CATÓLICA
CATÓLICA PORTO BUSINESS SCHOOL

PORTO

Portfolio Performance Evaluation and Gender Diversity in Europe

João Eduardo Lima da Costa

Católica Porto Business School
November 2020

PORTO

Portfolio Performance Evaluation and Gender Diversity in Europe

Dissertation submitted at Universidade Católica Portuguesa in order to fulfill the requirements for the MSc in Finance

By

João Eduardo Lima da Costa

Written under the supervision of Professor João Pinto and co-supervision of Professor João Novais

Católica Porto Business School

November 2020

Acknowledgments

I would like to express my gratitude to all the professors who helped along my journey, in special to Professor Dr João Pinto, Professor Dr Ricardo Ribeiro and Professor Dr João Novais for their availability for listening to my diverse doubts as well as their constructive feedback, always pushing me to think further and try to understand where improvements could be made.

To my parents, Asdrúbal Costa and Carla Lima, for their relentless support throughout my entire life, never leaving me lost and giving me the freedom to decide by myself, even though my decisions could be wrong. We do not always agree, but I could not have asked for better guidance than the one I received.

To Mariana, for allowing me to be the older brother and strive to give the best example. For all the conversations we had, I am sure you will find this thesis of your interest and I sincerely hope that you grow up into the future that you dream of. I will always be watching you.

To all my friends, which gave that extra energy on the moments where the motivation was not present. In special, to Inês for her patience and for always being there, no matter what or who, to give me that so needed advice.

Last but not least, to my grandfather, Fernando, which is the person that would go miles and miles just to put a smile on his grandchildren. You are the person I wanted more to be here and look into this work, but you are miles and miles away now. Nevertheless, this is for you.

Abstract

In the recent decade, gender diversity has become an ever-standing subject. Inside this theme, many factors have been mentioned and studied, making headlines in the media, like gender pay gap. Given this, a shift in the companies' objectives has happened, where they are concerned not only with their financial goals but also with their social impact. Notwithstanding, the financial and social components are not separated and that is exactly what this paper intends to analyze by taking into account the gender diversity and checking whether such factor has an impact in the company's financial performance and, if so, whether it is positive or negative.

In order to do that, we need to first introduce the support for portfolio performance evaluation which can be divided between the conditional and the unconditional performance evaluation. The former takes into account the dynamic economy we live in, although with more complex models (also more prone for misspecification errors), while the latter assumes a static reality, allowing for more practical models, that are easily applicable to a set of data. From these, it is highlighted the models that make a portfolio's returns to depend on three, four and five factors, which are also the ones applied to the set of data analyzed.

Our portfolios were built using a document from EWOB (*European Women on Boards*) where the companies from Stoxx 600 are given a score, based on the ratio of women that occupy a board position on these companies' boards. With this document, two types of portfolios were formed: Top 20 (with a high ratio of women on board) and Bottom 20 (with a low ratio of women on board).

In the end, we found that both the four and the five-factor model suited the data adequately, with our analysis suggesting that the presence of more women on the companies' boards is not connected with a better performance.

Keywords: Conditional/Unconditional Portfolio Performance; Gender Diversity; Stoxx 600;

Resumo

Na década passada, a diversidade de género tornou-se num tema omnipresente. Diversos fatores foram analisados sendo que alguns, como a diferença salarial entre géneros, teve grande impacto mediático. Dado isto, observou-se uma alteração nos objetivos das empresas, verificando-se maior preocupação não só com a performance financeira, mas também com o impacto social. Não obstante, ambas as partes estão interligadas e é exatamente isso que este trabalho de final de mestrado pretende analisar, verificando se a diversidade de género tem impacto na performance financeira da empresa e, caso tenha, se o mesmo é positivo ou negativo.

De modo a atingir esse objetivo, é necessário introduzir o suporte teórico para a análise de performance de portefólios, que pode ser dividido em condicional ou incondicional. O primeiro tem em conta a economia dinâmica na qual estamos inseridos, apresentando modelos mais complexos (também mais propícios a erros de especificação), enquanto o último assume uma realidade estática, permitindo modelos mais práticos, facilmente aplicados a um conjunto de dados. Destes últimos, são realçados os modelos que fazem o retorno de um portefólio depender de três, quatro e cinco fatores, sendo os modelos aplicados ao conjunto de dados que analisamos.

Construímos o nosso portefólio através do documento da organização EWOB (*European Women On Board*) onde as empresas que constituem o Stoxx 600 são classificadas de acordo com o rácio de mulheres que ocupam posições nos seus Conselhos de Direção. Assim, foram criados dois tipos de portefólios: Top 20 (com um rácio elevado de mulheres no Conselho de Direção) e o Bottom 20 (com um rácio baixo de mulheres no Conselho de Direção).

No final, concluímos que ambos os modelos de quatro e cinco fatores se encontravam bem ajustados aos nossos dados, sugerindo que a presença de mais mulheres nos Conselhos de Direção das empresas não está associado a um incremento de performance.

