



UNIVERSIDADE CATÓLICA PORTUGUESA

To “sin” or to be ethical?

Comparing the performance of FTSE4Good index with Sin Stocks

Ana Rita Fragateiro

Católica Porto Business School

2020



UNIVERSIDADE CATÓLICA PORTUGUESA

To “sin” or to be ethical?

Comparing the performance of FTSE4Good index with Sin Stocks

Final work in the form of Dissertation
presented to the Portuguese Catholic University
to obtain a master's degree in Finance

By
Ana Rita Fragateiro

Under the guidance of Professor
João Filipe Monteiro Pinto

Universidade Católica Portuguesa – Porto
Católica Porto Business School
Porto, March of 2020

Acknowledges

First, I would like to express my deepest thanks to my supervisor, Professor João Filipe Monteiro Pinto and also to the Professors João Novais and Ricardo Ribeiro for accompanying me on the last steps of the Master. Their effort to answer my questions and their helpful guidance and comments during the process of this work, made possible the successful completion of this Dissertation.

I would like to express my sincere gratitude to my family, specially my parents and my brother and also my friends who were patient and showed endless support throughout this journey.

A thanks to my colleagues who accompanied me throughout this Master and who supported me through this process.

Finally, I also would like to thank my teachers who, during my time as a student at Católica, contributed to the development of my knowledge and for the impact they had in my academic life.

Abstract

This study focuses on the differences between being a socially responsible investor and a vice investor by investigating the performance of the FTSE4Good index and compare it with the performance of a “sin” portfolio. The dataset includes stocks from the European and American market over the period of 2001-2019. The results reveal that FTSE4Good portfolios outperform when compared with appropriate benchmarks. Additionally, by employing different models, the findings point out to a superior performance in the European “Sin” portfolio. Consistent with previous evidence, the results do not show an underperformance or outperformance in the difference of these portfolios, when using models such as the CAPM, the Fama and French three-factor model (1993), the Carhart (1997) four factor model and the Fama and French (2015) five-factor model.

Key words: Sin investments, Socially responsible investments, Performance evaluation, FTSE4Good, CAPM, Fama and French models, Carhart’s model

Resumo

Neste estudo, o principal foco é analisar as diferenças entre um investidor socialmente responsável e um investidor que investe em stocks “sin”, ao comparar a performance do índice FTSE4Good com a performance de um portfólio “sin”. Os stocks analisados pertencem ao mercado Europeu e ao mercado dos EUA, durante o período de 2001 a 2019. A análise aos portfólios FTSE4Good revelam um desempenho superior quando comparados com outros comparáveis. Adicionalmente, ao utilizar diferentes modelos, os resultados apontam para um desempenho superior do portfólio “Sin” da Europa. Consistente com estudos anteriores, os resultados não mostram um desempenho inferior ou superior na diferença desses portfólios, quando aplicados modelos como o CAPM, o modelo de três fatores de Fama e French (1993), o modelo de quatro fatores de Carhart (1997) e o modelo Fama e French de cinco fatores (2015).

Palavras-chaves: Investimentos sin, Investimentos Socialmente Responsáveis, avaliação do desempenho, FTSE4Good, CAPM, modelos de Fama e French, modelo de Carhart

Index

Acknowledges	i
Abstract	ii
Resumo	iii
Index of tables	vi
List of Abbreviations	vii
Introduction	8
1. Literature Review	12
1.1 Portfolio Performance Evaluation Models	12
1.1.1 CAPM.....	12
1.1.2 APT.....	16
1.1.3 Fama and French Three-Factor Model.....	18
1.1.4 Carhart Four-Factor Model.....	19
1.1.5 Fama and French Five-Factor Model.....	20
1.2 Recent trends on investing	20
1.2.1 SRI.....	20
1.2.2 Sin Stocks.....	25
1.2.3 SRI Versus Sin Stocks.....	27
2. Methodology and Data	30
2.1 Methods.....	30
2.2 Data	33
2.2.1 FTSE4Good portfolios	33
2.2.2 Sin portfolios.....	34
2.2.3 Summary statistics with excess returns	36
3. Results	39
3.1 Results of the FTSE4Good portfolios	39
3.1.1 Discussion of the FTSE4Good portfolios	41
3.2 Results of the Sin portfolios	44
3.2.1 Discussion of the Sin portfolios.....	46
4. Analysis in recession and expansion periods	50
4. 1 Analysis in the American market	51

4.2. Analysis in the European market	62
5. Conclusion	72
References	75

Index of tables

Table 1: Summary statistics of the variables of the equations: Europe.....	36
Table 2: Summary statistics of the variables of the equations: US.....	37
Table 3: Performance measurement of FTSE4GOOD portfolio in the US with different performance models: July 2001–September 2019, 218 months.....	39
Table 4: Performance measurement of FTSE4GOOD portfolio in Europe with different performance models: July 2001–September 2019, 218 months.....	40
Table 5: Performance measurement of sin portfolio in the US with different performance models: July 2001–September 2019, 218 months.....	44
Table 6: Performance measurement of sin portfolio in the Europe with different performance models: July 2001–September 2019, 218 months.....	45
Table 7: Summary statistics in expansion periods in the US.....	51
Table 8: Summary Statistics in recession periods in the US.....	52
Table 9: Descriptive statistics in recession periods in the US: FTSE4Good portfolio and Sin portfolio.....	53
Table 10: Descriptive statistics in recession periods in the US: Difference portfolio	54
Table 11: Descriptive statistics in expansion periods in the US: FTSE4Good Portfolio and Sin Portfolio	57
Table 12: Descriptive statistics in expansion periods in the US: Difference Portfolio	58
Table 13: Summary statistics in expansion periods in the Europe.....	62
Table 14: Summary statistics in recession periods in Europe.....	63
Table 15: Descriptive statistics in recession periods in Europe: FTSE4Good portfolio and Sin portfolio.....	64
Table 16: Descriptive statistics in recession periods in Europe: Difference portfolio.....	65
Table 17: Descriptive statistics in expansion periods in Europe: FTSE4Good portfolio and Sin Portfolio.....	68
Table 18: Descriptive statistics in expansion periods in Europe: Difference portfolio.....	69

List of abbreviations

SRI	Socially Responsible Investing
US SIF	US Social Investment Forum
MktRF	Market minus the risk-free rate
SMB	Small minus Big
HML	High minus Low
RMW	Robust minus Weak
CMA	Conservative minus Aggressive
SIC	Standard Industry Classification

Introduction

In general, investors' decisions are taken with the main goal of maximizing their return for a certain level of risk. However, these decisions depend on a diversity of factors such as investors' budget, investors' risk aversion, investors' ethical profile and other criteria.

Socially Responsible Investments (SRIs) are investments that seek to combine diverse criteria such as environmental, social, and corporate governance criteria, as mentioned by Bello (2005), Barnett and Salomon (2006) and Becchetti et al. (2015). According to the 2016 US SIF trends report, these investments may have the following consequences: long-term financial returns and a positive social impact. This approach to investing started being used in the seventies, in which one of the first studies was made by Moskowitz (1972), but only in recent years, SRIs have become more popular. SRIs' screens depend on the investor's criteria. There are: negative screens (excludes companies from industries such as tobacco, alcohol, gambling and armaments); positive screens (includes companies with good practices such as treatment of customers, corporate governance, environment and others) and best-in-class (best companies with the best practices in each sector), as mentioned by Renneboog et al. (2008) and Statman and Glushkov (2009).

In the US, SRI mutual fund assets have increased 34 percent since 2016 from \$1.72 trillion to \$2.58 trillion and the total number of SRI mutual funds has increased from 475 to 636, according to the 2018 US SIF report. This report also finds that more than one out of every four dollars under professional management in the US (\$12.0 trillion or more) was invested according to SRI strategies at the end of 2017. In Europe, socially responsible investments increased by 11% in two years (from 2016 to 2018) and reached \$14.1 trillion.

In the first studies on SRIs, most authors focused on analysing SRI funds. For example, Luther et al. (1992) and Luther and Matatko (1994), studied UKs' funds but did not find significant results between the performance of the funds with the benchmark used in each work. The reason that may have led to these conclusions is the incoherence of the results in the performance of these funds comparing with conventional benchmarks, due to the lack of ethical nature and transparency. These results incentivized other researchers to analyse the performance of SRI indices instead of funds.

This study focuses on the analysis of FTSE4Good Index, in particular, the FTSE4Good Europe and FTSE4Good US indices. The FTSE4Good Index was launched in 2001 and adopts "negative screens" to decide which securities are included in their indices. Several firms are excluded due to their involvement in industries that are considered unethical, such as tobacco, nuclear weapon systems, weapon systems and nuclear power stations.

In contrast with SRIs, there are the so called "sin" stocks. The "sin" stocks are primarily from industries such as tobacco, alcohol, and gambling (gaming) stocks, which are industries composed by companies that are usually associated with unethical activities.

The classification of "sin" stocks in the studies about this topic depends on each author. For example, Lobe and Walkshäusl (2014), besides including the three sectors that were previously mentioned, also include adult entertainment, nuclear power, and weapons¹.

According to Richey (2016), the research dedicated to the investment on "sin" stocks remains scant. One of the earliest studies was published by Luck and Wood (1992), who found a significant outperformance of US tobacco and alcohol stocks comparing to gambling stocks. But only in the 2000s, other authors like

¹ The authors Hong and Kacperczyk (2009) focus on the so called triumvirate of Sin that classifies alcohol, tobacco and gambling as "sin" stocks, in contrast with Lobe and Walkshäusl (2014) who considered adult entertainment, alcohol, gambling, nuclear power, tobacco, and weapons as "sin" stocks and called it Sextet of Sin.

Fabozzi et al. (2008), found that “sin” stocks outperformed the market. One of the most mentioned works in this field was published by Hong and Kacperczyk (2009), who found that “sin” stocks in the US have higher abnormal risk-adjusted returns compared to other similar stocks, prove that the “sin” stocks outperform the market. They also show evidence that social norms have in fact, an impact in the markets, influencing the expected returns of these stocks.

Due to the ambiguity of results presented in the empirical works about the SRI portfolios and “sin” stocks, the main purpose of this research is to fill some existing gaps in the literature by investigating the performance of the FTSE4Good index portfolios and “sin” portfolios. In this study, the SRI portfolio is composed by the FTSE4Good index and the “sin” portfolio is composed by “sin” stocks of the following industries: gambling, tobacco, alcohol and defense. A more detailed analysis of the performance of these portfolios will provide new insights on the comparison between these two different portfolios. In this study, an emphasis on the differences between being a socially responsible investor and a vice investor in terms of the performance and using models such as the CAPM, the Fama and French (1993) three-factor model, the Carhart (1997) four-factor model and the Fama and French (2015) five-factor model, will be examined. Such analysis, employing the Fama and French (2015) five-factor model to compare these two investment strategies, have not been done previously. Additionally, this work analyses the differences of performance between the FTSE4Good portfolio and the “sin” portfolio on a period of recession and a period of economic growth. Finally, a discussion of what motivates investors to choose either one of the portfolios (SRI versus “sin”) will be presented.

Reviewing the literatures of both, vice investing and socially responsible investing, it is suggested in some researches, that “sin” stocks are likely to have superior returns comparing to socially responsible stocks. Although this hypothesis cannot be ignored, it is important to note that Lobe and Walkshäusl (2014), by using a hedge portfolio approach, did not find an underperformance

or outperformance. This study confirms Lobe and Walkshäusl (2014) results that do not support the outperformance of “sin” portfolios vis-à-vis FTSE4Good portfolios.

Through the work, it is analysed the “sin” portfolios and FTSE4Good portfolios separately, and when comparing with appropriate benchmarks and controlling for the different factors, it is found an outperformance in the FTSE4Good portfolios and in the “Sin” Europe portfolio.

This study proceeds as follows. Section 1, discusses the relevant literature, mentioning the different Portfolio Performance evaluation models and the recent investing trends such as SRI and “sin” stocks. In section 2, it is described the methodologies, explaining the different models that will be tested such as the CAPM, the Fama and French three-factor model, the Carhart four-factor model and the Fama and French five-factor model. It is also described the sample selection. Section 3 presents the results and discusses those results. Section 4 presents the results considering two sub-samples for expansion and recession periods. Section 5 concludes the study.

1. Literature Review

1.1 Portfolio Performance evaluation models

In the research, the focus will be on the performance of two different portfolios, therefore it is essential to understand how the performance evaluation of portfolios is proceeded. In the first part of this section, it will be mentioned the different models and measures to evaluate the performance of portfolios.

One of the studies that have completely revolutionized the way investors choose its securities is the Modern Portfolio Theory (MPT), developed by Markowitz (1952), in which the author explains how investors could construct portfolios through diversification in order to optimize the expected return given a certain level of market risk. Although the author states that diversification cannot eliminate all the risk due to the complexity and the inter-correlation between securities, Markowitz (1952) also quantifies the concept of portfolio diversification by introducing the concepts of covariance, return and volatility. The author argues that the investors should choose mean-variance efficient portfolios, which are portfolios that provide the highest expected return on a given level of variance or the minimum variance for a given expected return. These efficient portfolios are obtained by solving the mean-variance optimization problem with different values of the risk-aversion coefficient.

1.1.1 CAPM

The previous theory is considered a cornerstone to the models that came afterwards, such as the Capital Asset Pricing Model (CAPM) developed by Jack Treynor (1961), William Sharpe (1964) and John Lintner (1965). In a world, where investors are risk-averse, they will prefer to choose an investment with a lower risk and they will only accept more risk, if there is additional compensation.

CAPM is a single factor model that describes and gives the prediction of the relationship between the risk of an asset and its expected return, by assuming that the expected return only depends on the asset's systematic risk. CAPM is widely used because of its insight although so far it is not fully proven empirically.

This model follows the following assumptions: i) Investors are price-takers, ii) Investors have the same decision horizon, iii) There are no taxes and transactions costs, iv) Investors maximize the expected utility depending on expected returns and variance, v) Investors agree on the expectations of its assets, vi) Existence of a risk-free asset, vii) Investors can borrow and lend at the same rate.

The first published work about this model was developed by Sharpe (1964), in which the author tried to solve the problems caused by the lack of studies about the conditions of risk, in the literature. Sharpe (1964) shows that the rates of return from efficient combinations would be perfectly correlated and the investor would be able to eliminate all the risk through diversification except for the systematic risk that would remain even in efficient combinations.

Sharpe (1964) also states that the standard deviation measures the risk associated with the assets. However, in the following year, Lintner (1965) publishes a different statement by considering variance and covariance with other securities, the most important measures of the risk of a security within a portfolio of risky assets.

