadouni, B., and Briston, R.J. (1995). The financial performance of ethical investment funds. *Journal of Business Finance & Accounting*, 22, pp. 483-496.

Poterba, J.M and Shoven J.B. (2002). Exchange Traded Funds: A new investment option for taxable investors. Working paper, National Bureau of Economic Research.

Rosella, M.R and Pugliese, D. (2006). The evolution of the exchange-traded fund: is active management on the horizon?, *Journal of Investment Compliance*, 7 (3), pp. 44-50.

Van Dijk, M.A. (2011). Is size dead? A review of the size effect in equity returns. *Journal of Banking & Finance,* 35, pp. 3263-3274.

Wermers, R. (1996). Momentum investments strategies of mutual funds, performance persistence, and survivorship bias. Working paper, Graduate School of Business and Administration, University of Colorado at Boulder, Boulder, Col.

Databases

Bloomberg. (2016) *Bloomberg*. (Online). Available at: Subscription Service (Accessed: April 15th 2016).

Kenneth R. French. (2016) *Kenneth R. French Data Library.* (Online). Available at: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html (Accessed: April 15th 2016).

Morningstar. (2016) *Morningstar Direct*. (Online). Available at: Subscription Service (Accessed: April 15th 2016).

Reuters. (2016) *Thomsom Reuters Eikon.* (Online). Available at: Subscription Service (Accessed: April 15th 2016).

Yahoo. (2016) *Yahoo Finance.* (Online). Available at: Subscription Service (Accessed: April 15th 2016).

Websites

ETFdb.com, (2016) (Online). Available at: http://etfdb.com (Accessed: April 15th 2016).

Unpri.com, (2016) (Online). Available at: https://www.unpri.org (Accessed: April 15th 2016)

U.S Department of the Treasury, (2016) (Online). Available at: https://www.treasury.gov/resource-center/data-chart-center (Accessed: April 15th 2016)

Appendix

1. a. Summary Information of the Clean and Dirty ETF Portfolios

Table 4

Clean ETF	*Ticker*	*Size (USDm) 2016.04.25*	*% allocated in US*	*Inception date*
First Trust NASDAQ Clean Edge Green Energy	QCLN US	65.75	90.92	2007.02.14
Guggenheim Solar	TAN US	249.17	41.94	2008.04.15
iShares Global Clean Energy	ICLN US	84.76	26.81	2008.06.25
Market Vectors Global Alternative Energy	GEX US	93.68	54.94	2007.05.09
Market Vectors Solar Energy	KWT US	14.90	25.24	2008.04.23
PowerShares Global Clean Energy	PBD US	60.39	27.70	2007.06.13
PowerShares WilderHill Clean Energy Portfolio	PBW US	102.83	76.31	2005.03.03

Table 5

Dirty ETF	*Ticker*	*Size (USDm) 2016.04.25*	*% allocated in US*	*Inception date*
Guggenheim S&P 500 Equal Weight Energy	RYE US	208.62	97.77	2006.11.07
Market Vectors Coal	KOL US	52.83	15.01	2008.01.14
PowerShares DB Energy	DBE US	78.68	44.27	2007.01.05
United States Brent Oil	BNO US	124.20	61.86	2010.06.02
United States Gasoline	UGA US	87.56	66.95	2008.02.27
United States 12 Month Oil	USL US	118.84	56.08	2007.12.06
WisdomTtree Global Natural Resources	GNAT US	16.85	21.07	2006.10.13