Palavras-Chave: Performance Condicional/Incondicional de Portefólio; Diversidade de Género; Stoxx 600;

Table of Contents

List of tables

Chapter 1 - Introduction

Diversity on boards is not a new issue on corporate governance. Across Europe, many countries have addressed it throughout the years, by taking different paths: thorough legislation, through quotas and even through voluntary targets. Notwithstanding, it is now reaching some important milestones, setting the standards higher. As of May 2020, FTSE 350 Index reached 33% of women on boards, being this accomplishment built over 9 years, from a 10% average on 2011. In contrast, the S&P500, as of June 2020, has only 31 female CEO's, as opposed to 81 in 2000. Even though the time intervals aren't comparable, it still shows how different the development has been among markets in the whole world.

Corporate governance's impact in financial performance has been studied in past years, but such a concept has many factors included which may not allow to retrieve the actual effect of each factor. Having said that, more recent literature has focused into one of its components, gender diversity, and its impact on the financial performance of the companies and its subsequent returns to the company's shareholders.

By itself, financial performance has been a theme for a long time, with works like the one developed by Markowitz (1952) igniting the financial world into asset pricing. The former author introduced a normative theory of how an investor should behave, reinforcing the importance of diversification and the trade-off between expected return and variance. Afterwards, several authors added to this work, presenting one model that, even nowadays, is used: CAPM (Capital Asset Pricing Model). Such model can be said to be originated in Treynor (1961), Sharpe (1964) and Lintner (1965). Nevertheless, this model was not perfect and, while some authors tested its robustness, as in Black et al. (1972), or tried to extend it, like in Merton (1973), others tried to propose different alternatives, as the APT model by Ross (1976).

However, the CAPM made the excess return of an asset to be linearly dependent only on the asset's sensibility to market variations (measured by its Beta) and some authors discovered that more factors should be considered when trying to explain assets' returns, as size in Banz (1981) and debt-to-equity ratio in Bhandari (1988). All these works culminated in the birth of further models who aimed at explaining the average returns, starting with Fama & French (1993) three-factor model (adding the size of the company and its book-to-market ratio as factors to be considered, besides the market factor already proposed in the CAPM). Such a model was followed by the four-factor model, developed by Carhart (1997), by joining the work of Jegadeesh & Titman (1993) on the momentum factor to the previous model. Notwithstanding, Fama & French (2015) came back again with an improvement on their previous model, while not using the momentum factor since the tests they realized did not reveal a significant impact from this factor. Instead, they added factors that reflected the company's profitability (RMW) and investment profile (CMA).

Despite being different among themselves, these models assume a static reality where loading of factors, like Beta, remain constant over time. Some authors disagreed with this view and decided to propose an alternative where the performance of the assets was conditional and Beta would vary over time, in the hope of reflecting the impact that our dynamic economy has. Jagannathan & Wang (1996) did exactly that by extending CAPM to a conditional performance, applying what Ferson & Harvey (1991) already had mentioned regarding time-variation in Beta, and reinforced by their empirical analysis in Ferson & Harvey (1999). Additionally, certain authors also tested these models on mutual funds, as Ferson & Schadt (1996), while making use of Jensen (1968) constant, but modifying it in order to make it conditional on the information available and finding that mutual funds' Beta are correlated with public information.

Nonetheless, these models had a complex methodology and certain authors proposed different alternatives to address this issue, like Wang (2003) who applied a simpler methodology to allow these models to not have errors derived from misspecifications and retain their validity, or Ferson & Siegel (2001) who proposed nonlinear functions of conditional parameters.

More recently, the board composition has also been studied along with its impact on the financial performance of a company. Adams & Ferreira (2007) investigate whether the boards should be divided in order to perform different roles: advisory and monitoring. Despite not specifically studying the impact of having women on boards, it was a support for Campbell & Mínguez-Vera (2008), where such a study was made, in the context of the Spanish economy, showing us that the presence of women on the board, on average, does not destroy value for the shareholders. In a way, this issue is presented as a dichotomy since, on the one hand, women may bring new perspectives and ideas, improving the financial performance of the company, while, on the other hand, they may bring an undesired outcome due to being selected in order to respect a mandatory quota, for example, and not because of their qualifications.

Furthermore, researchers also found that their results were not unbiased, since some companies did not have enough women on their board to do a proper analysis, as reported in Chapple & Humphrey (2014), supporting the actions that were being taken across the world to address gender diversity in boards. At an European level, we can see an example of this measures in Ahern & Dittmar (2012) where they study the impact of Norway's mandatory quota of 40% of the board being composed by women, introduced in 2003. As expected, firms reacted differently to the imposed quota, either reshuffling their members or going private to avoid mandatory compliance with such law.

Connecting gender diversity with stock prices, Gul et. al (2011) studied the impact of the presence of women on the board on the stock price informativeness and found that there was a positive relation between them. Notwithstanding, such positive connections was shown to not only exist on developed markets but also on developing markets, like in Vietnam by Nguyen et. al (2015).

In order to try to contribute to this subject of the impact of gender diversity on firm's performance, we decided to use a document from *European Women On Boards* (EWOB), where they publish a Gender Diversity Index (GDI) on firms from the Stoxx 600, and build two portfolios with the Top 20 firms (the ones with a GDI score above 0.8) and the Bottom 20 firms (GDI score below 0.2), where 1 is gender equality, above and below 1 is female and male dominance, respectively.