Lintner (1965), once again mentions the importance of finding the optimal portfolio. The author uses regressions of annual rates of returns for 301 large industrial companies in a period of ten years (1953-1963) and then the regressions of 70 large mutual funds (listed in Weisenberger's Investment Trusts 1953-1963). The author concludes that the risks associated with the given expected returns cannot be eliminated, but only reduced through selection and diversification.

Comparing the models of Sharpe (1964) and Lintner (1965), both models are based on the same set of assumptions and the equilibrium is the same. However,

the second author considers that the risk measure derived from his model is more general and different than the one proposed by the first author. But later, Fama (1968) realizes that both papers state the same, therefore there is no conflict between the two authors' models. Fama (1968) concludes that both authors made errors in the equations of risk premium and that the major difference between them was the residual variances. Although Sharpe (1964) does not intend for the residual variance to be zero, Lintner (1965) makes the previous assumption, which is the cause of inconsistency.

Another author that focuses his studies on the CAPM is Treynor (1961, 1962). In his work, Treynor (1961, 1962) explains the relationship between the expected return and covariance with the market portfolio, by developing the mean-variance formulas. The author concludes that the market is the single optimal mean-variance efficient portfolio. The author also concludes that the securities held by investors were proportional to the number of shares available in the market.

Jensen (1968, 1969) tests the CAPM. In order to evaluate the performance of risky assets, the author develops a model using the Sharpe-Lintner model results and tries to propose a solution to CAPM, in a situation where the markets have privileged information. The author uses a new measure of performance called alpha, which represents the difference between the effective and the expected return of the asset, given the portfolio's beta which also depends on the systematic risk and the risk-free rate. In order to realize the empirical estimation of the market model and the systematic risk, the author uses a sample that included portfolios of 115 open-end mutual funds from the Wiesenberger's Investment Companies over the period of 1955-1964. The author concludes that, the returns of the funds are consistent with CAPM and that there is a relationship between the measure of portfolio performance and the measure of efficiency.

Although Jensen's alpha is the most used performance measure, it is also the most controversial due to the possibility of identifying a wrong market timer.

Despite of its current use, critics believe that the market is efficient and state that superior returns derive from luck. Another study was developed by Grinblatt and Titman (1989), in which the authors introduce a new measure, the positive weighting measure. One of the solutions presented to solve the problem was to identify investors with all the correct information as positive performers, demonstrating that the mean-variance efficient portfolio is the most suitable portfolio. The empirical tests show that both Jensen's alpha and positive weighting measure are similar, but the second measure is not used because it is not intuitive and due to the complexity of the computations to determine the weights.

Despite of being the most used model, Roll (1977) presents a critique of CAPM. The author argues that this model is untestable, due to the following reasons: i) this model only predicts that the market portfolio is mean-variance efficient; ii) the existence of the linear relation between expected returns and beta could be found in any sample, independently of how returns were determined in the market; iii) the existing tests in the literature did not add new information about the efficiency of the market portfolio and iv) the composition of the market portfolio is not known. Roll (1997) concludes that it is only possible to test CAPM if all the information about the individual assets that compose the sample is available and even using proxies for the market can cause different inferences.

Another critique to CAPM is presented years later by Admati and Ross (1985). The authors investigate a method to evaluate a portfolio's performance, considering that a superior performance is the result of better information. It is shown the importance of the asymmetries of information and heterogeneity of beliefs in the markets and how to measure performance. This contradicts previous works, where it is assumed that the beliefs are homogeneous in the context of the CAPM.

Jagannathan and Wang (1996), instead of using a constant beta, they assume a conditional CAPM where betas and risks vary over time and it is included the

return on human capital to measure the return on aggregate wealth. The authors claim that the lack of support on CAPM by the empirical research is due to the fact of being a one-period model, whereas the investors' periods are longer in the real world. In this study, Jagannathan and Wang (1996) vary beta because the relative firm's risk is likely to vary. As a proxy for the market portfolio, they considered the CRSP'S value-weighted index. In their results, by allowing these variations in the CAPM, it is shown, using a US sample over the period 1962-1990, that this model succeeds in explaining the average stocks returns and the size effects.

CAPM assumes that the investors invest in the portfolio with higher returns but more recently, Frazzini and Pedersen (2014) widen the analysis across markets by constructing "betting against beta" (BAB) portfolios which buys low-beta stocks and sells high-beta stocks. Their data is primarily focused on the US market over the period from 1926 to 2012. They find evidence of underperformance of high beta stocks, confirming the presence of a beta anomaly in the US stock market. It is documented in the US stock market and in 18 of 19 international markets, that when beta increases in the portfolios, the alpha declines. Furthermore, the authors show that when the availability of funding is reduced, the BAB strategy should lead to losses. The previous statement could be explained by the increase in the required return and investors might need to rebalance their BAB positions, in the future.

1.1.2 APT

As Admati and Ross (1985), the authors Dybvig and Ross (1985) question the veracity of the CAPM, which is only valid if the market is efficient and the information is homogeneous. In their paper, the authors test the Asset Pricing Theory (APT) proposed by Ross (1976), where the authors try to reduce the

confusion between the theoretical model and the empirical tests and explain the relationship between the APT and the CAPM. APT is a model composed by factors that is used to obtain the returns, but the number of factors is not specified, which leads to derived models with certain assumptions that are approximations. The authors state that the models derived from APT cannot be held because these assumptions are not appropriate to use in arbitrary portfolios, which happens because of the existence of a proxy error in the models. Comparing APT and CAPM, the authors conclude that CAPM implies APT with the existence of a factor structure.

As previously mentioned, Ross (1976), who proposed the model of capital asset pricing (APT), explains the cross-sectional variation in asset returns and tries to eliminate the CAPM's unrealistic assumptions. APT may appeal to researchers because this model includes several risk premia as a compensation for enduring risk in contrast with the CAPM, where there is only one risk premium. APT is also applicable to a subset of assets. Ross (1976) demonstrates that the existence of a linear pricing relation is a necessary condition to achieve equilibrium.

As stated before, in the APT model, the number of factors is not specified, so Roll and Ross's (1980) purpose is to use the APT framework to calculate the assets prices, in order to find empirically at least three and probably four factors (although the evidence is less conclusive) in the dataset of equities traded on the New York and American Exchanges that are priced from 1962 to 1972. However, Roll and Ross (1980) do not offer an economic interpretation of these factors and admit that their test is weak but on the other side, they conclude that the APT is the best model to measure the expected returns.

Chen (1983) estimate the parameters of Ross's (1976) APT and compare the evidence on APT and CAPM. The sample used in this study is composed by daily returns over the period from 1963 to 1978. Based on the evidence gathered, APT can better explain the expected returns comparing to CAPM and the author

concludes that variables such as own variance and firm size do not have additional explanatory power.

In conclusion, the previous studies show how the APT is a substitute for the CAPM. The APT only requires utility assumptions and it is hold in both single period and multiperiod, opposed to CAPM that can only be held in a single period. Unlike the CAPM, there is no requirement for the market portfolio to be mean variance efficient. The major differences between the CAPM and APT are: i) APT allows more than one generating factor; ii) In the APT, there is a linear relationship between expected returns and its loading on common factors.

1.1.3 Fama and French three-factor model

Previous studies had shown that there are other factors than can explain average returns besides beta, such as size, book-to-market, leverage and earning-prices. Fama and French (1992) evaluate size and book-to-market equity in the cross-section of average returns on NYSE, AMEX, and NASDAQ stocks during the 1963-1990 period. They conclude through their tests, that the evidence found do not exhibit a relation between average returns and market betas, contradicting Shaper (1964) and Lintner's (1965) results. For the 1963-1990 period, it is shown that the factors, size and book-to-market equity capture the cross-section of average stock returns. This leads to the following conclusions: i) beta does not seem to explain the cross-section of average stock returns because of the lack of information about these returns and ii) the combination of size and book-to-market equity seem to have an important role on the average returns because these factors seem to be able to absorb the effects of the earning-prices and leverage, during the period of the sample used.

In the following year, Fama and French (1993) deepen their previous study and identify five common risk factors in the returns on stocks and bonds, in

which the three stock-market factors are: market, size and book-to-market. In this model, the authors analyse 25 portfolios and capture the relation between average returns and size and the relation between average returns and price ratios. By using the time-series regression developed by Black, Jensen and Scholes (1972), the authors show that size and book-to-market are a proxy for sensitivity in returns. Another conclusion is that, by including the excess market returns, it is shown that the market factor and the proxies for the risk factors related to firm size and book-to-market equity are able to explain the cross-section of average returns.

1.1.4 Carhart four-factor model

Carhart (1997) develops a four-factor model using Fama and French's (1993) three-factor model and adds a one-year momentum factor, in order to explain the average returns. In this study, the author shows that a mutual fund performance can be explained by the investment style chosen by the managers of the fund. The data of this study is composed by U.S equity funds over the period from January 1962 to December 1993. The results imply that, the common-factor sensitivities, expenses and transaction costs capture the predictability of the expected returns of the funds. The author also concludes that in order to maximize the returns, the investor should not invest in funds that present a poor performance, where transaction costs and other expenses have a negative impact in performance. This model exhibits smaller price's errors than comparing to the CAPM and the Fama and French (1993) three-factor model.

1.1.5 Fama and French five-factor model

Fama and French (2015) developed another model: the five-factor model that expands the Fama and French (1993) three-factor model and adds two new factors: operating profitability and investment. The authors decide to add the profitability factor due to the higher expected earnings that imply higher expected returns. The investment factor indicates that higher expected growth in book-to-market equity implies a lower expected return. One of the main problems of the Five-Factor model is its difficulty to capture the low average returns on small stocks and the model's performance is not sensitive to the way its factors are defined. However, this model was more successful in capturing the returns of US firms over than period 1963-2013 than comparing to the Fama and French (1993) three-factor model. Comparing with the Carhart (1997) four-factor model, this model does not apply alternative factors. The variable momentum was not included in Fama and French (2015) five-factor model, because of the possible existence of bad diversification in some of the portfolios.

1.2 Recent trends on investing

1.2.1 SRI

Akerlof (1980) states that a social norm is dependent on the beliefs of the individuals and these beliefs will affect the choice made by the individual who makes the decision. Therefore, when choosing an investment strategy, social norms may impact the economic behaviour. The focus of this study is the SRI (Socially Responsible Investing), more specifically the FTSE4GOOD index and vice investing or investing in "sin" stocks, which are two different strategies of investment on an ethical level.

Prior literature on socially responsible investments has shown that by excluding "sin" stocks (negative screens) from a portfolio will reduce

performance and increase risk, while investing in stocks of firms with positive social responsibility scores (positive screening) will improve performance and reduce risk.

The empirical analysis of SRI funds date back to 1972, in which Moskowitz studies the research question whether social issues should be considered when making investments and analyses companies with positive responses to social problems, assuming that social responsibility is an important factor for the investor.

Most of the early studies on the performance of the SRI investments funds do not consider the costs of these funds. These studies also tend to only use regressions based on one or two-factor model, which may neglect other factors such as size, which can influence the performance results. Many authors started analysing ethical funds, such as Luther et al. (1992) who analyses British ethical funds, Hamilton et al. (1993), Goldreyer and Diltz (1999), Statman (2000), and Bello et al. (2005), which analyse US ethical funds and White (1995), who focuses on both German and US ethical funds. These studies, by using a simple regression and performance measures as Jensen's alpha against a market index, were not able to find significant results on the underperformance or outperformance of SRI funds comparing to conventional funds. On the other hand, Luther and Matatko (1994) instead of just comparing ethical funds' performance in the UK with a market index, as made previously by Luther et al. (1992), the authors try to find the most adequate benchmark and compare the performance with a small company index. Over the period from 1985 to 1992, using the CAPM, they find a neutral performance, independently of the benchmark used.

Mallin et al. (1995) and Gregory et al. (1997) also analyse the ethical funds of the UK market and they were the first authors to apply a matched pair approach. Both studies conclude that SRI and non-SRI funds have a similar performance and that the ethical funds present a lower risk comparing to non-ethical.

Instead of just analysing one country, Kreander et al. (2005) focus their research on four European countries. The authors, which also follow the matched pair approach (matching size and age), analyse ethical funds of the European market over the period of 1995-2001 and compare the performance of SRI mutual funds and non-SRI funds using Jensen, Treynor and Sharpe measures. This research once again shows evidence of a similar performance of the SRI funds and the non-SRI funds.

Statman (2000) focuses on the US market and compares the performance of 31 other US ethical mutual funds with the performance of 62 non-ethical funds. The author also compares the performance of the ethical funds with the Domini Social Index, over the period of 1990-1998, using the Jensen's alpha and the Sharpe measures. The results showed no significant differences between the performance of the SRI funds and the non-SRI funds.

Authors started using more complex models in order to evaluate the performance of mutual funds, such as Bauer et al. (2005), who applied a multi-factor model to investigate the performance of 103 British, German and US mutual funds and 4384 conventional mutual funds over the period from 1990 to 2001. The authors employ the Carhart (1997) four-factor model and their results indicate an underperformance of German and US SRI mutual funds comparing to their relevant indices and the conventional funds, and an outperformance of the UK funds. The authors conclude that both ethical and conventional funds present a similar performance in the difference in return due to the statistically insignificant results found in their sample.

Scholten (2005) investigates the performance of 12 Dutch SRI funds, over the period from 2001 to 2003. The author employs the CAPM and the Carhart (1997) four-factor model and finds that, although these funds outperform conventional funds, the difference is not statistically significant. Another important conclusion, taken from this study, is that it is found an advantage in using SRI

indexes (e.g FTS4Good Global, DJSI STOXX) than conventional indexes (e.g DJ STOXX, AEX, DJ World).

Cortez et al. (2009) also find evidence of similar performance for socially responsible funds comparing with conventional and socially responsible benchmarks, in seven European countries. In this study, the authors analyse the performance of a sample of 88 socially responsible funds over the period of 1996 to 2007, using unconditional (CAPM) and conditional approaches (following the approach of Ferson and Schadt, 1996). They conclude that socially responsible funds are more exposed to conventional indices and that the conventional benchmarks are the more appropriate to explain the funds returns comparing to the socially responsible benchmarks. This study shows that the performance of the European socially responsible funds is similar to the performance of conventional or socially responsible benchmarks.

Instead of analysing investment funds as in the previous studies, Schröder (2007), analyses SRI indices, including the FTSE4Good indices. By analysing 29 international SRI indices and using the CAPM to estimate the performance through Jensen's alpha, the author demonstrates that there is no significant evidence of under or over performance. It is important to note that, according to this author, there is no need to employ models such as Fama and French (1993) three-factor model and Carhart (1997) four-factor model. The high adjusted r-squared values, unspecific investment styles and the infrequent update of the indices are some of the reasons why the author justifies that the use of the CAPM with an appropriate benchmark will be a close approximation of the investment universe.