1. b. Individual ETF Characteristics

Table 6

Portfolio	*ETF Ticker*	*Average Return*	*Standard Deviation*	*Beta*	*Sharpe Ratio*	*Treynor Ratio*	*Jensen's Alpha*
Clean ETF	KWT US	-0.103 %	2.637 %	1.562	-3.902 %	-6.585 %	-6.892 %
Clean ETF	ICLN US	-0.068 %	1.664 %	1.251	-4.102 %	-5.454 %	-5.165 %
Clean ETF	GEX US	-0.044 %	1.633 %	1.342	-2.693 %	-3.278 %	-2.233 %
Clean ETF	PBD US	-0.056 %	1.487 %	1.271	-3.780 %	-4.424 %	-3.854 %
Clean ETF	QCLN US	-0.043 %	1.783 %	1.352	-2.385 %	-3.146 %	-2.031 %
Clean ETF	TAN US	-0.089 %	2.717 %	1.672	-3.271 %	-5.317 %	-4.881 %
Clean ETF	PBW US	-0.098 %	1.855 %	1.431	-5.285 %	-6.849 %	-7.139 %
Dirty ETF	KOL US	-0.167 %	1.800 %	1.361	-9.263 %	-12.252 %	-14.401 %
Dirty ETF	GNAT US	-0.090 %	1.529 %	1.212	-5.874 %	-7.412 %	-7.543 %
Dirty ETF	USL US	-0.113 %	1.672 %	0.786	-6.754 %	-14.357 %	-12.228 %
Dirty ETF	DBE US	-0.114 %	1.489 %	0.645	-7.642 %	-17.644 %	-13.109 %
Dirty ETF	UGA US	-0.074 %	1.833 %	0.653	-4.051 %	-11.373 %	-9.110 %
Dirty ETF	RYE US	-0.053 %	1.718 %	1.293	-3.077 %	-4.090 %	-3.398 %
Dirty ETF	BNO US	-0.107 %	1.840 %	0.744	-5.822 %	-14.396 %	-11.888 %

1. c. Carhart Four-Factor Model

Table 7

Variables	*QCLN US*	*TAN US*	*ICLN US*	*GEX US*	*KWT US*	*PBD US*	*PBW US*	*S&P 500 Index*
Four-factor alpha	-0.092*	-0.147	-0.119*	-0.097*	-0.160*	-0.105*	-0.146*	-0.052*
	(0.029)	(0.059)	(0.031)	(0.024)	(0.057)	(0.021)	(0.029)	(0.003)
Market	1.200*	1.484*	1.178*	1.238*	1.405*	1.166*	1.234*	0.995*
	(0.044)	(0.063)	(0.033)	(0.033)	(0.062)	(0.026)	(0.033)	(0.003)
SMB	0.755*	0.838*	0.203*	0.415*	0.715*	0.388*	0.974*	-0.137*
	(0.074)	(0.123)	(0.066)	(0.058)	(0.123)	(0.049)	(0.068)	(0.005)
HML	-0.376*	-0.194	0.076	-0.144	-0.178	-0.050	-0.010	0.007
	(0.087)	(0.142)	(0.077)	(0.069)	(0.157)	(0.055)	(0.078)	(0.007)
Mom	-0.239*	-0.405*	-0.202*	-0.230*	-0.309*	-0.218*	-0.221*	0.000
	(0.047)	(0.092)	(0.046)	(0.040)	(0.093)	(0.036)	(0.052)	(0.005)
Observations	1319	1319	1319	1319	1319	1319	1319	1319

Notes: Presented are estimates for each clean ETF and the benchmark index during the examination period January 2011–March 2016. The estimates are the results from regressing daily excess return on the daily factor returns. Newey-West standard errors, correcting for serial correlation and heteroscedasticity, are in parentheses. The OLS estimation was made using the following equation:

$$R_{it} - R_{ft} = \alpha_i + \beta_{1i}(R_{mt} - R_{ft}) + \beta_{2i}SMB_t + \beta_{3i}HML_t + \beta_{4i}Mom_t + \varepsilon_{it} \quad (5)$$

Where R_{it} is the portfolio return, R_{ft} is the risk-free rate, $(R_{mt} - R_{ft})$ is the excess return of the market, SMB_t captures the size effects, HML_t captures book-to-market effects, Mom_t captures the momentum effects

* Significant at the 1 % level.