For this, we gathered monthly information on the stock price and market value of the companies in the document, from Datastream, for the period beginning on January 2000 and ending on December 2019, and, after organizing the time series data, applied the unconditional three, four and five-factor models, already mentioned above, to the two portfolios, dividing them into two categories: equally-weighted (EW) and value-weighted (VW).

The subsequent outcome suggested that the four and five-factor models are both a good fit, without one having a clear advantage over the other. Nevertheless, on both, our outcome suggested that the presence of women on the boards did not benefit the companies, with the Bottom 20 portfolios appearing to have an advantage. Furthermore, this analysis also took into account different market proxies (the one used in Fama & French (1993) and Fama & French (2015), the Stoxx 600 index and the MSCI index), in order to complement it and allow for further analysis on the sensibility of our portfolios to the market factor.

To summarize, this work consists of 4 chapters. Chapter 1 is the literature review and theoretical background, where the theory of portfolio performance evaluation is exposed, from its first steps until the most recent, taking into account the dynamic economy we are inserted. Still in here, the gender diversity is addressed by making use of the literature available on the subject. Afterwards, on Chapter 2, we describe the methodology of our research, with an explanation of the data gathered, along with the models we intend to estimate. The outcome of this is displayed in Chapter 3, where our regressions are presented, and the subsequent results are discussed. Finally, on Chapter 4, the main points of our work are summarized, the limitations exposed and suggestions for further research are made.

Chapter 2 - Literature Review

2.1 Portfolio Performance Evaluation

Over the last decades, a lot of research has been done regarding security analysis and asset pricing. Markowitz (1952), assessed the process of selecting a portfolio, focusing on the investor's beliefs about the future performance of securities. While acknowledging diversification as a must-have investment behavior, the author also recognizes that the returns from securities are highly intercorrelated, not allowing diversification to fully eliminate the variance of returns in stock prices, which is a measure of risk.

Having said that, at a certain point, the investor is faced with a decision: either he/she tries to increase the expected return, accepting an increase in variance, or he/she tries to reduce variance, giving up expected return in exchange. Given that, Markowitz (1952) shows that investors should try to obtain the maximum expected return possible, for a given level of variance, or the minimum variance possible, for a given level of expected return. This can be seen as a guideline to rational investment rather than speculative behavior.

Such a procedure implies, for a large and representative sample, diversification. Furthermore, even in the case of two portfolios with equal variance, any combination of both will result in a decrease in variance, unless both portfolios are perfectly correlated. Therefore, this diversification requires some insight from investors since it does not depend exclusively on the quantity of different securities held but also on how they correlate with each other. In a way, Markowitz (1952), through a mean-variance approach, presents us with a normative theory of how investors should behave.

An empirical explanation for diversification was then provided by Tobin (1958), which inquired about the holding of a risk-free asset due to expectations of losses in other assets. The investor's decision for mitigating risk, in the form of interest rate risk, will eventually fall in the holding proportion of a risk-free asset and a portfolio, which is composed by diverse assets, following Markowitz (1952) rational procedures.

Notwithstanding, Treynor (1961) disagreed with Tobin (1958) since, in the former's opinion, the investor attempts diversification to mitigate equity risk instead of interest rate risk. For Treynor (1961), the problem lies in selecting a combination between fixed-return and equity assets, hence not considering both price level and interest rate risks. Nevertheless, the author explains that despite both risks being important, when we consider the US economy, equity risk has a bigger impact, and the author uses that as a justification for his focus on it.

With these assumptions properly established, Treynor (1961) develops his model and realizes that a decisive factor for investment, besides the uncertainty present in every investment, is the risk premium. In fact, each investment has its own risk premium, so the overall value is achieved taking into account the portion that each investment represents. Furthermore, the author also states that the risk premium and the standard error share a relationship when we think about efficient combinations, since each of them has the same ratio of risk premium to standard error.

In line with Markowitz (1952), Treynor (1961) concludes that the risk premium in a share belonging to a certain investment is proportional to the covariance of such investment with the total amount of investments in the market. The author also suggests that a distinction between risks that are affected by changes in the general conditions of the market and those that are not, should be made. This makes investments that are exposed to the former to possess a cost of capital equal to the risk-free rate, while the investments that are exposed to the latter require a more thoughtful approach.

On a similar note, and not being aware of Treynor (1961) work, Sharpe (1964) focuses on explaining the relationship between the price of an asset and its risk. It was known that diversification would mitigate some of the risk intrinsic to the asset but that only revealed that the biggest factor for an asset's price was not its total risk, saying nothing about the risk component that is specific to each asset and, therefore, relevant to determine its price.

By joining the works of Tobin (1958) and Markowitz (1952), Sharpe (1964) shows that investors are expected to choose an optimal portfolio, that provides them with the highest utility. However, since assets are correlated among themselves, this makes the portfolio formation to not depend solely on the combinations of expected return and risk. Furthermore, investors' actions in the market causes price movements and, adding the fact that many combinations of risky assets are efficient, investors may opt for different combinations among the options available.

Notwithstanding, when the equilibrium conditions are met, the combinations considered have to possess perfect, and positive, correlations among themselves, since they all lie in the capital market line (constructed by the author to show all the possible combinations following rational procedures). This also corroborates the existence of a simple linear relationship between the expected return and the standard deviation for efficient combinations.