Other authors have also concluded that the relationship between the returns of stocks of socially responsible companies and the returns of stocks of conventional companies can be better analysed by comparing indexes of stocks of socially responsible companies to indexes of stocks of conventional companies. Sauer (1997) and Statman (2006) compare the performance of the DS 400 Index

with the S&P 500 Index and both present a neutral performance. Sauer (1997) estimates Jensen's alpha whereas Statman (2006) uses Fama and French's (1993) three-factor model. Statman (2006) not only analyses the performance of the DS 400 Index but also focuses its studies on exploring the characteristics that define socially responsible companies. By using the Fama and French (1993) three-factor model and comparing the returns of the four SRI indexes to the returns of the conventional S&P 500 Index, the results show that returns of the DS 400 Index are higher than those of the S&P 500 Index during the sample period of 1990-2004, but not in every sub-period. In general, SRI indexes had better results than the S&P 500 Index during the boom of the late 1990s than during the early 2000s, which demonstrated that the returns of the SRI indices exceeded the returns of the S&P 500, but the results were not statistically significant.

Collison et al. (2008) analyse the financial performance of the global FTSE4GOOD indices, using a sample for the 1996-2005 period. The authors calculate daily returns and use Sharpe, Treynor and Jensen performance measures. Results suggest that returns of a portfolio of companies included in FTSE4Good indices are not lower when compared to the performance of portfolios that include stocks of companies that do not follow a social responsible strategy. The authors find an outperformance of these indices when comparing to the benchmarks.

One recent study by Belghitar et al. (2014), focus on comparing the performance of the FTSE4Good index of four geographic regions (Europe, US, UK and Global) with conventional indexes, over the period of July 2001 to November 2010. The authors analyse the performance by using the MCSD theorem and the mean-variance analysis (by calculating Sharpe ratio, Treynor ratios and Jensen's alphas). They conclude that, there is a financial price to be paid for socially responsible investing and risk-averse investors can improve their returns by reducing their investments on SRI stocks and buying conventional stocks.

1.2.2 Sin Stocks

“Sin” portfolios offer investors a different investment strategy comparing to SRI portfolios. “Sin” portfolios are composed by stocks of industries such as alcohol, tobacco and gaming that are considered undesirable, given their social norms, although the returns may be higher than comparing to SRI portfolios.

One of the first studies about “sin” stocks, was published by Merton (1987), in which the author develops a “neglected” stocks model that states that the smaller is the investor base, the higher will be the expected returns.

Salaber (2007) focuses his studies on the “sin” stocks. The author studies the behaviour of “sin” stocks’ returns over the business cycle, using a sample of 183 US “sin” stocks over the period 1926-2005. By applying conditional models, the author finds evidence that “sin” stocks earn excess returns comparing to the market. However, it is shown that the excess returns disappear when “sin” stocks are compared to a portfolio of stocks of the same industry, with similar characteristics. The author also states that the time-varying risk premiums may be the reason of why “sin” stocks outperform in long-term. Another conclusion is that risk premiums of “sin” stocks are higher during recession periods, which leads to the conclusion that “sin” stocks outperform during recession periods and underperform during expansion periods.

Another similar study was developed by Fabozzi et al. (2008) who analyse a sample of “sin” stocks (alcohol, tobacco, defense, biotech, gaming, and adult services industries) of 21 countries over the period of 1970-2007 and apply the CAPM. They show evidence that “sin” stocks outperform the benchmark. The authors explain that the higher excess returns in the “sin” stocks may be due to these stocks not complying with social norms.

Among the authors who focused their research on the “sin” stocks, one of the most important papers is by Hong and Kacperczyk (2009) who investigate the effect of social norms on markets. They analyse “sin” stocks (companies involved

in the production of tobacco, gambling and alcohol, also known as Triumvirate of Sin) of a sample of 156 US stocks over the period of 1962-2006. The authors identify the “sin” stocks using a NAICS/SIC code. In order to measure the performance of the SIN-COMP portfolio (long in “sin” stocks and short in comparable stocks), they use different models including a four-factor model which shows evidence that a portfolio of “sin” stocks, significantly outperforms comparable stocks. These results imply that investors with SRI stocks may be at a disadvantage and investing only on SRI stocks may not be the best option. In their research, the authors also document that “sin” stocks might be under-priced, due to exhibiting a high book-to-market ratio.

In recent years, there have been more studies revolving around “sin” stocks. Richey (2016) examines the risk-adjusted returns of a portfolio of firms from industries such as alcohol, tobacco, gambling and defense (vice portfolio), in the US market. Over the period of 1995-2015, the author employs the Carhart (1997) four-factor model and the Sortino ratio and concludes that the vice portfolio outperforms the market.

In the following year, the same author, Richey (2017), analyses the performance of a “sin” portfolio (includes stocks of alcohol, gambling, tobacco, defense, adult entertainment and payday lenders firms) from the US market over the period from 1987 to 2016. Comparing with its previous study, Richey (2017) also considers adult entertainment and payday lenders firms as unethical or “sin” industries. In this study, the author applies CAPM, Fama and French (1993) three-factor model, Carhart (1997) four-factor model and Fama and French (2015) five-factor model. Richey (2017) finds a positive and significant alpha which is an indicator of the existence of abnormal returns in all the models, except in the Fama and French (2015) five-factor model. This indicates that the evidence of a superior performance disappears.

In the same year, Blitz and Fabozzi (2017) analyse the returns of “sin” stocks from different parts of the world (US, Europe and Global), over the period of

1990-2016. The authors find an outperformance when employing the CAPM specification. Although the authors apply other models, alpha turns statistically insignificant in the other regressions, when controlling for the factors such as size, value, momentum, profitability and investment.

1.2.3 SRI Versus Sin Stocks

Although the studies that compare “sin” stocks and SRI is still scant, the existing literature show some promising conclusions.

Similar results as in Hong and Kacperczyk’s (2009) study, were obtained by Chong et al. (2006), who compare a Domini Social Equity Fund (SRI fund) and a Vice Fund (vice investing fund) and use S&P 500 as a benchmark. The dataset is composed by 758 observations over the period of 2002-2005. This study demonstrates by using performance measures such as Jensen’s alpha, Sharpe ratio and a less common used measure based on an autoregressive conditional heteroskedasticity (ARCH) model, that the Vice Fund outperforms both Domini Social Equity Fund and the benchmark.

Statman and Glushkov (2009) examine the returns of SRI portfolios constructed based on SRI characteristics and the returns of “sin” portfolios of industries such as tobacco and firearms, for the US market. The sample of this research is composed by the returns of stocks with SRI characteristics, rated by KLD, over the period of 1992-2007. The stocks that are classified by KLD as “sin” stocks (industries such as alcohol, tobacco, gambling, firearms, nuclear operations and military) are included in the “sin” portfolios. In order to measure the performance, the authors apply the CAPM, the Fama and French (1993) three-factor model and the Carhart (1997) four-factor model. The results of Statman and Glushkov (2009) indicate that the “sin” portfolio outperforms comparing to a portfolio that does not include “sin” stocks (“accepted” stocks). The authors

also find that the expected returns of the socially responsible stocks are lower comparing to the conventional stocks.

Liston and Soydemir (2010) investigate the performance between a “sin” portfolio and a faith-based portfolio (using religious screens) by collecting data over the period of 2001-2007. By using the CAPM, the Fama and French (1993) three-factor model and the Carhart (1997) four-factor model, they find positive alphas in the “sin” portfolio, which indicate that this portfolio outperforms the faith-based portfolio and it also has higher risk-adjusted returns comparing to the faith-based portfolio.

Some years later, Durand et al. (2013), replicate Hong and Kacperczyk (2009) research, by analysing a sample composed by a portfolio with alcohol, tobacco and gambling stocks (“Sinners”) and a portfolio with the constituents of the MSCI KLD400 Social Index (“Saints”), over the period of 1990-2008. When applying the Fama and French (1993) three-factor model, the authors find an outperformance of the portfolio of “Sinners”, comparing to the “Saints” portfolio. However, the alpha turns statistically insignificant when applying the Carhart (1997) four-factor model. Their results also indicate that “Sinners” are under-priced, supporting Hong and Kacperczyk’s (2009) evidence.

In contrast with previous studies, after controlling for managerial skills, transaction costs and fees, Humphrey and Tan (2013) investigate the impact on the performance of positive and negative screenings, over the period of 1966-2010, in the US market. In order, to construct the negatively screened portfolio, the “sin” stocks (stocks of the following industries: tobacco, alcohol, gambling, firearms and defense/weapons) are excluded from the S&P 500 and to construct the positively screened portfolio, the stocks of the S&P 500 with a positive KLD scores are included. In this study, the authors apply the CAPM and the Carhart (1997) four-factor model and find no evidence of an outperformance of portfolios that include “sin” stocks.

Other paper, that compares the performance between “sin” stocks and social responsible stocks was published by Lobe and Walkshäusl (2014). The authors analyse 755 “sin” stocks of six industries that includes adult entertainment, alcohol, gambling, nuclear power, tobacco, and weapons and 31 international social responsible indexes, including the FTSE4GOOD index at a global level over the period of 1995-2007. They calculate the Sharpe ratio and then to evaluate the performance, they use the following models: the CAPM, Fama and French (1993) three-factor model and the Carhart (1997) four-factor model. They show the inexistence of statistically significant evidence in the performance the hedge portfolio (long in “sin” stocks and short in SRI). These results point out that the choice to whether to invest in one or both types of portfolio depends on the profile of the investor.

One of the most recent studies on this topic is by Trinks and Scholtens (2017). The authors study the impact of negative screening on the financial performance. They use a Carhart (1997) four-factor model, in a sample composed by 1934 controversial global stocks over the period of 1991-2012. Instead of using industry classifications to construct “sin” portfolios, which may lead to an incomplete representation of the controversial issues, the authors study the size and the financial performance of controversial stocks from 14 different industries. The industries that are considered controversial in this study are: abortion, adult entertainment, tobacco, alcohol, gambling, pork, stem cells, animal testing, weapons, nuclear power, fur, meat, contraceptives and genetic engineering. In this paper, the authors test the differences between the “sin” (controversial) stocks and the market, concluding that most of the controversial stocks outperform the market. Additionally, while analysing the risk-adjusted returns of the controversial stocks, the authors conclude that excluding these stocks may reduce financial performance.

2. Methodology and Data

2.1 Methods

The aim of the present study is to compare the performance of two portfolios composed by the FTSE4Good index (FTSE4Good US and FTSE4Good Europe) and two portfolios composed by “sin” stocks of the STOXX 600 for Europe and S&P 500 for the US. Different models are estimated by using OLS regressions techniques to analyse monthly historical returns over the period from July 2001 to September 2019 (see section 2.2).

Four asset pricing models will be tested on their ability to measure the performance of the “sin” portfolio and the FTSE4Good index portfolio, namely: the CAPM, the Fama and French (1993) three-factor, the Carhart (1997) four-factor, and the Fama and French (2015) five-factor models.

2.1.1 CAPM

In order to analyse the returns, the first model to be applied is the CAPM:

$$R_{it} - R_{Ft} = \alpha_i + \beta_{1i} (R_{Mt} - R_{Ft}) + \varepsilon_{it} \quad (1)$$

where R_{it} denotes the return of the portfolio i at moment t , α_i is the Jensen’s alpha (Jensen, 1968) of the portfolio i (denotes the abnormal performance of portfolio i), R_{Mt} represents the return of the market benchmark associated to portfolio i in moment t , R_{Ft} represents the risk-free rate at moment t and ε_{it} is the error term. β_{1i} denotes beta that measures the sensibility to the risk of the portfolio i , in which values above (lower) than 1 imply that the portfolio has a higher (lower) risk. This model will allow to determine if the behaviour of portfolio i can be replicated by the benchmark.

Schröder (2007) conducts a spanning test with the following hypotheses:

$$H_0: \alpha = 0 \wedge \beta = 1$$

$$H_1: \alpha_i \neq 0 \vee \beta_i \neq 1$$

If results reject the null hypothesis, it will imply that portfolio i cannot be replicated by the respective benchmark. This would indicate that both portfolio i and the benchmark do not have similar risks and returns. Consequently, it would point to the need of applying a more complex specification.

Although several authors such as Capelle-Blancard and Monjon (2014) apply the CAPM, this specification has some limitations and it may not be the most adequate to explain the portfolio's returns.

2.1.2 Fama and French three-factor model

The second model is the Fama and French (1993) three-factor model, where it is added two new factors: the size of the firm and the book-to-market:

$$R_{it} - R_{Ft} = \alpha_i + \beta_{1i} (R_{Mt} - R_{Ft}) + \beta_{2i} \text{SMB}_t + \beta_{3i} \text{HML}_t + \varepsilon_{it} \quad (2)$$

where the second factor is the small-minus-big (SMB) that denotes the size factor and it is a zero-investment portfolio² with a long position in small cap stocks and a short position in large cap stocks at moment t . The third factor is the high-minus-low (HML) that denotes the value factor and it is also a zero-investment portfolio with a long position in high book-to-market stocks and a short position in low book-to-market stocks at moment t .

² A zero-investment portfolio is a portfolio where the sum of the weights of each stock is equal to zero. According to Fama and French (1993) three-factor model, the SMB portfolio has the following composition: $\text{SMB} = 1/3 (\text{Small Value} + \text{Small Neutral} + \text{Small Growth}) - 1/3 (\text{Big Value} + \text{Big Neutral} + \text{Big Growth})$.

2.1.3 Carhart four-factor model

As in previous literature such as Salaber (2007), Richey (2017), and others, it will be used the Carhart (1997) four-factor model to analyse the returns of the portfolios of the sample. The Carhart's (1997) model is one of the most commonly used models in the literature about performance evaluation and it is an extension of the Fama and French (1993) three-factor model. This model adds the winners-minus-losers (MOM) factor, which is based on the momentum effect of stock returns. According to Jegadeesh and Titman (1993), the momentum factor states that investors can achieve a superior performance by buying stocks that have good performance (winners) and selling stocks that have a worst performance (losers) in the last 3 to 12 months.

$$R_{it} - R_{Ft} = \alpha_i + \beta_{1i} (R_{Mt} - R_{Ft}) + \beta_{2i}SMB_t + \beta_{3i}HML_t + \beta_{4i}MOM_t + \varepsilon_{it} \quad (3)$$

2.1.4 Fama and French five-factor model

The fourth model and last model used in this study is the most recent response of Fama and French (2015) to critics that claim that their previous model is insufficient. The authors add two new factors to the Fama and French (1993) three-factor model: Robust-minus-weak (RMW) that represents the profitability factor and conservative-minus-aggressive (CMA) that denotes the investment factor.

$$R_{it} - R_{Ft} = \alpha_i + \beta_{1i} (R_{Mt} - R_{Ft}) + \beta_{2i}SMB_t + \beta_{3i}HML_t + \beta_{5i}RMW_t + \beta_{6i}CMA_t + \varepsilon_{it} \quad (4)$$

where RMW denotes the difference between the returns of stocks with a high and low profitability (long position in companies that present robust profits and short position in companies with weak profits) at moment t. CMA denotes the difference between returns of stocks with low and high levels of investment at moment t (long position in companies that present a conservative investment

strategy and short position in companies with an aggressive strategy). The coefficients of the factors, as in the previous models denote the portfolio's i exposure to the factor.