Table 8

Variables	*RYE US*	*KOL US*	*DBE US*	*BNO US*	*UGA US*	*USL US*	*GNAT US*	*S&P 500 Index*
Four-factor alpha	-0.092*	-0.208*	-0.126*	-0.120*	-0.086	-0.129*	-0.132*	-0.052*
	(0.027)	(0.030)	(0.034)	(0.042)	(0.045)	(0.037)	(0.023)	(0.003)
Market	1.200*	1.230*	0.583*	0.681*	0.582*	0.708*	1.147*	0.995*
	(0.036)	(0.039)	(0.045)	(0.057)	(0.059)	(0.054)	(0.030)	(0.003)
SMB	0.105	0.353*	0.020	-0.043	0.075	0.056	-0.010	-0.137*
	(0.071)	(0.073)	(0.081)	(0.103)	(0.108)	(0.096)	(0.062)	(0.005)
HML	0.388*	0.169	0.293*	0.358	0.248	0.323	0.259*	0.007
	(0.090)	(0.084)	(0.112)	(0.142)	(0.152)	(0.125)	(0.078)	(0.007)
Mom	-0.669*	-0.654*	-0.609*	-0.769*	-0.605*	-0.693*	-0.514*	0.000
	(0.062)	(0.052)	(0.070)	(0.091)	(0.089)	(0.080)	(0.042)	(0.005)
Observations	1319	1319	1319	1319	1319	1319	1319	1319

Notes: Presented are estimates for each dirty ETF and the benchmark index during the examination period January 2011–March 2016. The estimates are the results from regressing daily excess return on the daily factor returns. Newey-West standard errors, correcting for serial correlation and heteroscedasticity, are in parentheses. The OLS estimation was made using the following equation:

$$R_{it} - R_{ft} = \alpha_i + \beta_{1i}(R_{mt} - R_{ft}) + \beta_{2i}SMB_t + \beta_{3i}HML_t + \beta_{4i}Mom_t + \varepsilon_{it} \quad (5)$$

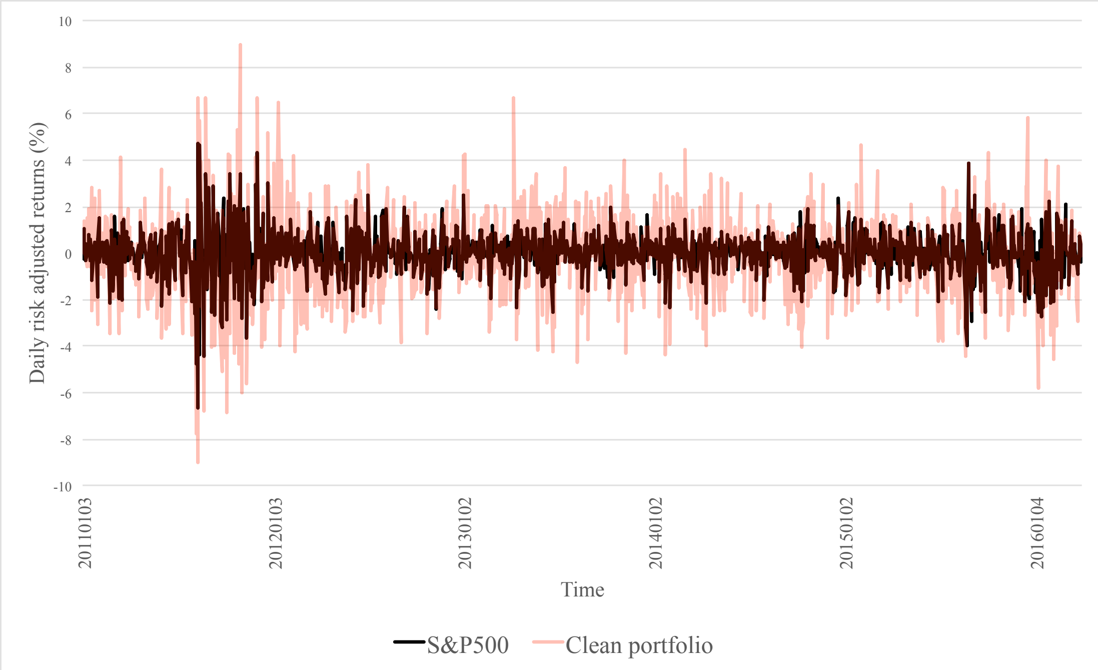
Where R_{it} is the portfolio return, R_{ft} is the risk-free rate, $(R_{mt} - R_{ft})$ is the excess return of the market, SMB_t captures the size effects, HML_t captures book-to-market effects, Mom_t captures the momentum effects