Therefore, Sharpe (1964) characterizes an asset's expected return as dependent on a risk-free rate and the responsiveness of the asset to variations in the economic activity. High responsiveness is expected to have higher expected returns and the opposite happens when there is a low level of responsiveness. To this variability, the author refers to as the systematic risk, which diversification allows the mitigation of, and states that the remainder of the asset's return variations is caused by specific, and unsystematic, components, to which the investor is exposed.

Nevertheless, Lintner (1965) disagreed with Sharpe (1964) regarding the effects from the responsiveness of a stock to market variations. In the former's opinion, the latter had only considered the risk effect, leaving behind the income effect. Both impact the securities expected rates of return the same way but that is not the case when we consider the impact on prices. On the one hand, the risk effect increases the expected rate of return, due to increased uncertainty, decreasing the stock price. On the other hand, the income effect comes as investors may consider that the general conditions of the market may be favorable for a particular stock to have a great performance, raising its expected return and causing an increase in purchases which, in their turn, increase the stock price. In a way, the responsiveness to variations in the market index is not as straightforward as expected and depends on the net result of both risk and income effects.

Furthermore, the author also argues that, to reach Sharpe (1964) results, it would be required to consider an hypothetical situation where all systematic risks are neutralized and all specific risks, are equal to zero. As a consequence, Lintner (1965) also thinks that his diversification approach differs from Sharpe (1964), since the latter would not consider specific risk components in his analysis. For the former, the objective of diversification is not to mitigate risk but rather to select the best combination of expected return and risk, in a portfolio.

To support that, Lintner (1965) establishes criteria to hold stocks within a portfolio, which was related with the portfolio expected rate of return, the variance of its return and how the stock being analyzed would behave if selected (i.e. how it correlates with the portfolio). The author considers this condition to be valid for every stock in the portfolio but brings our attention to the fact that it is not the same as the return per unit of portfolio risk which, according to him, has created a misunderstanding in previous situations.

The disagreement also reached the linear relationship between a security return and its standard deviation, where Lintner (1965) thought that it does not hold when we are considering the security to be within a portfolio. Instead, the author proposes that, under these circumstances, the risk of an individual security is determined by its correlation with other securities.

However, Fama (1968) showed that both, Sharpe (1964) and Lintner (1965) models, were right and that, actually, both were based on the same set of assumptions from Markowitz (1952), having the same conclusions regarding the measurement of risk and its relationship with the asset's expected return. According to Fama (1968), the differences that authors mentioned were caused by each focusing on different stochastic processes. More precisely, the author thinks that both Sharpe (1964) and Linter (1965) misunderstood the implications underlying Sharpe (1964) model and he demonstrates how it is possible to reach Lintner (1965) model and conclusions starting from the former's model, showing that both are approaching the problem of capital asset pricing in a similar way.

With all the confusion sorted out by Fama (1968), the Capital Asset Pricing Model (CAPM) was born from the combined works of Treynor (1961), Sharpe (1964) and Lintner (1965). A model which predicted that the excess return of an asset was proportional to its beta (covariance of its returns with the returns of the market portfolio). In its simplest form, this model establishes a linear relationship between the asset's beta and their excess returns.

Applying these developments, Jensen (1968) evaluated 115 open end mutual funds, in which he measured the forecasting ability of the portfolio manager to correctly predict the future prices of securities. To fulfill his objective, Jensen (1968) compares the risk premium realized by the portfolio with the level of risk that the portfolio's composition suggests it should have. Since the portfolio risk was allowed to vary over time, this comparison regarding manager's ability was decomposed in two components: prediction of market's behavior (which the manager may want to do, in order to outperform it) and selection of individual securities (which, by evaluating the portfolio's returns, also allows us to withdraw conclusions regarding the manager's prowess).

In order to properly measure it, and consider the effects of risk on return, he used a non-zero constant in the model, denominated α, that would have a positive value if the manager had a good forecasting ability (i.e. the risk premium attained by the portfolio was higher than the expected risk premium, for a certain level of risk), would be null if the manager is doing as well as a random selection and negative if the manager were to perform worse than the random selection. However, the author calls our attention for the latter situation, which should be taken with caution, since it is not a common outcome and may have underlying factors that are worth being analyzed before coming up with a conclusion.

The evidence found by Jensen (1968) points out that, on average, the chosen funds, and consequently their managers, were not able to forecast prices in such a way that would allow them to beat the market. Actually, the analysis suggests that their performance was not enough to cover their brokerage expenses, even when all other expenses are not considered. Notwithstanding, the author mentions that these results should be taken with caution since the dimension that concerns diversification was not considered.

Subsequently, Jensen (1969) complemented his work by performing a more in-depth analysis of the same set of funds, adding criteria for measuring the efficiency of a portfolio as well, while also showing that this and the performance's measures were connected. To properly address efficiency, the model developed allowed the testing of the semi- strong form of the efficient market hypothesis, in the sense that current prices already have reflected on them all the information available.