2.2 Data

In order to conduct this study, it was selected two SRI indices, the FTSE4Good Europe (FT4GBEU) and FTSE4Good US (FT4GBUS). In addition, two "sin" portfolios for both Europe and US markets were constructed. These portfolios will be used to compare the performance between the two different types of investing.

Following the approach of Lobe and Walkshäusl (2014), it was collected monthly closing prices for the "Sin" portfolios and the FTSE4GOOD Index portfolios. Data was extracted from Thomson Reuters DataStream Database, over the period of July of 2001 to September of 2019.

2.2.1 FTSE4Good portfolios

The FTSE4Good Index Series was created to measure the performance of companies that show strong Environmental, Social and Governance (ESG) practices. Investors who choose to invest in these indices are exposed to firms that present high global corporate responsibility standards and earn competitive returns. The selection of the securities included in the FTSE4Good indices is made through the adoption of "negative screens", in which several firms are excluded due to their involvement in industries that are considered unethical (such as tobacco, nuclear weapon systems, weapon systems, nuclear power stations).

The start date of this study is July 31 of 2001 because the launch date of the FTSE4Good indices was in July 1 of 2001, whereas the finish date is September 30, 2019 due to being the most recent data available in the Kenneth French website, in the moment, the data was obtained.

The prices of the index in dollars of the last day of each month between 31.07.2001 and 31.09.2019 were obtained.

According to the information available on the FTSE website, all the SRI indexes are value-weighted, so to calculate the returns of the FTSE4Good portfolios, it was only used the prices of the indexes of each month over the period of the sample. In the Datastream, it was found two indices for the purpose of this study: the FTSE4Good US (LFT4GBUS) that contains 267 companies and the FTSE4Good Europe (LFT4GBEU), which is composed by 408 companies.

2.2.2 Sin portfolios

In vice investing, investors exclude SRI stocks and invest primarily in unethical industries such as on tobacco, alcohol, and gambling (gaming) stocks. These stocks are known as “sin” stocks.

One of the most important studies about “sin” stocks was published by Hong and Kacperczyk (2009) where, the authors analysed “sin” stocks as alcohol, tobacco and gambling. Since then, alcohol, tobacco and gambling industries which are called “sin triumvirate” are the most often negatively screened stocks by socially responsible investors. The classification of a “sin” stock may vary, depending on the investor’s beliefs, cultural values and social, economic and climatic norms, which may lead to the expansion of the industries included in a “Sin” portfolio. Some previous studies, for example by Ghoul et al. (2011), have been adding other industries to the “sin” stocks such as firearms, military and nuclear power and defense. Richey (2017) also add other controversial businesses: adult entertainment, and payday lenders. Although these last industries are not as homogeneous as the “sin triumvirate”, more researchers have been including these industries to their vice portfolios.

This study focuses on the “sin” stocks of the following industries: alcohol, tobacco, gambling and defense. Although FTSE4Good Index also excludes companies related with nuclear power activities, it was not found any company

when identifying the “sin” stocks. In order to identify the “sin” stocks, it was analysed the constituents of the STOXX 600 (Europe) and S&P 500 (US) over the period of the study.

Hong and Kacperczyk (2009) and Richey (2016), in order to identify the “sin” stocks, they use Fama and French’s (1997) classification of stocks based on their SIC codes. For example, alcohol stocks fall under the Fama and French industry group 4 and tobacco stocks are in industry group 5. In this study, to identify the “sin” stocks, it was selected the variable industry and analysed the SIC codes in the Datastream. It was found 32 “sin” stocks from Stoxx 600 and 17 stocks from S&P 500.³

Just as in the FTSE4Good portfolios, it was obtained for each “sin” stock, the prices of the stocks in dollars of the last day of each month between July 31, 2001 and September 31, 2019. Another important information for the construction of the “Sin” portfolio, was the market value in dollars. In order to simplify the construction of the portfolios, it was assumed that the market value does not vary significantly over the month, hence it was only collected the market value for the last day of each month, of the period in study. The main purpose of the market value was to calculate the weight for each stock, in order to construct value-weighted portfolios.

The returns of the FTSE4Good portfolios and the portfolios composed by “sin” stocks were calculated using the following formula:

$$r_{i,t} = \frac{p_{i,t}}{p_{i,t-1}} - 1$$

where $p_{i,t}$ denotes the closing price of the index/stock i at month t .

In order to achieve uniformity in the presentation of the results, the values of the portfolios were collected in dollars.

³Stoxx 600 is a European index composed by 600 stocks from 17 countries of the European region: Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, Luxembourg, the Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland and the United Kingdom. S&P 500 is an American index composed by 500 stocks.

2.2.3 Summary statistics with excess returns

Excess returns were calculated by subtracting the monthly risk-free rate from the monthly returns in all the stocks that compose the portfolios (both “sin” portfolios and FTSE4Good portfolios), using the data available in the Kenneth French website.

Table 1: Summary statistics of the variables of the equations: Europe

Variables	Mean	Median	Std. Dev.	Min	Max
Portfolios excess returns					
<i>LFT4GBEU</i>	0.1653	0.1140	5.2882	-20.7200	13.3780
<i>sinEU</i>	0.8391	1.2303	4.5188	-19.6278	12.2391
<i>benchEU</i>	0.9414	0.9825	5.8034	-25.7889	21.5832
Risk factors returns					
<i>mktrf</i>	0.5523	0.6850	5.2214	-22.0600	13.7400
<i>Smb</i>	0.1650	0.1750	1.8889	-6.8200	4.8600
<i>Hml</i>	0.1878	0.2350	2.1524	-4.7200	8.2900
<i>wml</i>	0.8706	1.1150	4.1039	-26.2700	13.6900
<i>rmw</i>	0.3214	0.3850	1.6341	-4.8800	6.3800
<i>cma</i>	0.1834	0.0100	1.5186	-5.4700	5.3700

The statistics presented are across 218 observations over the period of 2001-2019. The factors size (SMB), value (HML), momentum (WML), profitability (RMW) and investment (CMA) for Europe are from the Kenneth French website. This table shows the summary statistics of the FTSE4Good Europe Portfolio, the “Sin” Europe Portfolio and the Benchmark portfolio. The “Sin” portfolio is composed by 32 companies of the STOXX 600 selected per industry (Alcohol, Gambling, Tobacco and Defense). There are 7 stocks of alcohol, 12 stocks of gambling, 5 stocks of gambling and 8 stocks of Defense. According to the Datastream, the FTSE4GOOD Europe (LFT4GBEU) is composed by 408 companies. The benchmark used was the STOXX 600

In table 1, it is shown that the “Sin” Europe portfolio has higher returns than the FTSE4Good Europe portfolio (0.8391>0.1653) on average. The proxy of the market portfolio (STOXX 600) has on average superior returns comparing to the “Sin” portfolio (0.9414>0.8391). The “Sin” portfolio has a lower risk comparing to the FTSE4Good portfolio (4.5188<5.2882) and the benchmark (4.5188<5.8034).

Table 2: Summary statistics of the variables of the equations: US

Variables	Mean	Median	Std. Dev.	Min	Max
Portfolios excess returns					
<i>LFT4GBUS</i>	0.3884	0.9084	4.3563	-16.6431	10.8322
<i>sinUS</i>	0.9178	1.2330	3.7955	-10.3304	14.8434
<i>benchUS</i>	1.0146	1.3724	4.6777	-20.1211	18.6233
Risk factors returns					
<i>mktrf</i>	0.5973	1.0600	4.2761	-18.4100	11.5600
<i>Smb</i>	0.1044	-0.0150	2.3117	-6.1100	6.9600
<i>Hml</i>	0.0169	0.0600	2.3825	-7.2500	7.5000
<i>wml</i>	0.2769	0.4650	4.1315	-24.9700	11.3500
<i>rmw</i>	0.3269	0.2900	1.8208	-7.3300	7.1200
<i>cma</i>	0.1393	-0.0800	1.9009	-6.4900	8.1900

The statistics presented are across 218 observations over the period of 2001-2019. The factors size (SMB), value (HML), momentum (WML), profitability (RMW) and investment (CMA) for the US are from the Kenneth French website. This table shows the summary statistics of the FTSE4Good US Portfolio, the “Sin” US Portfolio and the Benchmark portfolio. The “Sin” portfolio is composed by 17 companies of the S&P 500 selected per industry (Alcohol, Gambling, Tobacco and Defense). There are 2 stocks of alcohol, 4 stocks of gambling, 5 stocks of gambling and 6 stocks of Defense. According to the Datastream, the FTSE4Good US (LFT4GBUS) is composed by 267 companies. The benchmark used was the S&P 500.

In table 2, it is shown that the “Sin” US portfolio presents superior returns than the FTSE4Good US portfolio ($0.9178 > 0.3884$), on average. The proxy of the market portfolio (S&P 500) has on average superior returns comparing to the “sin” portfolio ($1.0146 > 0.9178$). The “sin” portfolio has a lower risk comparing to the FTSE4Good portfolio ($3.7955 < 4.3563$) and the benchmark ($3.7955 < 4.6777$).

Analysing both summary statistics tables, it can be concluded that all the variables present a positive excess return. Tables 1 and 2 show that the FTSE4Good Europe (LFT4GBEU) and the FTSE4Good US (LFT4GBUS) portfolios have lower excess returns than the sinEU and sinUS portfolios.

3. Results

In the following tables, in order to test the alphas to see if they are statistically significant, it is used the Newey–West standard errors⁴ which are robust to heteroskedasticity and serial correlation.

3.1 Results of the FTSE4Good portfolios

Table 3: Performance measurement of FTSE4Good portfolio in the US with different performance models: July 2001–September 2019, 218 months

	CAPM	FF 3-factor	Carhart	FF 5-factor
Alpha	- 0.0235 (0.1002)	0.2646*** (0.019)	0.1928** (0.0804)	0.2655*** (0.0903)
Beta	1.0104*** (0.1110)	0.7682*** (0.00009)	0.8208*** (0.0853)	0.7488*** (0.0963)
SMB		-0.2884*** (0.085)	-0.2673*** (0.0314)	-0.3214*** (0.0336)
HML		-0.0634* (0.006)	-0.0686* (0.0366)	-0.1098*** (0.0420)
WML			-0.0527* (0.0278)	
RMW				-0.1851*** (0.0468)
CMA				0.0942* (0.0530)
Adjusted R-squared	0.9350	0.9502	0.9519	0.9559

* All specifications include a constant term and are based on 218 observations. Newey–West standard-errors in parenthesis. *** denote p-values <0.01, ** denote p-values <0.05, and * denote p-values <0.10.

⁴ According with Newey and West (1987), in order to set lag length m to the integer part of $T^{1/4}$, in which T represents the number of observations. In this case, the number of observations is 218, so the lag settled was 3, because m is lower than $T^{1/4}$.

Table 4: Performance measurement of FTSE4Good portfolio in Europe with different performance models: July 2001–September 2019, 218 months

	CAPM	FF 3-factor	Carhart	FF 5-factor
Alpha	- 0.0957* (0.0509)	0.2051*** (0.0391)	0.1807*** (0.0384)	0.1915*** (0.0359)
Beta	1.1071*** (0.0593)	0.7690*** (0.0417)	0.7904*** (0.0406)	0.7725*** (0.0387)
SMB		-0.3129*** (0.0299)	-0.3038*** (0.0277)	-0.3075*** (0.0244)
HML		0.0114 (0.0248)	0.0097 (0.0242)	-0.0203 (0.0349)
WML			-0.0150 (0.0139)	
RMW				-0.1060*** (0.0332)
CMA				-0.0218 (0.0423)
Adjusted R-squared	0.9799	0.9888	0.9888	0.9893

* All specifications include a constant term and are based on 218 observations. Newey–West standard-errors in parenthesis. *** denote p-values <0.01, ** denote p-values <0.05, and * denote p-values <0.10

3.1.1 Discussion of the FTSE4Good portfolios

- **General Discussion**

Tables 3 and 4 show the results of the FTSE4Good portfolios, in the US and Europe, respectively. Although the CAPM presents a negative and statistically significant alpha in the European market, the alphas in the remaining regressions indicate a confident prediction of an outperformance, after controlling for the different factors. Lobe and Walkshäusl (2014) shows in their results that in the FTSE4Good US portfolio, the alpha for CAPM is negative, contrary to the FTSE4Good Europe, in which the alpha is positive. The authors find positive alphas in the Fama and French (1993) three-factor model and Carhart (1997) four-factor model. This supports the results obtained in this study, including in the Fama and French (2015) five-factor model.

Beta is higher than 1 in both markets, in CAPM, but it becomes lower than 1, in the Fama and French (1993) three-factor model, Carhart (1997) four-factor model and Fama and French (2015) five-factor model.

In this work, the SMB factor exhibits a negative value, pointing to a higher proportion of big-cap companies, but comparing with Lobe and Walkshäusl (2014), they did not reach a conclusion due to their heterogenous results.

The adjusted r-squared vales are relatively high (>90%), which indicate that these portfolios are a close approximation of the benchmark.

- **CAPM**

The alpha in CAPM is negative in both tables but it is only significant in the FTSE4Good Europe portfolio. Analysing the betas, in the SRI portfolios, both are higher than 1, which means that they are more sensitive to market risk. This is supported by the evidence found by Schröder (2007).

We reject the null hypothesis ($\alpha=0$ and $\beta=1$), mentioned in section 2.2.1. Although the portfolios reproduce very closely the movements of their

benchmarks, as can be observed by the high adjusted r-squared value, the rejection of the null hypothesis indicates the need of applying more complex models.

- **Fama and French three-factor model**

Although in CAPM, the adjusted r-squared value is high in both portfolios, by adding the new factors of the Fama and French three-factor model, there are some changes in both portfolios. The alphas turn positive and they are statistically significant in both portfolios with a significance level of 1%.

Both betas are now inferior to 1 and with a significance level of 1%. The smaller value of the betas shows that the risk-exposure of these portfolios is partly explained by the size and value factors.