* Significant at the 1 % level.

1. d. Descriptive Statistics

Graph 1

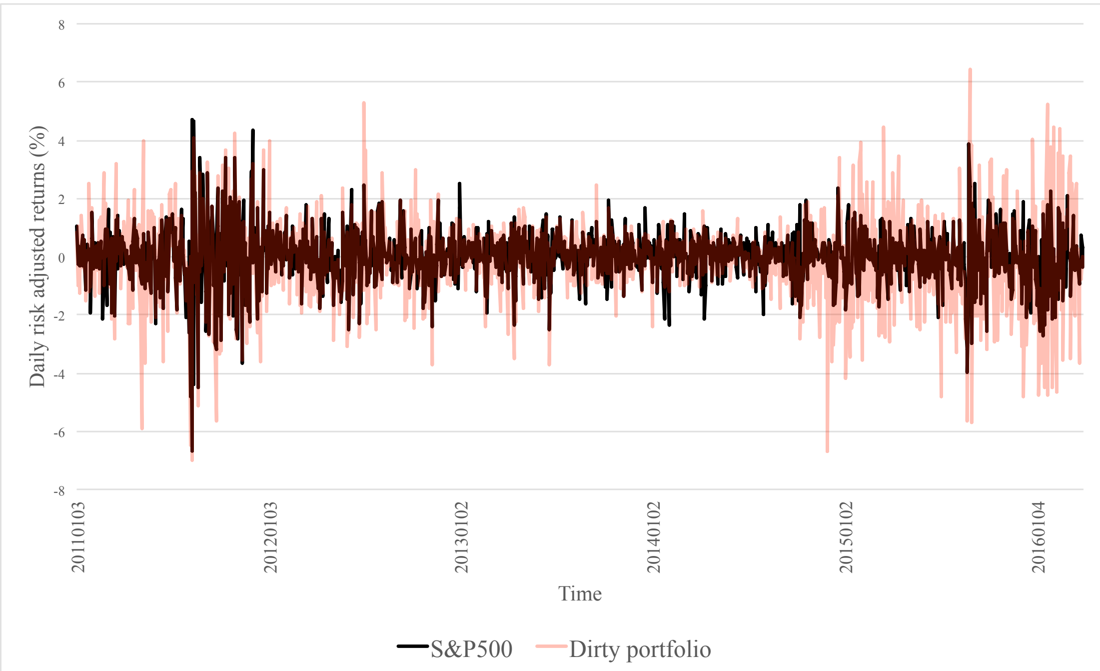
Daily returns for the clean portfolio



Notes: Presented are daily risk adjusted returns for the clean portfolio and the benchmark index, S&P500, during the examination period January 2011–March 2016.