Starting with the relationship between performance and efficiency, Jensen (1969) considers three different situations: the scenario of a perfectly diversified portfolio, where both measures of performance and efficiency overlap (not forgetting the fact that even a diversified portfolio may become inefficient due to the actions of the manager), a scenario with a bad-performing portfolio, implying that it is also inefficient, and a scenario of a good-performing portfolio, which brings us to an ambiguous position where a portfolio will be efficient if the forecasting ability displayed by the manager is enough to overcome possible inefficiencies present in it. Nevertheless, these definitions should not be considered lightly since an investor may combine investments and, under those circumstances, a good-performing portfolio, even when not efficient, may be useful.

Jensen (1969) had similar conclusions to his previous work, in Jensen (1968), regarding the performance of funds, with the addition that their behavior resembled the expected from the semi-strong form of market efficiency hypothesis, although funds were, on average, inefficient. This implies that the manager's attempts to analyze the current and past prices has not yielded higher returns. However, in contrast with his previous findings, the author states that, if the funds' expenses are added back to them, they would reveal a neutral performance instead of underperformance. Nonetheless, Jensen (1969) work favors the maintenance of a perfectly diversified portfolio.

As seen above, the CAPM was practical and was used in studies regarding portfolio performance evaluation. However, some authors took a different approach on the CAPM and decided to relax or extend some of its assumptions.

On this matter, Black et al. (1972) performs additional tests of the model in order to contribute to a better understanding of security returns. With that in mind, the authors decide to estimate the value of an average intercept ($\hat{\alpha}$) which, if the model was correctly pricing the assets, should have an outcome that was not significantly different from zero.

The first results showed a negative $\hat{\alpha}$ for the portfolios with higher risk and a positive $\hat{\alpha}$ for the low-risk portfolios which means that, over the analysis period, securities with higher risk, on average, earned less than what the CAPM would predict while securities with lower risk, on average, earned more than what the CAPM would predict. After this first test, the authors used a two-factor model that, given certain circumstances, would yield the traditional model (CAPM), to perform cross-sectional tests on portfolios again.

In their model, the return of a security was a linear function of two factors: the market portfolio and a zero-beta portfolio (i.e. its covariance with the market portfolio is null) with minimum variance. Both factors have coefficients, but only the second factor has its coefficient depending on the security's beta, earning it the name of beta factor. Under these circumstances, the CAPM would be consistent with the latter factor having a zero mean on its excess returns (meaning it would be equal to the risk-free rate).

The outcome of the empirical tests revealed strong evidence in favor of rejecting the CAPM, since both, the excess returns and the factor itself, were significantly different from zero and economically significant, respectively. Nevertheless, the authors call our attention to the fact that, while they have asserted the significance of this factor, they have not proved its existence, since a direct test for it was not performed. They have, however, provided the proper economic rationale for its finding to be consistent with capital market equilibrium, suggesting the existence of other factors, besides the market, that systematically affects assets' returns.

On the other hand, Merton (1973) opts to develop a model that, whilst respecting the simplicity and practicality of the CAPM and being consistent with its assumptions, would provide a more accurate description of the relationship among yields, in line with the empirical data. The author decides to extend the model, by actually making it intertemporal, arguing that this transformation would increase the model's ability to capture effects that, otherwise, wouldn't be noticed on the single-period CAPM.

This extension brings extra considerations, especially when selecting the portfolio, since the investor now has to take into account the relationship between current and future returns. Given this, the demand for assets will have a new component showing the investor's desire for a certain asset in order to hedge against possible harmful changes in the investment opportunities. In fact, Merton (1973) states that, due to the interest rate changes over time, we cannot assume a constant set of opportunities for investment and should, instead, incorporate such changes in our analysis and interpret the effects that they cause.

By doing that, the author presents a situation where the investment decision is characterized by holdings that would provide both an efficient combination as well as the hedge already mentioned above. This will require a higher expected return for the incurrence of risk that is composed by the market risk and the risk of possible harmful changes in the investment opportunities set. From here, we can see that a security whose beta equals zero does not have an expected return that equals the risk-free rate, differing from the CAPM and making Merton (1973) remark that his model seemed to better suit the data. The author even suggested his and Black et. al (1972) models could have potential for agreement, although his analysis was very simple and did not provide any solid conclusion.

Notwithstanding, the empirical tests of Black et. al (1972) and the extension of Merton (1973) have showed that there is a relationship between risk and average returns. Fama & MacBeth (1973) decided to test it and considered three hypotheses: linear relationship between the expected return in a security and its risk, beta being the only measure of risk of a security and, assuming that there are risk-averse investors in the market, higher risk should mean higher returns.

The authors consider that the linear relationship hypothesis has not been evaluated as it should, given its importance. According to them, this relationship would exist if the process of price formation in the capital market showed the investor's desire to hold efficient portfolios. As for the second one, if beta is not the only measure of risk, then there would be premiums that were being paid for risks that did not contribute for the overall risk of an efficient portfolio, making the assumption that investors want to hold them fail. Last but not least, the remaining hypothesis was seen as a connector between the previous two and was expected to reveal the existence, on average, of a positive tradeoff between risk and return.

The outcome revealed that it was not possible to reject the hypothesis of risk-averse investors wanting to hold efficient portfolios and that there was, on average, a positive tradeoff between return and risk, measuring the latter from the portfolio perspective. Having said that, it implies that investors should consider a linear relationship between a security's expected return and its portfolio risk, expecting a positive tradeoff between these two parameters and only needing the security's beta to measure risk. Furthermore, since the hypotheses seemed to be corroborated, they can be seen as supportive of the existence of an efficient capital market.