The estimates for the factors SMB and HML are overall statistically significant, except for the HML, in the FTSE4Good Europe portfolio. The negative and statistically significant HML estimate in the FTSE4Good US portfolio shows that this portfolio tends to contain more growth stocks than value stocks in relation to the benchmark.

In both portfolios, the negative SMB indicate that the FTSE4Good portfolios is mostly composed by large capitalization companies, comparing to the benchmark.

- **Carhart four-factor model**

In both portfolios, alpha is positive, and statistically significant, whereas beta is inferior to 1 and with a significance level of 1%.

In the FTSE4Good US portfolio, it is shown that the inclusion of the momentum factor is statistically significant in contrast with the FTSE4Good Europe portfolio, and it does not influence the other factors (HML and SMB). Although authors such as Lesser et al. (2014) do not find significant estimates for this factor, the negative WML found in these results are supported by Lobe and

Walkshäusl's (2014) findings. The WML factor in the FTSE4Good US portfolio is of small magnitude, negative and statistically significant, which indicates that there are less winner stocks and more loser stocks.

The estimate for SMB is negative and statistically significant in both portfolios. In contrast, the HML factor is not significant once again in the FTSE4Good Europe portfolio.

- **Fama and French five-factor model**

In both portfolios, it is found alphas with a positive value and a significance level of 1% in both portfolios. The same level of significance is found in the betas that are both inferior to 1. The estimate of the betas shows that the risk-exposure of these portfolios is explained by the factors.

The estimate of the SMB factor in both tables, does not have a significant change from the Carhart (1997) four-factor model to the Fama and French (2015) five-factor model.

The estimate of the HML remains negative and now with a significance level of 1% in the FTSE4Good US portfolio, which is an indicative of a higher weight of growth stocks.

The inclusion of the variables RMW and CMA in the FTSE4Good Europe portfolio, affects the HML factor, that becomes negative.

The estimate of RMW for both portfolios is negative and significant, which is an indicator that both portfolios have a weak operating profitability. The estimate for the factor CMA is positive and significant in the FTSE4Good US portfolio, which indicates that it is used a more conservative investment strategy.

3.2 Results of the Sin portfolios

Table 5: Performance measurement of sin portfolio in the US with different performance models: July 2001–September 2019, 218 months

	CAPM	FF 3-factor	Carhart	FF 5-factor
Alpha	0.2410 (0.2060)	0.5453** (0.2284)	0.7732*** (0.2112)	0.3979 (0.2206)
Beta	0.3261 (0.2103)	0.0975 (0.2233)	-0.0694 (0.2131)	0.3503 (0.2232)
SMB		-0.4350*** (0.1080)	-0.5021*** (0.1081)	-0.2935*** (0.1103)
HML		0.1507 (0.1069)	0.1674* (0.0930)	0.0219 (0.1352)
WLM			0.1672** (0.0766)	
RMW				0.5276*** (0.1805)
CMA				0.2879* (0.1716)
Adjusted R-squared	0.4300	0.4848	0.5074	0.5351

* All specifications include a constant term and are based on 218 observations. Newey–West standard-errors in parenthesis. *** denote p-values <0.01, ** denote p-values <0.05, and * denote p-values <0.10

Table 6: Performance measurement of sin portfolio in the Europe with different performance models: July 2001–September 2019, 218 months

	CAPM	FF 3-factor	Carhart	FF 5-factor
Alpha	0.2081 (0.1958)	0.2001 (0.2156)	0.6410*** (0.2139)	0.3459* (0.1859)
Beta	0.4700** (0.2155)	0.5329** (0.2227)	0.1464 (0.2266)	0.4841** (0.2023)
SMB		0.0621 (0.1729)	-0.1016 (0.1441)	0.0297 (0.1385)
HML		-0.3281** (0.1399)	-0.2976** (0.0930)	-0.1599 (0.1608)
WML			0.2711*** (0.0636)	
RMW				0.8096*** (0.1886)
CMA				0.3391** (0.1363)
Adjusted R-squared	0.6487	0.6670	0.7044	0.7251

* All specifications include a constant term and are based on 218 observations. Newey–West standard-errors in parenthesis. *** denote p-values <0.01, ** denote p-values <0.05, and * denote p-values <0.10

3.2.1 Discussion of the sin portfolios

- **General Discussion**

Both table 5 and 6 that present the results of the “sin” portfolios, in the US and Europe, respectively, show positive alphas in all the models. In the European market, it is found an outperformance of the “Sin” portfolio, in the more complex models, the Carhart (1997) four-factor model and Fama and French (2015) five-factor model, that present a significance level of 1% and 10%, respectively.

The Fama and French (1993) three-factor model and the Carhart (1997) four-factor model exhibit significant alphas in the “Sin” US portfolio. Although these two regressions, point to the existence of a superior performance, the alpha turns statistically insignificant, when applying the Fama and French (2015) five-factor model.

Hong and Kacperczyk (2009) states that the values-driven investors neglect these stocks and the authors consider the outperformance of “sin” stocks is a mechanism to compensate these investors for the sharing less risk.

Beta is lower than 1 in all the models, but it is statistically significant only in the European market, in the CAPM and Fama and French (1993) three-factor model, presenting the following values 0.4700 and 0.5329, respectively. This is an indication that “sin” portfolios exhibit less market risk.

In conclusion, it is found an outperformance of the “sin” portfolios for the Carhart (1997) four-factor model and Fama and French (2015) five-factor model in Europe. These results can be compared with the findings of Hong and Kacperczyk (2005), in which the authors report an outperformance over the period of 1965–2004, when employing the four-factor model. In this study, the “Sin” US portfolio fails to outperform the market index (S&P 500) in the Fama and French (2015) five-factor model. This is supported by the evidence shown by Richey (2017).

According to extant empirical literature, such as Liston and Soydemir (2010) and Hong and Kacperczyk (2005), “sin” industries usually present a positive HML. In contrast, this study presents mixed results in the HML factor. The “Sin” Europe portfolio presents a negative estimate and the American portfolio exhibits values with a positive sign.

As in Richey (2017), we also found a positive and significant RMW and CMA, which indicate that the firms of these portfolios tend to have a robust operating profitability and a more conservative investment strategy.

The adjusted r-squared values are lower than 90%, which indicate that these two portfolios cannot be well approximated by the benchmark.

- **CAPM**

Analysing the tables 5 and 6, the alpha in CAPM is positive in both portfolios (“Sin” Europe portfolio and “Sin” US portfolio), but they are non-statistically significant. These results support the evidence presented by Lobe and Walkshäusl (2014) and Liston and Soydemir (2010) which also show positive alphas.

Analysing both betas in the “Sin” portfolios, they are lower than 1, but only the beta of the “Sin” Europe portfolio is statistically significant, which indicates that these stocks have a lower risk comparing to the market.

In conclusion, the equation of the “Sin” US portfolio is not relevant due to non-statistically significance and the low value of the adjusted r-squared, therefore this model cannot explain everything. The equation of the “Sin” Europe portfolio can reproduce closer the movements of its benchmark because of its higher adjusted r-squared value but it should be applied more complex specifications.

- **Fama and French three-factor model**

The alphas are positive in both portfolios, but the only statistically significant is from the “Sin” US portfolio. Both betas are smaller than 1, but only the one from the “Sin” Europe portfolio is statistically significant (5% significance level). The low value of beta shows that the risk-exposure of these portfolios is partly explained by the factors added in this regression.

In the “Sin” US portfolio, the estimate for the SMB factor is negative and shows a significance level of 1%, indicating a large cap tilt. In contrast, the same factor is positive and non-statistically significant in the “Sin” Europe portfolio. Lobe and Walkshäusl (2014), also indicate in their results, a negative SMB in the US portfolio.

The estimate for the HML factor is positive and non-statistically significant in the “Sin” US portfolio. The negative and significant HML factor in the “Sin” Europe portfolio, shows that this portfolio is composed by a higher proportion of growth stocks, in relation to the benchmark.

The adjusted r-squared value is higher in the European portfolio than in the American portfolio, but in both portfolios is lower than 90%. This indicates that this model is not able to fully explain the performance of both portfolios.

- **Carhart four-factor model**

In both portfolios, the alphas are positive, and with a significance level of 1%. Analysing the betas, both are non-statistically significant.

The inclusion of the momentum (WML) variable is positive and statistically significant in the American and European portfolios, which shows that there are more winner stocks and less loser stocks comparing to the benchmark. Liston and Soydemir (2010), also show a positive WML.

In the “Sin” US portfolio, the estimates of the SMB and HML factors are statistically significant, whereas in the “Sin” Europe portfolio, only the estimate for the HML is significant. The SMB factor in the “Sin” US portfolio is negative

which means that, the portfolio is mostly composed by big-cap companies. The HML coefficient is not coherent in both portfolios, in the European portfolio is negative and the American portfolio is positive. Hence, the “Sin” Europe portfolio is mostly composed by growth stocks comparing to the benchmark and the “Sin” US portfolio has a higher proportion of value stocks.

- **Fama and French five-factor model**

In both portfolios, the alphas of the regressions are positive, but it is only significant in the “Sin” Europe portfolio. The “Sin US portfolio does not exhibit an out or underperformance compared to the benchmark.

In the “Sin” US portfolio, the SMB estimate is coherent with the previous models, being negative and with a significant level of 1%. The HML factor is insignificant in both portfolios.

The inclusion of the variables RMW and CMA in the “Sin” US portfolio, affect the beta, that turns positive, whereas, in the “Sin” Europe portfolio, it affects SMB, which becomes positive but still insignificant.

In both portfolios, RMW factor is positive and it has a significance level of 1%, which indicates that they have a robust operating profitability. The positive and significant CMA factor indicate that both portfolios, invest more conservatively.

4. Analysis in recession and expansion periods

Some previous researches have been also focusing on analysing the returns of stocks in recession and expansion periods, due to the importance of business cycles for stock return predictability. Authors such as Chordia and Shivakumar (2002) and Fama and French (1989) by using conditional models show the existence of a negative relationship between expected returns and business cycles, which means that the expected returns are higher at troughs than comparing to the peaks of the business cycles. According to Fama and French (1989), the expected returns are higher in periods of recession and lower in periods of expansion since investors tend to have bigger savings in expansion periods than comparing to recession periods. Hong and Kacperczyk (2009) also show that “sin” firms use more private debt to finance their operations comparing to other firms, implying that “sin” stocks may be less sensitive to periods of recession than other stocks.

In this study, the period of the sample is from July 2001 to September 2019, and according to the NBER classification (based on the business cycles available in the National Bureau of Economic Research website⁵), the periods of recession are 04.2001-11.2001 (in this study, only the information since July 2001 is accounted for, since the period of analysis starts in this month) and 12.2007-07.2009. In contrast, the periods of expansion are 12.2001-11.2007 and 08.2009-09.2019.

⁵ The information about the different recession and expansion periods is available in the following website: <https://www.nber.org/cycles.html>

So far, this study has only been focused on the analysis of FTSE4Good and “sin” portfolios, individually. In this section, the hedge portfolio approach will be used. It will be presented the alphas, the betas and the other coefficients of the factors, of the asset pricing models previously mentioned, in order to explain the hedge returns. This will allow to compare the results with Lobe and Walkshäusl’s (2014) and Hong and Kacperczyk’s (2009) works, who use the same approach.

In this current section, it will be presented the regressions for the CAPM, the Fama and French (1993) three-factor model, the Carhart (1997) four-factor model and Fama and French (2015) five-factor model, in order to analyse the impact of recession and expansion periods in the performance of the “sin” portfolios, the FTSE4Good portfolios and the difference portfolios.

4.1 Analysis in the American market

Table 7: Summary statistics in expansion periods in the US

Variables	Mean	Median	Std. Dev.	Min	Max
Dcrise	0	0	0	0	0
LFT4GBUS	0.6470	0.9583	3.6783	-10.7285	10.1570
sinUS	1.0769	1.2781	3.5489	-9.9227	14.8434
benchUS	1.2264	1.4583	3.8793	-10.3293	13.2210
sinUS-LFT4GBUS	0.3233	0.2074	3.2822	-11.1158	11.4702
mktrf	0.8535	1.1400	3.6818	-10.1200	11.5600
Smb	0.0666	-0.0600	2.2260	-5.1300	5.7600
Hml	0.0629	0.0700	2.2254	-6.3600	7.5000
wml	0.46613	0.4650	3.2652	-15.5600	9.6600
rmw	0.2622	0.2650	1.7642	-7.3300	7.1200
cma	0.1220	-0.0500	1.6223	-3.5600	7.5000

* The statistics presented are computed across 218 observations.

Table 8: Summary Statistics in recession periods in the US

Variables	Mean	Median	Std. Dev.	Min	Max
Dcrise	1	0	1	1	1
LFT4GBUS	-1.7025	0.0523	7.7712	-16.6431	10.8322
sinUS	-0.3681	-0.8868	5.3191	-10.3304	9.6090
benchUS	-0.6975	-0.1418	8.7593	-20.1211	18.6323
sinUS-LFT4GBUS	1.2140	1.5986	5.0298	-8.5242	12.9426
mktrf	-1.4738	-0.9250	7.3306	-18.4100	10.4800
Smb	0.4100	0.4350	2.9547	-6.1100	6.9600
Hml	-0.3550	-0.2450	3.4401	-7.2500	4.9300
wml	-1.2525	-0.9050	8.2965	-24.9700	11.3500
rmw	0.8504	1.3250	2.2007	-3.0100	5.3400
cma	0.2792	-0.9100	3.4617	-6.4900	8.1900

* The statistics presented are computed across 218 observations.

In the American market, comparing both tables, when examining the data is visible that in periods of expansion ($dcrise=0$), the excess returns on average of FTSE4Good and “Sin” portfolios are higher comparing to returns in periods of recession ($dcrise=1$). In the difference portfolio (difference between the “Sin” portfolio and FTSE4Good portfolio), the excess returns are higher, on average, during periods of recession than in periods of expansion. This fact may indicate that in periods of recession, the excess returns of the FTSE4Good portfolios are much lower comparing to the returns of the “sin” portfolios.

These summary statistic tables allow to perform a preliminary analysis of the dataset. In the next tables, the coefficients of each variable of the following

models are presented: CAPM, Fama and French (1993) three-factor model, Carhart (1997) four-factor model and Fama and French (2015) five-factor model.