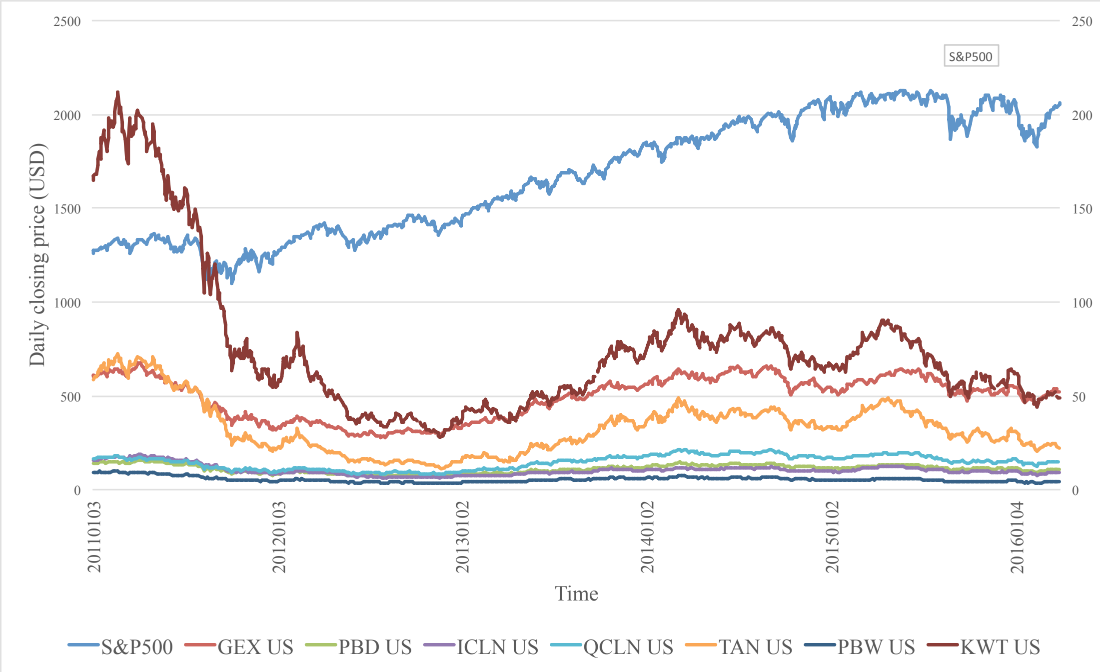
Graph 2

Daily returns for the dirty portfolio



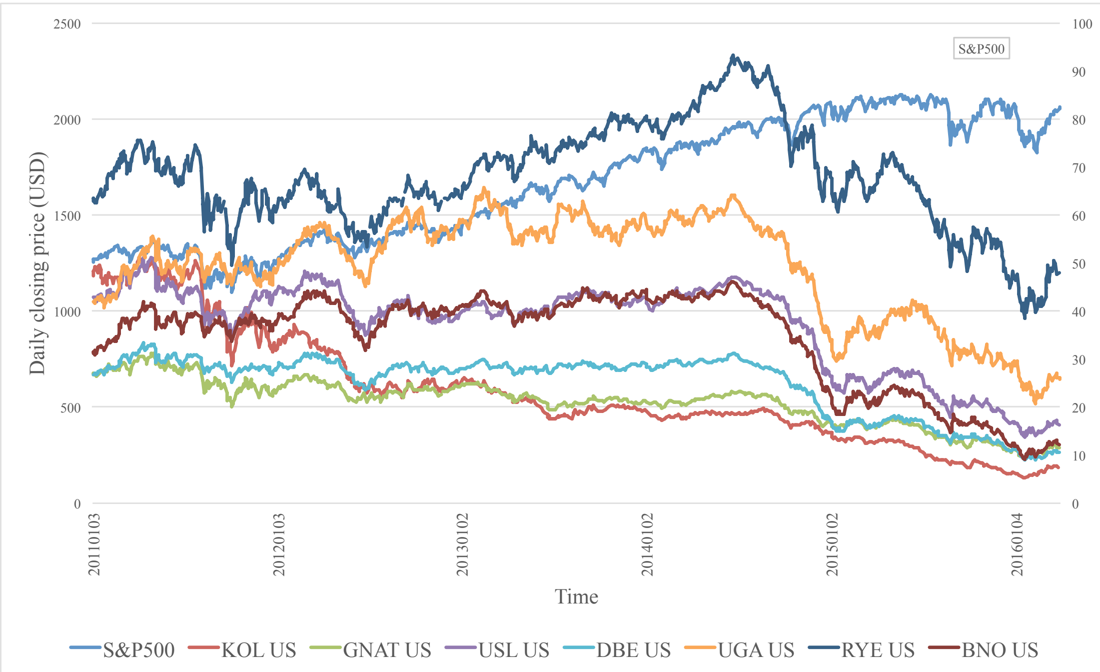
Notes: Presented are daily risk adjusted returns for the dirty portfolio and the benchmark index, S&P500, during the examination period January 2011–March 2016.