The relationship between risk and return was also presented, although in a different perspective, in Ross (1976). In here, the author intended to focus on an alternative model to CAPM because, even though he recognized that the tractability and linearity between expected return and risk had conferred popularity to it, he points out the difficulty of theoretically justifying certain assumptions that it makes, such as the normality in returns.

Given that, the author starts by mentioning a condition, which is the arbitrage alternative to CAPM, that is expected to hold in both equilibrium and disequilibrium situations, with the distinguishing characteristic of the market portfolio not playing a special role. Developing this condition, Ross (1976) introduces a model where the expected return of an asset is dependent on different risk factors and their load (i.e. the sensitivity of the portfolio to changes in those risk factors, measured by betas). Additionally, such a model also makes the risk premium of an asset to be the sum of the risk factors premiums, weighted by their respective betas.

Having said that, Ross (1976) proposal is more closely connected to an arbitrage relation than an equilibrium relation and is expected to be robust. However, even though it does not require homogeneous investors' anticipations, it requires that they share the same beliefs regarding the load of each factor (beta). Furthermore, it is still exposed to the possibility of the agents disagreeing on the nature, and the probability distributions, of the risk factors since they are not easily identifiable.

Nevertheless, Ross (1976) model brought yet another variant for academic empirical studies that were evaluating portfolio performance. Among them, Jensen's α was one of the most widely used metrics to measure performance. However, it wasn't free from controversy as documented in Grinblatt & Titman (1989), which makes a simple, yet insightful, reunion of the three main critics attached to these kind of measures: lack of a concrete and adequate market portfolio, overestimation of risk due to biased estimators and negative risk-adjusted returns caused by risk aversion levels that are higher than perceived.

In their work, the authors reach some conclusions regarding such criticism. More precisely, they justify the existence of a market portfolio by its mathematical, rather than equilibrium, properties, propose a new and broader measure, from which Jensen's α is a particular case, that avoids the overestimation of risk and, ultimately, give a possible explanation for the risk aversion levels which they consider to be dependent on the information available to the investor at the time of decision, as well as whether it is already priced or not.

But the performance measures are not the only ones targeted by critics, and Grinblatt & Titman (1989) also highlighted one of the main critics pointed at CAPM's empirical studies based on observable data that shows that, on average, smaller firms outperformed larger firms by a significant amount, as stated in Banz (1981). Nevertheless, Banz (1981) also warns that such results should be taken carefully, since it is not proven that the size effect, as he called the outcome, exists. In fact, there is the possibility that it is just a proxy for other factors that are not being considered and are correlated with size.

Notwithstanding, these last findings suggest the need to consider additional variables to explain a stock's expected return. Several studies addressed this issue, like Bhandari (1988) who focused on an intuitive approach to the risk of common equity in a firm, by considering debt/equity ratio (DER) as a proxy for it. Bhandari (1988) suggests that, even with beta estimates being calculated with a proper method, as in Black et al. (1972), and the existence of a size effect, pointed by Banz (1981), DER still has a role in explaining a stock's expected return.

DER was found to have a positive, and significant, impact on the expected return of a stock, while controlling for both, beta and the firm size. Furthermore, these results were not dependent on the chosen market proxy and suggested that observed premiums could be caused by other factors besides risk, supporting findings from previous studies.

These new findings put in perspective the expected positive simple relation between the average stock returns and market beta, supported by the works of Black et. al (1972) and Fama & Macbeth (1973) and motivated the work of Fama & French (1992) where the impact of a combination of factors (market beta, size, earnings-to-price ratio, leverage and book-to-market equity) in the expected stock returns is studied.

Since all the variables mentioned, despite being different, are connected to price, it was expected that there would be some redundancy. The results showed just that, with the combination of only two factors (size and book-to-market equity, the latter having more significance) representing the impact of the remaining (leverage and earnings-to-price ratio). Furthermore, they also suggested that beta does not seem to contribute in explaining the expected stock returns, supported by the fact that it had no explanatory power when used alone in the tests made.

These outcomes sustained a rational framework for asset-pricing, which was extended by the same authors in Fama & French (1993). In here, they decided to test whether their model would be appropriate to explain average returns in both stocks and bonds, while adding two more factors related to the term structure and default risk to better evaluate the latter. Additionally, they interchanged the factors between both type of securities to check whether there was an overlap between them.

The authors built six types of portfolios from the combinations of size and book-to-market values, allowing them to calculate the difference in returns between small and big firms (SMB), between high and low book-to-market companies (HML) and the excess market return (MKT). These were used as factors and, along with the previously mentioned bond-factors, tests were performed to evaluate the fit and significance of the five factors (MKT, SMB, HML and the two bond factors) and the three factors (MKT, SMB, HML).

The results revealed that the three-factor model seemed to work well when the portfolio evaluated was only composed by stocks, making the five factors only required when there was a mix of stocks and bonds. Even though there was some theoretical background for the inclusion of the bond factors when performing regressions on average stock returns, it was revealed that their effects were well captured by the MKT factor, hence supporting the choice of the three-factor model to explain average stock returns. Despite, in their previous work, having shown that the market factor did not have explanatory power when used alone, the authors find that there is the need for its inclusion in this model due to the fact that SMB and HML explain the differences in average returns across stocks, but not the differences on average returns between stocks and the risk-free asset, leaving this to be explained by the market factor.