Table 9: Descriptive statistics in recession periods in the US: FTSE4Good portfolio and Sin portfolio

	FTSE4GOOD Portfolio (1)				Sin Portfolio (2)			
	CAPM	FF 3-factor	Carhart	FF 5-factor	CAPM	FF 3-factor	Carhart	FF 5-factor
Alpha	0.2346 (0.2119)	0.3273 (0.2148)	0.2243 (0.2254)	0.3465 (0.1983)	0.3585 (0.2455)	0.3138 (0.4300)	0.3011 (0.7067)	0.1428 (0.3509)
Beta	0.7534*** (0.2652)	0.6805** (0.2245)	0.7609** (0.2162)	0.6864** (0.1809)	0.1436 (0.2825)	0.2052 (0.4507)	0.2151 (0.6641)	0.6525 (0.3578)
SMB		-0.2031 (0.1698)	-0.1732 (0.1720)	-0.2678 (0.1819)		-0.1621 (0.4990)	-0.1584 (0.5749)	-0.0673 (0.4088)
HML		0.1113 (0.1430)	0.1172 (0.1553)	-0.1589 (0.2462)		0.3089* (0.1681)	0.3096* (0.1740)	-0.2444 (0.3013)
WML			-0.0455 (0.0796)				-0.0056 (0.1477)	
RMW				-0.5870*** (0.0929)				0.1815 (0.2465)
CMA				0.3162 (0.1900)				0.8440** (0.3461)
Adjusted r-squared	0.9352	0.9380	0.9359	0.9627	0.5786	0.5991	0.5769	0.6631

* All specifications include a constant term and are based on 218 observations. Newey-West Standard-errors in parenthesis. *** denote p-values <0.01, ** denote p-values <0.05, and * denote p-values <0.10

Table 10: Descriptive statistics in recession periods in the US: Difference portfolio

Difference Portfolio (2-1)				
	CAPM	FF 3-factor	Carhart	FF 5-factor
Alpha	0.1172 (0.2321)	-0.0123 (0.5157)	0.0584 (0.8310)	-0.2004 (0.4692)
Beta	-0.5990* (0.3531)	-0.4721 (0.5588)	-0.5273 (0.8008)	-0.0354 (0.4636)
SMB		0.0368 (0.4961)	0.0163 (0.5667)	0.1942 (0.4030)
HML		0.1887 (0.2568)	0.1847 (0.2706)	-0.0929 (0.4626)
WML			0.0313 (0.1992)	
RMW				0.7569* (0.3086)
CMA				0.5241 (0.4346)
Adjusted r-squared	0.4053	0.3576	0.3233	0.4556

* All specifications include a constant term and are based on 218 observations. Newey-West Standard-errors in parenthesis. *** denote p-values <0.01, ** denote p-values <0.05, and * denote p-values <0.10

- **General Discussion**

In table 9, alpha is positive and statistically insignificant in both portfolios, thus it is not found an underperformance or outperformance, during a recession period in the US.

There are four betas which are below 1 in the FTSE4Good portfolio, with a significance level between 1% and 5%. The results indicate that in a period of recession, the FTSE4Good portfolio tends to be less sensitive to market risk.

The adjusted r-squared values in the FTSE4Good portfolio tend to increase as a more complex model is applied, therefore all models present a value higher than 90%, which indicate that this portfolio can be well approximated by the benchmark.

In table 10, the hedge portfolio presents two positive and two negative alphas but none of them is significant.

The betas are all negative in the different models, but only beta from the CAPM specification is significant. The negative betas indicate that the “sin” portfolio has a beta smaller than 1. These results are also supported in the evidence shown by Salaber (2007), in which the author found betas lower than 0, when analysing different hedge portfolios (long in the “sin” stock portfolio and short in two zero-investment portfolios).

Overall, it is not found evidence of an outperformance or underperformance during the periods of recession for our sample period.

- **CAPM**

In table 9, the alpha in CAPM is positive in both portfolios (FTSE4Good portfolio and “Sin” portfolio), but they are non-statistically significant.

- **Fama and French three-factor model**

In both portfolios, alpha is positive but not significant, whereas beta is lower than 1 and significant (significance level of 5%), in the FTSE4Good portfolio, which shows that this portfolio presents a lower volatility. The SMB factor is negative but insignificant in both portfolios. In the “Sin” portfolio, the HML coefficient tilts more towards value stocks than growth stocks.

In the difference portfolio, although none of the coefficients is significant, alpha is negative. Lobe and Walkshäusl (2014) also found negative alphas in hedge portfolios using this model augmented with factor (WML) of the Carhart (1997) four-factor model.

- **Carhart four-factor model**

In the FTSE4Good and “Sin” portfolios, the alphas are positive, but they are not statistically significant and only beta of the FTSE4Good portfolio is statistically significant and inferior to 1. This indicates that the stocks of this portfolio present a lower risk comparing to the market. The estimates of the SMB and WML are non-significant in both portfolios. The “Sin” portfolio has a higher proportion of value stocks, resulting in a significant and positive HML factor.

In table 10, although none of the coefficients are significant, beta is lower than 0 and the HML estimate is positive, which is supported by the Lobe and Walkshäusl’s (2014) results.

- **Fama and French five-factor model**

In this model, both portfolios have non-statistically significant alphas. The beta of the FTSE4Good portfolio is inferior to 1 and significant, which once again indicates that this portfolio is less sensitive to the market risk. In the “Sin” portfolio, beta is not significant.

Analysing the FTSE4Good portfolio, it is observed that neither of the coefficients for SMB, HML and CMA factors are significant. In contrast, the RMW factor is negative and with a significance level of 1%, which shows that a lower operating profitability has a higher weight in the portfolio.

In the “Sin” portfolio, the coefficients of the factors that are statistically insignificant are SMB, HML and RMW. The estimate of the CMA is positive and significant, which indicates a more conservative strategy of the firms that compose the portfolio.

Table 11: Descriptive statistics in expansion periods in the US: FTSE4Good Portfolio and Sin Portfolio

	FTSE4GOOD Portfolio (1)				Sin Portfolio (2)			
	CAPM	FF 3-factor	Carhart	FF 5-factor	CAPM	FF 3-factor	Carhart	FF 5-factor
Alpha	-0.2297*** (0.0682)	0.0945 (0.0739)	0.0695 (0.0789)	0.0386 (0.0864)	0.2349 (0.3286)	0.6514 (0.3792)	0.7784** (0.3159)	0.5582 (0.3471)
Beta	1.2008*** (0.0765)	0.9260*** (0.0732)	0.9360*** (0.0763)	0.9648*** (0.0851)	0.3497 (0.3314)	0.0409 (0.3596)	-0.0097 (0.3159)	0.2210 (0.3398)
SMB		-0.2587*** (0.0319)	-0.2426*** (0.0306)	-0.2705*** (0.0325)		-0.4749*** (0.1105)	-0.5565*** (0.1043)	-0.3295*** (0.1230)
HML		-0.1004*** (0.0297)	-0.1078*** (0.0288)	-0.1517*** (0.0299)		0.1457 (0.1261)	0.1831* (0.1097)	0.0961 (0.1414)
WML			-0.0469** (0.0198)				0.2379*** (0.0992)	
RMW				-0.1111*** (0.0324)				0.5692*** (0.1912)
CMA				0.1247** (0.0475)				0.0889 (0.1924)
Adjusted r-squared	0.9354	0.9552	0.9565	0.9591	0.3725	0.4351	0.4748	0.4893

* All specifications include a constant term and are based on 218 observations. Newey-West Standard-errors in parenthesis. *** denote p-values <0.01, ** denote p-values <0.05, and * denote p-values <0.1

Table 12: Descriptive statistics in expansion periods in the US: Difference Portfolio

Difference Portfolio (2-1)				
	CAPM	FF 3-factor	Carhart	FF 5-factor
Alpha	0.4683 (0.3406)	0.5598 (0.4039)	0.7113** (0.3420)	0.5196 (0.3839)
Beta	-0.8524** (0.3412)	-0.8861** (0.3800)	-0.9464*** (0.3362)	-0.7418* (0.3735)
SMB		-0.2154* (0.1271)	-0.3127*** (0.1183)	-0.0577 (0.1357)
HML		0.2458* (0.1391)	0.2906** (0.1192)	0.2452* (0.1523)
WML			0.2837*** (0.1085)	
RMW				0.6797*** (0.1998)
CMA				-0.0302 (0.2131)
Adjusted r-squared	0.1812	0.2142	0.2810	0.3095

* All specifications include a constant term and are based on 218 observations. Newey-West Standard-errors in parenthesis. *** denote p-values <0.01, ** denote p-values <0.05, and * denote p-values <0.1

- **General Discussion**

In table 11, the FTSE4Good portfolio presents a negative and significant alpha in the CAPM specification, whereas in the other asset pricing models, it was obtained positive but statistically insignificant alphas. In the “Sin” portfolio, all the alphas are positive, but only the one from the Carhart (1997) four-factor model points to the existence of higher abnormal returns. There is an impact in the estimate, when adding the factors of the Fama and French (2015) five-factor model and alpha becomes insignificant.

In the FTSE4Good portfolio, all the models present a beta with a significance level of 1%. In CAPM, beta is higher than 1 but in the other models is lower than 1. In the “Sin” portfolio, the results show that all the betas are insignificant.

Comparing table 9 with table 11, it is observed that during recession periods, FTSE4Good portfolio is less sensitive to the market risk and that during periods of expansion, beta is closer to 1, indicating that the portfolio tends to move more like the market.

It is important to note that according to the adjusted r-squared values in the FTSE4Good portfolio, which presents values higher than 90%, this portfolio can be well approximated by the benchmark because the adjusted r-squared increases with the usage of a more complex model.

In table 12, the results obtained in the Carhart (1997) four-factor model point to an outperformance of the difference portfolio but by adding the two factors of the Fama and French (2015) five-factor model, there is a decrease in the estimate and it turns statistically insignificant.

The negative and significant betas in the different models reflect the fact that “Sin” portfolio has a beta smaller than 1.

The hedge portfolio shows a positive WML factor, which according to Lobe and Walkshäusl (2014) is because the WML estimate of the FTSE4Good is negative and this factor does not have an impact in the “Sin” portfolio.

In this portfolio, it is also found a significant and positive HML factor in three of four models. Furthermore, it is found a negative and significant SMB estimate in two of the models, Fama and French (1993) three-factor model and Carhart (1997) four-factor model. In his regressions Salaber (2007) also found an HML factor higher than 0 and a SMB factor lower than 0, when applying the Fama and French (1993) three-factor model and the momentum factor. This evidence also supports our findings.

- **CAPM**

In the FTSE4Good portfolio, the alpha is negative and with a significance level of 1%, which points to an underperformance of this portfolio, whereas in the “Sin” portfolio, the alpha is positive, but it is not significant.

In the FTSE4Good portfolio, the beta is higher than 1, so it is considered riskier because the stocks prices may tend to vary wildly. Comparing with the beta of the “Sin” portfolio, this one is lower than 1, but it is not significant.

- **Fama and French three-factor model**

Analysing alpha in both portfolios, they are positive but not significant. In contrast, both betas are lower than 1, but only the beta from the FTSE4Good portfolio has a significance level of 1%, which indicates that the portfolio is less sensitive to the market risk.

The SMB factor for both portfolios is negative and significant, hence it indicates a big-cap tilt. The HML factor, in the FTSE4Good portfolio, presents a 1% significance level and shows that this portfolio is mostly composed by growth stocks.

- **Carhart four-factor model**

In the FTSE4Good and “Sin” portfolios, alphas are positive but only the “Sin” portfolio’s alpha is statistically significant. The beta of the FTSE4Good portfolio is lower than 1 and it is significant at a 1% level, which means that the stocks of the portfolios are less volatile comparing to the market. The SMB factor is negative, and it indicates that the portfolios contain a higher ratio of large-cap companies. The estimates of the HML is also statistically significant (at 1% level) in both portfolios, but in the FTSE4Good portfolio, the value is negative, which shows that the portfolio is composed mostly by growth stocks, comparing to the benchmark. In the “Sin” portfolio, the value is positive, indicating that this portfolio is tilted towards value stocks.

The estimate of the WML factor is also statistically significant in both portfolios but in the FTSE4Good portfolio, the factor is negative and in the “Sin” portfolio is positive. Hence, the FTSE4Good portfolio is mostly composed by loser stocks comparing to the benchmark and the “Sin” portfolio has a higher proportion of winner stocks.

- **Fama and French five-factor model**

It is not found an outperformance or an underperformance in both portfolios. The beta of the FTSE4Good portfolio is statistically significant at the 1% level and lower than 1, which means that the risk of this portfolio is lower comparing to the benchmark.

In both portfolios, the estimate for the SMB factor indicates that both portfolios are mostly composed by big-cap companies. Furthermore, the estimate for the HML factor is only statistically significant (at 1% level) in the FTSE4Good portfolio and its value is negative. This indicates that this portfolio is mostly composed by growth stocks. In the FTSE4Good portfolio, the estimate for the RMW and CMA are statistically significant, although the RMW factor is negative and the CMA is positive, which indicate a weak operating profitability and a more conservative strategy, respectively. In the “Sin” portfolio, the estimate for the RMW factor is positive and significant at the 1% level, which shows that this portfolio has a higher number of firms with a robust operating profitability.

4.2 Analysis in the European market

Table 13: Summary statistics in expansion periods in the Europe

Variables	Mean	Median	Std. Dev.	Min	Max
Dcrise	0	0	0	0	0
LFT4GBEU	0.4621	0.3281	4.6740	-10.7285	13.3766
sinEU	1.0739	1.5659	4.1212	-9.9227	11.4055
benchEU	1.2208	1.1266	4.9056	-10.3293	14.9691
sinEU-LFT4GBEU	0.5053	0.4909	3.0850	-11.1158	9.8441
mktrf	0.8514	0.7650	4.4926	-10.1200	13.1600
SMB	0.1995	0.1700	1.7634	-5.1300	4.6700
HML	0.2126	0.2500	2.0976	-6.3600	8.2900
WML	1.0084	1.1150	3.3282	-15.5600	13.6900
RMW	0.3190	0.3750	1.6430	-7.3300	6.3800
CMA	0.1680	1.2510	1.2510	-3.5600	5.3700

* The statistics presented are computed across 218 observations.

Table 14: Summary statistics in recession periods in Europe

Variables	Mean	Median	Std. Dev.	Min	Max
Dcrise	1	0	1	1	1
LFT4GBEU	-2.2343	-2.1608	8.5891	-20.7200	13.3780
sinEU	-0.8776	-0.0834	6.8261	-19.6278	12.2391
benchEU	-1.3174	-1.0672	10.4802	-25.7889	21.5832
sinEU-LFT4GBEU	1.2363	1.0870	4.5048	-8.3903	14.2140
mktrf	-1.8650	-1.9650	9.0001	-22.0600	13.7400
SMB	-0.1142	0.2850	2.7343	-5.2700	4.8600
HML	-0.0133	-0.0300	2.5960	-4.1100	6.3400
WML	-0.2429	1.1550	8.0310	-26.2700	10.4500
RMW	0.3404	0.4300	1.5932	-2.2400	4.3000
CMA	0.3079	0.1300	2.9338	-5.4700	5.3300

* The statistics presented are computed across 218 observations.