Graph 3
Daily closing price for the clean portfolio



Notes: Presented are daily closing prices for each individual ETF in the clean portfolio along with the benchmark index, S&P500, during the examination period January 2011–March 2016. ETF prices are shown on the secondary axis and the benchmark on the primary axis.

Graph 4
Daily closing price for the dirty portfolio



Notes: Presented are daily closing prices for each individual ETF in the dirty portfolio along with the benchmark index, S&P500, during the examination period January 2011–March 2016. ETF prices are shown on the secondary axis and the benchmark on the primary axis.

1. e. Tests of OLS Assumptions

Table 9
Correlation Matrix

	Market	SMB	HML	Mom
Market	1.000			
SMB	0.377	1.000		
HML	0.013	-0.187	1.000	
Mom	-0.098	-0.074	-0.429	1.000

Notes: Presented are the cross-correlations between the factors used in the Carhart (1997) four-factor model. The results indicate that there is no perfect collinearity between the factors. Hence, there is no problem with multicollinearity.

Table 10
Variance Inflation Factor Test

Variable	VIF	1/VIF
Market	1.180	0.850
SMB	1.240	0.804
HML	1.310	0.765
Mom	1.270	0.790
Mean VIF	1.250	

Notes: Presented are the results from a variance inflation factor (VIF) test. Values of VIF>10 or 1/VIF<0.1 would indicate a problem with multicollinearity. Thus, the results do not indicate problems with multicollinearity.

Table 11
Heteroscedasticity Test

Portfolios	Breusch-Pagan	White
Clean portfolio	Homoscedasticity	Homoscedasticity
Dirty portfolio	**Heteroscedasticity**	**Heteroscedasticity**
Difference portfolio	Homoscedasticity	**Heteroscedasticity**

Notes: Presented are the results from estimation of heteroscedasticity through Breusch-Pagan and White test. In order to make results and estimations as accurate and robust as possible, correction for heteroscedasticity was made by using Newey-West standard errors. Standard errors were estimated through Newey-West with one lag consistently throughout the thesis.

1. f. Residual Analysis

In order to test for serial correlation among the residuals, Durbin-Watson and Breusch-Godfrey tests have been conducted. The results indicate that the clean portfolio exhibits serial correlation while the dirty portfolio and the difference portfolio do not. In order to correct for serial correlation, Newey-West standard errors with one lag have been utilized throughout the thesis.

Furthermore, a visual inspection of the residuals has been made for all the regressions in order to verify that they are normally distributed.

1. g. Robustness and Sensitivity Analysis

This study has not had any reason to include an analysis of attrition bias due to the fact that there was no attrition among the portfolios.

When removing the best and the worst performing ETF from the clean and dirty portfolio respectively, the results in table 12 are generated from the regression.
Clearly, the results are in essence the same as before. The variables that were statistically insignificant before this adjustment are still statistically insignificant and vice versa.