Notwithstanding, this three-factor model was extended to a four-factor model by Carhart (1997), where the author added one momentum factor based on the work of Jegadeesh & Titman (1993). For Carhart (1997), the motivation for this extension came from the fact that the three-factor model did not seem to provide a proper explanation of cross-section variation when considering portfolios sorted by their momentum.

Regarding the new factor, Jegadeesh & Titman (1993) considered it by taking into account the behavior of individuals, which tend to over or underreact to information. Given that, it is possible to consider profitable trading strategies by selecting stocks, using their past returns as a criterion. Such a strategy is exactly what the authors proposed, holding portfolios that focused on having a long position in the past winners and a short position in the past losers, while rebalancing them accordingly.

For the period they analyzed, they obtained abnormal returns and evidence that these types of performances were not guaranteed for the long run, giving support to the conclusion that the impact of over or underreaction to information is not as simple as it was thought.

Nevertheless, these results contrasted with the findings of Carhart (1997), which suggested that the momentum trading strategy was not profitable at the individual security level. On the other hand, the author does find that funds with persistent poor performance should be avoided while being cautious about good performing funds, since their performance is not guaranteed for the long run, in line with Jegadeesh & Titman (1993). Last but not least, these performances are dependent on the investment costs, which have a direct and negative impact on it.

Furthermore, Carhart (1997) also evaluates the improvement in analysis by comparing the outcomes of CAPM, the three-factor model and his own four-factor model. From the first to the last, there is an improvement on pricing errors, supporting the better fit of the four-factor model for explaining cross-section variation on average stock returns. Nonetheless, Fama & French (2015) proposed a five-factor model, although not extending the four-factor model proposed by Carhart (1997), since their tests to the momentum factor revealed that its impact in portfolio performance was not significant.

Their focus went to the evidence of the variation in average returns, connected with profitability and investment, that the three-factor model failed to explain, and decided to add them to their previous model. To take them into account, Fama & French (2015) introduced RMW and CMA as the difference between the returns of stocks with high and low profitability and returns of stocks with low and high investment profiles, respectively. Furthermore, they use combinations between the factors to produce different three and four-factor models and compare the results with the five-factor model initially proposed, finding that the latter suits the data better than the three-factor model that they had proposed before.

The authors also found that one of the factors, related with book-to-market equity, seemed to be redundant, since it did not improve the results from the four-factor model formed with the other factors. Notwithstanding, if the objective of the analysis is not only on the abnormal returns but also in the impact of each factor, then the five-factor model remains the best option.

Despite the five-factor model fitting better the data than the three-factor model, there was still a type of portfolio, composed by small size stocks with high investment profile and low profitability, whose average returns are not properly explained by the model. This outcome was also found in Fama & French (2017), where the authors apply the five-factor model on a global and regional level across North America, Europe, Asia Pacific and Japan and compare it, as before, with the three-factor model.

In North America, the five factors were considered important, in contrast with Europe, where only four factors, excluding investment, are significant for explaining the average returns. Turning to Japan, book-to-market and profitability share the spotlight, although the former has a greater impact, while Asia Pacific has a similar outcome to Europe, although the investment factor still plays a role, even though not as big as the remaining four factors.

In the end, the five-factor model revealed itself as an improvement over the three-factor model, even when applying it internationally, as long as the factors are calculated taking into account regional characteristics.

2.2. Conditional Performance

As we can see, CAPM has not been a model free from controversy, with attempts at improving, or replacing it, happening throughout its history. Having said that, there were some authors who preferred to take a different approach than the previous ones and, instead of accepting a static reality as an assumption, considered the possibilities of the dynamic patterns that characterize our economy.

Among them, were Jagannathan & Wang (1996) who looked at results as the ones presented in Banz (1981) and Bhandari (1988) as scenarios where it is possible to construct portfolios which CAPM cannot explain properly. These also highlight the absence of time variation in betas which, according to Ferson & Harvey (1991), should be considered due to the fact that time variation in the expected risk premium is one of the main contributors to explain portfolio returns.

In fact, Jagannathan & Wang (1996) don't agree with CAPM's assumption of betas that remain constant over time, since the firm's risk may vary over the business cycle and, taking into account that the economy is dynamic, the relative shares of certain sectors may change over time, causing changes in the betas of firms that are present on those sectors. Therefore the authors consider that both, betas and expected returns, will be dependent on the amount and nature of information that is available at each moment in time, which leads to their evaluation of a conditional version of CAPM where an asset's expected return is linearly related to its conditional beta, making it dependent on the information available.

Notwithstanding, it is also possible to idealize an unconditional model where the unconditional expected return is linear in a market beta and a premium beta (measure of beta variation over time) but, despite such a model containing the static CAPM, facilitating direct comparison between them, there are some considerable differences in the way certain variables were built, namely the return on aggregate wealth. In here, the authors decided to introduce a measure of return on human capital, trying to improve the explanatory power of the model and avoiding interpretations where the market proxy selected is the main responsible for the undesirable outcome of CAPM, as it is possible in Fama & French (1992).