In the European market, during periods of expansion (dcrise=0) the excess returns of the FTSE4Good and “Sin” portfolios are higher comparing to returns in periods of recession (dcrise=1). These findings are in line with the results previously mentioned for the American market.

Just like in the analysis of the US data, when observing the difference between the “Sin” portfolio and FTSE4Good portfolio, it is higher in economic downturns than in periods of expansion, which may mean that in periods of recession, the returns of the FTSE4Good portfolios are much lower comparing to the returns of the “Sin” portfolios. The portfolios in periods of recession tend to have negative excess returns.

Again, this evaluation is just preliminary and, in the next tables, it will be presented the coefficients of the variables of the following models: CAPM, Fama and French (1993) three-factor model, Carhart (1997) four-factor model and Fama and French (2015) five-factor model.

Table 15: Descriptive statistics in recession periods in Europe: FTSE4Good portfolio and Sin portfolio

	FTSE4Good Portfolio (1)				Sin Portfolio (2)			
	CAPM	FF 3-factor	Carhart	FF 5-factor	CAPM	FF 3-factor	Carhart	FF 5-factor
Alpha	0.0076 (0.0711)	0.2491** (0.0498)	0.1276 (0.1361)	0.1984** (0.0458)	0.3030 (0.3757)	-0.3564 (0.5144)	0.5206 (0.6471)	-0.1867 (0.4950)
Beta	0.9409*** (0.0793)	0.6853*** (0.0590)	0.7925*** (0.1166)	0.7047*** (0.0501)	0.3119 (0.4400)	0.9831** (0.5562)	0.2096 (0.6785)	0.8937* (0.5117)
SMB		-0.3006*** (0.0590)	-0.2503** (0.0919)	-0.3108*** (0.0414)		0.7821* (0.4214)	0.4192 (0.4398)	0.7876* (0.4625)
HML		-0.0500 (0.0445)	-0.0175 (0.0569)	0.0660 (0.0464)		0.3529 (0.2297)	0.1183 (0.2091)	0.1166 (0.3125)
WML			-0.0390 (0.0414)				0.2811* (0.1176)	
RMW				0.0992 (0.0660)				0.0142 (0.3015)
CMA				-0.1670*** (0.0471)				0.3708 (0.3015)
Adjusted r-squared	0.9893	0.9945	0.9944	0.9954	0.7390	0.7662	0.7754	0.7498

* All specifications include a constant term and are based on 218 observations. Newey-West Standard-errors in parenthesis. *** denote p-values <0.01, ** denote p-values <0.05, and * denote p-values <0.1

Table 16: Descriptive statistics in recession periods in Europe: Difference portfolio

Difference Portfolio (2-1)				
	<i>CAPM</i>	<i>FF 3-factor</i>	<i>Carhart</i>	<i>FF 5-factor</i>
Alpha	0.3071 (0.4055)	-0.5773 (0.5242)	0.4298 (0.7343)	-0.3430 (0.4814)
Beta	-0.6397 (0.4739)	0.2726 (0.5635)	-0.6158 (0.7481)	0.1547 (0.4900)
SMB		1.0663*** (0.3829)	0.6496 (0.4394)	1.0806** (0.4313)
HML		0.3764 (0.2397)	0.1070 (0.2259)	0.0152 (0.3209)
WML			0.3228* (0.1406)	
RMW				-0.0598 (0.3584)
CMA				0.5553 (0.3098)
Adjusted r-squared	0.2861	0.4558	0.4930	0.4482

* All specifications include a constant term and are based on 218 observations. Newey-West Standard-errors in parenthesis. *** denote p-values <0.01, ** denote p-values <0.05, and * denote p-values <0.1

- **General Discussion**

In table 15, analysing the FTSE4Good portfolio, alpha is positive in all the models and although it is only significant in the Fama and French (1993) three-factor model and Fama and French (2015) five-factor model, this evidence indicates an outperformance of this portfolio, during a recession period, in Europe. In the “Sin” portfolio, the findings do not show any significant alphas.

Analysing the betas, in the FTSE4Good portfolio, all of them are lower than 1 and significance at 1% level. This indicates a lower risk comparing to the market. In the “Sin” portfolio, only betas of the Fama and French (1993) three-factor model and Fama and French (2015) five-factor model are significant and inferior to 1. The adjusted r-squared values that are above 90% point to close approximation of the FTSE4Good portfolio to the market.

In table 16, all the alphas are insignificant. This indicates that this hedge portfolio, by controlling for the different factors, does not show a different performance. In the difference portfolio, it was also found a positive and significant SMB factor, in two of the models, Fama and French (1993) three-factor model and Fama and French (2015) five-factor model. The momentum is positive at the 10% significance level because the FTSE4Good portfolio tilts towards a negative momentum, according to the authors Lobe and Walkshäusl (2014).

- **CAPM**

Analysing the table 15, both alphas are positive but not significant. The FTSE4Good portfolio is less sensitive to the market risk, exhibiting a beta that is lower than 1 and with a significance level of 1%, while the “Sin” portfolio exhibits a non-significant beta.

- **Fama and French three-factor model**

In the FTSE4Good portfolio, alpha is positive at the 5% significance level.

Beta is lower than 1, with a 1% and 5% significance level, in the FTSE4Good portfolio and “Sin” portfolio, respectively. The SMB factor, in the FTSE4Good portfolio, is negative and significant, which means that this portfolio indicates a tilt towards large-cap companies. In contrast, the “Sin” portfolio has a positive and significant SMB estimate, which show a small cap tilt. The HML factor, in both portfolios, is not significant.

- **Carhart four-factor model**

In the FTSE4Good and “Sin” portfolios, the alphas are positive, but they are not statistically significant and only the beta of the FTSE4Good portfolio is statistically significant and it is lower than 1. This evidence indicates that, the constituents of the portfolio have lower risk comparing to the market.

The estimates of the HML factor are not significant in both portfolios. The size factor in the FTSE4Good portfolio is negative and statistically significant, indicating a big size tilt. The estimate for the WML factor is positive and significant in the “Sin” portfolio, which shows that the constituents of this portfolio are mostly winner stocks comparing to the benchmark.

- **Fama and French five-factor model**

In this specification, the FTSE4Good portfolio presents a positive and significant alpha, which is an indicator of an outperformance. The beta in both portfolios is lower than 1 and it is statistically significant.

In the results, it is found a contrast in the value of the SMB factor in both portfolios. Although being both significant, the FTSE4Good portfolio tilts towards big size companies and the “Sin” portfolio towards small size companies. Furthermore, the estimate for the HML factor is not statistically significant in both portfolios, but they are both positive. In the “Sin” portfolio, the estimate for both RMW and CMA factors are not statistically significant. The FTSE4Good portfolio presents a negative and significant (at 1% level) CMA factor. This value indicates that this portfolio follows a more aggressive investment strategy.

Table 17: Descriptive statistics in expansion periods in Europe: FTSE4Good portfolio and Sin Portfolio

	FTSE4GOOD Portfolio (1)				Sin Portfolio (2)			
	CAPM	FF 3-factor	Carhart	FF 5-factor	CAPM	FF 3-factor	Carhart	FF 5-factor
Alpha	-0.1160** (0.0532)	0.2407*** (0.0443)	0.2192*** (0.0422)	0.2311*** (0.0434)	0.2064 (0.2151)	0.2873 (0.2765)	0.4909** (0.2683)	0.3722* (0.1914)
Beta	1.1528*** (0.0611)	0.7452*** (0.0472)	0.7619*** (0.0449)	0.7430*** (0.0473)	0.4934* (0.2333)	0.4724* (0.2918)	0.3145 (0.2896)	0.4900** (0.2133)
SMB		-0.3275*** (0.0286)	-0.3128*** (0.0268)	-0.3175*** (0.0230)		-0.0070 (0.1889)	-0.1462 (0.1474)	-0.0920 (0.10063)
HML		0.0193 (0.0260)	0.0127 (0.0251)	-0.0198 (0.0335)		-0.3755*** (0.1433)	-0.3122*** (0.1302)	-0.0547 (0.1765)
WML			-0.0292** (0.0150)				0.2765*** (0.0784)	
RMW				-0.1141*** (0.0291)				0.9749*** (0.2117)
CMA				-0.0271 (0.0367)				0.2664 (0.1652)
Adjusted r-squared	0.9768	0.9871	0.9874	0.9879	0.6056	0.6333	0.6713	0.7185

* All specifications include a constant term and are based on 218 observations. Newey-West Standard-errors in parenthesis. *** denote p-values <0.01, ** denote p-values <0.05, and * denote p-values 0.1

Table 18: Descriptive statistics in expansion periods in Europe: Difference portfolio

Difference Portfolio (2-1)				
	CAPM	FF 3-factor	Carhart	FF 5-factor
Alpha	0.3416 (0.2202)	0.0672 (0.2832)	0.2903 (0.2717)	0.1624 (0.1899)
Beta	-0.6813** (0.2419)	-0.2950 (0.2963)	-0.4681* (0.2923)	-0.2757 (0.2113)
SMB		0.3203* (0.2090)	0.1678 (0.1629)	0.2258** (0.1165)
HML		-0.4008*** (0.1563)	-0.3315*** (0.1416)	-0.0462 (0.1913)
WML			0.3030*** (0.0830)	
RMW				1.0852*** (0.2231)
CMA				0.3032* (0.1708)
Adjusted r-squared	0.2102	0.2803	0.3622	0.4703

* All specifications include a constant term and are based on 218 observations. Newey-West Standard-errors in parenthesis. *** denote p-values <0.01, ** denote p-values <0.05, and * denote p-values <0.1

- **General Discussion**

In table 17, the alphas, in the different asset pricing models of the FTSE4Good portfolio, show some differences. For starters, the alpha in the CAPM is negative, but this changes in the other models. The value of the alpha becomes positive and with a significance level of 1%, which represents an outperformance of this portfolio.

In the “Sin” portfolio, only the alphas of the two last specifications are significant and positive. In contrast to other models, the significance levels of 5%

and 10%, in the Carhart (1997) four-factor model and Fama and French (2015) five-factor model, respectively, point to an outperformance of this portfolio.

Once again, the adjusted r-squared values are higher than 90% in the FTS4Good portfolio, indicating a good approximation of the different specifications to the benchmark.

Analysing the estimates in table 18, alphas are not conclusive because they are insignificant. The betas in the hedge portfolio are negative in the different specifications but only the values of the CAPM and Carhart (1997) four-factor model are significant. This indicates that the “Sin” portfolio has a beta smaller than 1. The HML factor is negative and significant in the Fama and French (1993) three-factor model and Carhart (1997) four-factor model. This factor shows a tilt towards growth stocks that can be explained by the fact that the “Sin” portfolio exhibits a higher proportion of growth stocks in its composition. The difference portfolio also presents a positive momentum loading during an expansion period.

•CAPM

During an expansion period, the FTSE4Good portfolio exhibits a negative and significant alpha, which mean that this portfolio would underperform comparing to the market. In contrast, the alpha is positive in the “Sin” portfolio but not significant.

In the FTSE4Good portfolio, the beta’s estimate is above 1, which shows a higher risk of this portfolio comparing to the market. However, the “Sin” portfolio exhibits a beta below 1, indicating that this portfolio is less sensitive to the market risk.

•Fama and French three-factor model

In this model, the FTSE4Good portfolio presents a positive and significant alpha. In contrast with the previous model, this model points to an

outperformance of this portfolio. Beta is below 1 and significant in both portfolios, showing once again that these portfolios are less sensitive to the market risk.

Analysing the estimate for the HML factor, the “Sin” portfolio presents a negative value, which means that this portfolio is mostly composed by growth stocks. The FTSE4Good portfolio is tilted towards large-cap companies.

- **Carhart four-factor model**

In the FTSE4Good and “Sin” portfolios, the alphas are both positive and statistically significant. Only the beta of the FTSE4Good portfolio is significant and lower than 1, which means that the constituents of the portfolios present a lower risk comparing to the market. The SMB factor, in the FTSE4Good portfolio is negative and statistically significant, showing a large cap tilt. The “Sin” portfolio shows that the portfolio is composed mostly by growth stocks. The estimate of the WML factor is also statistically significant and its value is positive in the “sin” portfolio, which means that the portfolio is mostly composed by winner stocks. In contrast, the FSTE4Good portfolio, presents a negative value, indicating a higher proportion of loser stocks.

- **Fama and French five-factor model**

Both portfolios have statistically significant alphas and both betas are also statistically significant and lower than 1. As in the previous model, these results point to an outperformance.

Comparing with the previous results in Carhart (1997) four-factor model, the SMB factor of FTSE4Good portfolio keeps its tilt towards large-cap companies with a weak profitability. The HML factor presents negative values in both portfolios but they are not significant. Both portfolios present insignificant CMA estimates.

5. Conclusion

Extant literature presents mixed evidence regarding the performance of the SRI stocks and strong evidence concerning the outperformance of “sin” stocks. This study presents important findings that extend empirical literature.

While analysing the “Sin” portfolios, the results show a positive and statistically significant alpha in some of the specifications. Specially, in the “Sin” Europe portfolio, it is found an outperformance denoted by the positive and significant alpha found in the Carhart (1997) four-factor model and in the Fama and French (2015) five-factor model. The “Sin” US portfolio presents some significant results however, it loses its significance in the Fama and French (2015) five-factor model.

Although prior evidence is mixed regarding SRIs, the significant positive values obtained in different specifications support the statement of outperformance for these portfolios, especially in the FTSE4Good Index. The superior performance found, in both FTS4Good portfolios, comparing to the benchmark and controlling for the factors, reveals that investors who choose SRI’s may have profits in the long-term.

We extend the extant empirical studies by using the Fama and French (2015) five-factor model. This work shows that FTSE4Good’s stocks are prone to have a weak operating profitability, whereas “sin” stocks that compose the portfolios tend to have a robust operating profitability and a more conservative investment strategy.

Another important insight taken from these analyses is that, the FTSE4Good portfolios tend to bear more market risk than the “Sin” portfolios.

During economic upturns, the FTS4Good Europe portfolio and the “Sin” Europe portfolio present a superior performance. Despite these results, the same cannot be concluded from the American portfolio due to insignificant findings.

In this study it was not found evidence that supports Salaber (2007) argument that states that during periods of crisis, “sin” stocks outperform due to the fact of these firms not being as sensitive to changes of the economic conditions. In contrast it is found during recession periods, an outperformance of the FTSE4Good Europe portfolio in the Fama and French (1993) three-factor model and in the Fama and French (2015) five-factor model.