From a visual inspection of the data and graphs, the ETF KWT US stands out. It seems to have had a drastic decrease in price, which might have caused some bias in the results. When excluding the ETF KWT US from the clean portfolio and running the regression, the results in table 13 are generated. As can be seen from the table, the results are in essence the same. Thus, the conclusions drawn before this modification can be regarded as robust, and not sensitive to individual ETFs to a degree that an alternative conclusion would have been drawn.

Table 12

Variables	*Clean ETF portfolio*	*Dirty ETF portfolio*	*Difference portfolio*	*S&P500 Index*
Four-factor alpha	-0.123*	-0.119*	-0.004	-0.052*
	(0.029)	(0.032)	(0.042)	(0.003)
Market	1.260*	0.740*	0.520*	0.995*
	(0.032)	(0.044)	(0.048)	(0.003)
SMB	0.563*	0.019	0.544*	-0.137*
	(0.063)	(0.081)	(0.094)	(0.005)
HML	-0.064	0.296*	-0.360*	0.007
	(0.070)	(0.109)	(0.122)	(0.007)
Mom	-0.255	-0.638*	0.383*	0.000
	(0.045)	(0.067)	(0.077)	(0.005)
Observations	1319	1319	1319	1319

Notes: Presented are estimates for each portfolio and the benchmark index during the examination period January 2011–March 2016 where the best and worst performing ETF from each portfolio are removed. The estimates are the results from regressing daily excess return on the daily factor returns. Newey-West standard errors, correcting for serial correlation and heteroscedasticity, are in parentheses. The difference portfolio is created by subtracting the daily returns of the dirty ETF portfolio from the daily returns of the clean ETF portfolio. The OLS estimation was made using the following equation:

$$R_{it} - R_{ft} = \alpha_i + \beta_{1i}(R_{mt} - R_{ft}) + \beta_{2i}SMB_t + \beta_{3i}HML_t + \beta_{4i}Mom_t + \varepsilon_{it} \quad (5)$$

Where R_{it} is the portfolio return, R_{ft} is the risk-free rate, $(R_{mt} - R_{ft})$ is the excess return of the market, SMB_t captures the size effects, HML_t captures book-to-market effects, Mom_t captures the momentum effects

* Significant at the 1 % level.

Table 13

Variables	*Clean ETF portfolio*	*Dirty ETF portfolio*	*Difference portfolio*	*S&P500 Index*
Four-factor alpha	-0.118*	-0.119*	0.010	-0.052*
	(0.028)	(0.032)	(0.037)	(0.003)
Market	1.250*	0.740*	0.374*	0.995*
	(0.032)	(0.044)	(0.041)	(0.003)
SMB	0.595*	0.019	0.516*	-0.137*
	(0.062)	(0.081)	(0.081)	(0.005)
HML	-0.116	0.296*	-0.407*	0.007
	(0.070)	(0.109)	(0.103)	(0.007)
Mom	-0.252*	-0.638*	0.393*	0.000
	(0.043)	(0.067)	(0.066)	(0.005)
Observations	1319	1319	1319	1319

Notes: Presented are estimates for each portfolio and the benchmark index during the examination period January 2011–March 2016 where the ETF KWT US from the clean portfolio is removed. The estimates are the results from regressing daily excess return on the daily factor returns. Newey-West standard errors, correcting for serial correlation and heteroscedasticity, are in parentheses. The difference portfolio is created by subtracting the daily returns of the dirty ETF portfolio from the daily returns of the clean ETF portfolio. The OLS estimation was made using the following equation:

$$R_{it} - R_{ft} = \alpha_i + \beta_{1i}(R_{mt} - R_{ft}) + \beta_{2i}SMB_t + \beta_{3i}HML_t + \beta_{4i}Mom_t + \varepsilon_{it} \quad (5)$$

Where R_{it} is the portfolio return, R_{ft} is the risk-free rate, $(R_{mt} - R_{ft})$ is the excess return of the market, SMB_t captures the size effects, HML_t captures book-to-market effects, Mom_t captures the momentum effects

* Significant at the 1 % level.