Concerning the hedge approach, with a long position in the “Sin” portfolios and a short position in FTSE4Good portfolios, in periods of crisis and non-crisis, results do not show evidence of underperformance or outperformance. This evidence is supported by Lobe and Walkshäusl (2014), in which the authors state that vice investing and socially responsible investing are at the same level in terms of performance. This statement indicates that the decision in the type of investment depends on the preferences of the investor.

The motivations behind choosing between FTSE4Good stocks (SRI) and “sin” stocks may vary depending on the profile of the investor, that consists on its preferences and beliefs. For example, “sin” stocks may be tricky to classify as an unethical or ethical stock because this classification depends on an investor’s ethical values and beliefs towards an industry. Salaber (2007) points out that “sin” stocks may not be the most appealing to investors who are more inclined to invest in stocks that present certain ethical and social standards, even though “sin” stocks have higher risk-adjusted returns.

Hong and Kacperczyk (2009) once again, reinforces the fact that although “sin” stocks present higher abnormal returns and financial reports with a higher quality, investors more driven by social norms tend to decline these stocks.

Although this work is not able to make a clear statement on what may motivate an investor to choose between a “Sin” portfolio or a SRI portfolio (in this case, a FTSE4Good portfolio), the lack of indicator signs of underperformance or outperformance when comparing both, once again support the idea that it may depend on the investor’s profile.

In conclusion, the contributions of this work to the existent literature are: i) The fact that is not only employed the CAPM, the Fama and French (1993) three-factor model, the Carhart (1997) four-factor model but also the Fama and French (2015) five-factor model, which has not been done before in the context of the comparison between “sin” stocks and SRI stocks’ performance; ii) it is also conducted an analysis in the American and European markets of the FTSE4Good, “Sin” and the difference portfolios (long in the “sin” stocks and short in the FTSE4Good stocks) in downturns and upswings’ market periods, which allows a more detailed insight.

The results obtained in this study, were not statistically significant in some cases and future researches still have some questions to be answered or better explained. In order to try to avoid inconclusive results as seen in some of the regressions of this work, one of the solutions that could be proposed to future research would be to have a bigger sample in the “sin” portfolios or realize other robustness tests. For example, in addition to analyse the subsamples of crisis and non-crisis periods, it could be analysed industry adjusted regressions. Another possible contribution to future research, could be to employ conditional models to evaluate the performance of the portfolios and compare it with the results of the unconditional asset pricing models.

References

- Admati, Anat R., Ross, Stephen A., (1985), *Measuring Investment Performance in a Rational Expectations Equilibrium Model*, The Journal of Business, Vol. 58, No.1, pp. 1-26.
- Akerlof, G., (1980), *A Theory of Social Custom, of which Unemployment May be One Consequence*, Quarterly Journal of Economics, Vol 94, No.4, pp. 749-775.
- Barnett, M., Salomon, R., (2006), *Beyond dichotomy: the curvilinear relationship between social responsibility and financial performance*, Strategic Management Journal, Vol. 21, pp. 1101-1122.
- Bauer, R., Koedijk, K., Otten R., (2005), *International Evidence on Ethical Mutual Fund Performance and Investment Style*, Journal of Banking & Finance, Vol. 29, No. 7, pp. 1751–1767.
- Becchetti, L., R. Ciciretti, A. Dalò, S. Herzel, (2015), *Socially Responsible and Conventional Investment Funds: Performance Comparison and the Global Financial Crisis*, Applied Economics, Vol. 47, No. 25, pp. 2541-2562.
- Belghitar, Yacine; Clark, Ephraim; Deshmukh, Nitin (2014), *Does it pay to be ethical? Evidence from the FTSE4Good*, Journal of Banking and Finance, pp. 54-62.
- Bello, Z. Y., (2005), *Socially Responsible Investing and Portfolio Diversification*, Journal of Financial Research, Vol. 28, N. 1, pp. 41-57.
- Black F, Jensen M, Scholes M., (1972), *The capital asset pricing model: Some empirical tests*, Praeger: New York, pp. 79–121.
- Blitz, D., Fabozzi, F., (2017), *Sin Stocks Revisited: Resolving the Sin Stock Anomaly*, The Journal of Portfolio Management, Vol. 44, pp. 105-111.

- Capelle-Blancard, G., Monjon, S., (2014), *The Performance of Socially Responsible Funds: Does the Screening Process Matter?*, European Financial Management, pp. 1-27.
- Carhart, Mark M., (1997), *On Persistence in Mutual Fund Performance*, The Journal of Finance, Vol. 52, No.1, pp. 57-82.
- Chen, Nai-Fu, (1983), *Some Empirical Tests of the Theory of Arbitrage Pricing*, The Journal of Finance, Vol. 38, No 5, pp. 1394-1414.
- Chong, J., Her, M., Phillips, G., (2006) *To sin or not to sin? Now that's the question*, Journal of Asset Management, Vol. 6, No. 6, pp. 406–417.
- Chordia, T., Shivakumar, L. (2002), *Momentum, Business Cycle and Time-Varying Expected Returns*, Journal of Finance. Vol. 57, pp. 985-1019.
- Collison, David; Cobb, George; Power, David; Stevenson; Lorna (2008), *The Financial Performance of the FTSE Good Indices*, Corp. Soc. Responsib. Environ. Mgmt, Vol. 15, pp. 14-28.
- Cortez, MC., Silva, F., Areal N., (2009), *The performance of European socially responsible funds*, Journal Business Ethics, Vol. 87, No. 4, pp. 573–588.
- Durand, Robert; Koh, Szekee; Limkriangkrai, Manapon, (2013), *Saints versus Sinners. Does morality matter?*, Journal of International Financial Markets, Institutions & Money, Vol. 24, pp. 166-183.
- Dybvig, Philip H., Ross, Stephen A., (1985), *Yes, The APT Is Testable*, The Journal of Finance, Vol. 40, No.4, pp. 1173-1188.
- Fabozzi, F., Ma, K.C. and Oliphant, B. (2008), *Sin stock returns*, The Journal of Portfolio Management, Vol. 35, pp. 82-94.
- Fama, Eugene F., (1968), *Risk, Return and Equilibrium: Some Clarifying Comments*, The Journal of Finance, pp. 29-40.
- Fama, Eugene F., French, Kenneth R., (1989), *Business Conditions and Expected Returns on Stocks and Bonds*, Journal of Financial Economics, Vol. 25, pp. 23-49.

- Fama, Eugene F., French, Kenneth R., (1992), *The Cross-Section of Expected Stock Returns*, The Journal of Finance, Vol. 47, No.2, pp. 427-465.
- Fama, Eugene F., French, Kenneth R., (1993), *Common risk factors in the returns on stocks and bonds*, The Journal of Financial Economics, Vol. 33, pp. 3-56.
- Fama, Eugene F., French, Kenneth R., (2015), *A five-factor asset pricing model*, Journal of Financial Economics, Vol. 116, pp. 1-22.
- Ferson, W.E. and Schadt, R.W., (1996), *Measuring fund strategy and performance in changing economic conditions*, The Journal of Finance, Vol. 51, No. 2, pp. 425–461.
- Frazzini, Andrea; Pedersen, Lasse H., (2014), *Betting against beta*, Journal of Financial Economics, Vol. 111, pp. 1-25.
- Ghoul, S., Guedhami, O., Kwok, C., Mishra D., (2011), *Does corporate social responsibility affect the cost of capital?*, Journal of Banking and Finance, Vol. 35, pp. 2388–2406.
- Goldreyer, E.F., Diltz, J.D, (1999), *The Performance of Socially Responsible Mutual Funds: Incorporating Socio-political Information in Portfolio Selection*, Managerial Finance, Vol. 25, No. 1, pp. 23–26.
- Gregory A, Matatko J, Luther RG, (1997), *Ethical unit trust financial performance: small company effects and fund size effects*, Journal of Business, Finance and Accounting, Vol 24, No. 5, pp. 705–725.
- Grinblatt, Mark; Titman, Sheridan, (1989), *Portfolio Performance Evaluation: Old Issues and New Insights*, The Review of Financial Studies, Vol. 2, No.3, pp. 393-421.
- Hamilton S, Jo H, Statman M., (1993), *Doing well while doing good? The investment performance of socially responsible mutual funds*, Financial Analysts Journal, Vol 49, No. 6, pp. 62–66.
- Hong H., Kacperczyk M., (2005), *The price of sin: the effects of social norms on markets*, Working paper.

- Hong, Harrison; Kacperczyk, Marcin, (2009), *The price of sin: The effects of social norms on markets*, Journal of Financial Economics, Vol. 93, pp. 15-36.
- Humphrey, J.E., Tan, D.T., (2013), *Does it really hurt to be responsible?*, Journal of Business Ethics, Vol. 1055, pp. 375-386.
- Jagannathan, Ravi; Wang, Zhenyu, (1996), *The conditional CAPM and the cross-section of expected returns*, The Journal of Finance, Vol. 51, No. 1, pp. 3-52.
- Jegadeesh, N., Titman, S., (1993), *Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency*, Journal of Finance, Vol. 48, pp. 965-991.
- Jensen, Michael C., (1968), *The performance of mutual funds in the period 1945-1964*, The Journal of Business, Vol. 42, No. 2, pp. 167-247.
- Jensen, Michael C., (1969), *Risk, the pricing of capital assets, and the evaluation of investment portfolios*, The Journal of Business, Vol 42, No. 2, pp. 167-247.
- Kreander N, Gray R., Power D., Sinclair C., (2005) *Evaluating the performance of ethical and non-ethical funds: a matched pair analysis*, Journal Business, Finance and Accounting, Vol 32, No. 7, pp.1465–1493.
- Lesser, K., Lobe, S., Walkshäusl, C., (2014), *Green And Socially Responsible Investing In International Markets*, Journal of Asset Management, Vol. 15, No. 5, pp. 317-331.
- Lintner, John, (1965), *Security prices, risk and maximal gains from diversification*, The Journal of Finance, Vol. 20, No.4, pp. 587-615.
- Liston, D. P., Soydemir, G., (2010), *Faith-based and sin portfolios: An empirical inquiry into norm-neglect vs norm-conforming investor behaviour*, Managerial Finance Vol. 36, No. 10, pp. 876–885.
- Lobe, S., Walkshäusl, C., (2014), *Vice versus virtue investing around the world*, Review of Managerial Science, pp. 1-42.
- Luck, C., Wood, K., (1992), *Wages of sin: ethical investing and performance*, The Journal of Investing, Vol. 1, No. 2, pp. 15–21.

- Luther, RG., Matatko J., Corner, DC., (1992) *The investment performance of UK "ethical" unit trusts*, Account Audit Account Journal, Vol 5, No. 4, pp. 57–70.
- Luther, RG., Matatko, J., (1994) *The performance of ethical unit trusts: choosing an appropriate benchmark*, The British Accounting Review, Vol. 26, No. 1, pp. 77–89.
- Mallin, CA., Saadouni, B., Briston, RJ., (1995), *The financial performance of ethical investment funds*, Journal of Business Finance and Accounting, Vol. 22, No. 4, pp. 483–496.
- Markowitz, Harry, (1952), *Portfolio Selection*, The Journal of Finance, Vol.7, No.1, pp. 77-91.
- Merton, R. (1987), *A simple model of capital market equilibrium with incomplete information*, Journal of Finance, Vol. 42, pp. 483–510.
- Moskowitz, MR., (1972), *Choosing socially responsible stocks*, Business & Society, Vol. 1, No. 1, pp. 71–75.
- Newey, W. K., West, K. D., (1987), *Hypothesis Testing with Efficient Method of Moments Estimation*, International Economic Review, Vol. 28, N. 3, pp. 777-787.
- Renneboog, L., Horst, J., Zhang, C., (2008), *The Price Of Ethics And Stakeholder Governance: The Performance Of Socially Responsible Mutual Funds*", Journal of Corporate Finance, Vol. 14, No. 3, pp. 302-322.
- Richey, Greg, (2016), *Sin Is In: An Alternative to Socially Responsible Investing?*, Journal of Investing, pp. 136-143.
- Richey, Greg, (2017), *Fewer reasons to sin: a five-factor investigation of vice stock returns*, Managerial Finance Journal, pp. 1-36.
- Roll, Richard, (1977), *A critique of the asset pricing theory's tests- Part 1: On past and potential testability of the theory*, Journal of Financial Economics, Vol. 4, pp. 129-176.

- Roll, Richard; Ross, Stephen A., (1980), *An Empirical Investigation of the arbitrage pricing theory*, The Journal of Finance, Vol. 35, No. 5, pp. 1073-1102.
- Ross, Stephen A., (1976), *The arbitrage theory of capital asset pricing*, Journal of Economic Theory, Vol. 13, pp. 341-360.
- Salaber, J., (2007), *Sin stock returns over the business cycle*, unpublished manuscript, University of Bath School of Management, Bath.
- Sauer, D., (1997), *The impact of social-responsibility screens on investment performance: evidence from the Domini 400 social index and Domini equity fund*, Review of Financial Economics, Vol. 6, No. 2, pp. 23–35.
- Scholtens, B., (2005), *Style and performance of Dutch socially responsible investment funds*, The Journal of Investing, Vol. 14, No. 1, pp. 63–72.
- Schröder, Michael, (2007), *Is There a Difference? The Performance Characteristics of SRI Equity Indices*, Journal of Business & Accounting, pp. 331-348.
- Sharpe, William F., (1964), *Capital Asset Prices: A theory of market equilibrium under conditions of risk*, The Journal of Finance, Vol 19, No. 3, pp. 425-442.
- Statman, Meir., (2000), *Socially Responsible Mutual Funds*, Financial Analysts Journal, Vol. 56, No. 3, pp. 30–39.
- Statman, M., (2006), *Socially responsible indexes: composition, performance, and tracking error*, Journal of Portfolio Management, Vol. 32, pp. 100–109.
- Statman, M., Glushkov, D., (2009) *The wages of social responsibility*, Financial Analysts Journal, Vol. 65, pp. 33–46.
- Treynor, Jack L., (1961), *Market value, time, and risk*, unpublished manuscript.
- Treynor, Jack L., (1962), *Toward a theory of market value of risky assets*, unpublished manuscript.

- Trinks, Peter; Scholtens, Bert, (2017), *The Opportunity Cost of Negative Screening in Socially Responsible Investing*, Journal of Business Ethics, Vol. 140, pp. 193-208.
- White, M. (1995), *The Performance of Environmental Mutual Funds in the United States and Germany: Is There Economic Hope for “Green” Investors?*, Research in Corporate Social Performance and Policy, pp. 323-344.