The outcome revealed that the conditional CAPM was significantly better than the static CAPM, being corroborated by the fact that size and book-to-market variables did not have as much impact on determining the part that the model could not explain, unlike the static version of CAPM, with examples in Banz (1981) and Fama & French (1992).

Nonetheless, this result may be biased due to each model's assumptions. Based on that, Wang (2003) tested the conditional versions of CAPM, Fama & French (1993) three-factor model and Jagannathan & Wang (1996) conditional model, while applying a methodology that did not require the formal specification of as many parameters, without losing relevance or validity.

Among some of the outcome achieved, Wang (2003) finds that although size and book-to-market have an impact on explaining average returns, they reveal it through different patterns. Furthermore, the author also tested Jegadeesh & Titman (1993) momentum strategy, which the static CAPM failed to explain, and the results of applying the conditional version of the three-factor model were not rejected, suggesting that winners are inclined to have higher conditional expected returns.

Given this, Wang (2003) shows the importance of measuring the time-varying returns properly, which was in line with the work developed by Ferson & Harvey (1999), who were aware of the existing concerns about the factors used in the three factor model, since they were not derived from a theoretical background. Nevertheless, the authors, building upon their previous work, focused on whether the economy dynamic patterns were reflected in the model while knowing that, considering only the three factors presented by Fama & French (1993), it is expected that the patterns mentioned above will produce biased outcomes.

After conducting empirical tests, while using a set of economy-wide variables, Ferson & Harvey (1999) concluded that controlling for time-variation improved the model, which had implications to the evaluation of asset prices, since the results revealed that there was information not being captured by the most widely used models up until then.

On that subject, Ferson & Schadt (1996) had already shown that, given the time variation on expected returns and risk, it was likely that models with unconditional approaches would not be reliable. To counter this problem, the authors analyze the influence on investment performance of including lagged information variables on their tests, hoping to reduce the bias caused by public information.

Such a conditional approach was seen as advantageous due to the possibility of controlling for the dynamic patterns that are present in the economy and influence returns, unlike the traditional, and unconditional, approach. Furthermore, dynamism can also be introduced by managers' own behavior on selecting investment opportunities and, controlling for that, is also seen as an important advantage.

Since Ferson & Schadt (1996) analyzed mutual fund performance, they based their work on Jensen (1968) model, although modifying it in order to make it conditional on the information available. Such modification neutralized previously considered poor performance by Jensen and suggested that mutual funds' betas are correlated with public information, showing that variation on their risk exposure would be dependent on it. Such evidence, once more, gave support to the statistical and economical significance of conditioning returns on information.

These outcomes, together with the works of Jagannathan & Wang (1996) and Wang (2003), were clear examples of the use of conditional models to evaluate asset prices. Nevertheless, these models weren't bulletproof and their construction, mainly in the way the data was handled, was targeted by the work of Ferson et al. (2008).

In it, the authors intended to study the impact of data mining and spurious regression. The former refers to the utilization of the same datasets by many researchers, while attempting to estimate their models. The later comes from the existence of high autocorrelation from a variable which, if not taken into account, may wrongly lead to a conclusion of a statistically significant relation. Furthermore, such problems are strongly related and, when both are present, tend to magnify the impact of each other. These issues are important due to the mainstream use of lagged variables and datasets in conditional modelling that may fall prey to the problems mentioned above.

Their analysis suggested different outcomes for each conditional estimator. When it comes to betas, they seem to be relatively robust to the effects of either data mining or spurious regression. However, in order to properly take into account time variation, as Ferson & Harvey (1991) and Ferson & Harvey (1999) recommended, and get a consistent estimator, a linear term in the lagged variable must be included. On the other hand, while alphas reveal some robustness to the issues already mentioned, their time variation estimator seems to have a bias. Additionally, the authors also find that its absence from a model will cause the conditional estimator for beta to be biased.

Notwithstanding, the authors also test a regression where time variation in betas is suspended, similar to the work of Jagannathan & Wang (1996) who considered time variation but not in an explicit manner, in order to observe the impact on alpha. The evidence gathered shows that the importance given in literature to time variation in alphas may be excessive, leading the authors to propose that the current conditional asset pricing models can be a better fit than what it may be suggested from certain studies on them.

In fact, these models also require a different approach to achieve an optimal solution for portfolio construction since the presence of conditional information will make such a solution to be dependent on it, taking into account that it provides the probability distribution of future outcomes.

Ferson & Siegel (2001) addressed this issue by proposing solutions who are nonlinear functions of the conditional parameters of returns that are important for the formulation of the optimal portfolio: mean and variance. Their objective was to promote unconditional efficiency, due to the existence of information asymmetry. This problem can be exemplified by thinking about a manager and an investor. The former can use information about future returns when constructing the portfolio but, if the latter does not have access to that information, then the portfolio will not seem efficient.

The unconditionally efficient portfolio weights are very similar to other outcomes in traditional models such as CAPM, maximizing the utility achieved. However, such a realization only happens for central levels of the probability distribution that conditional information presents. For extreme realizations, there is the need for a more thoughtful approach, since very high expected returns are an opportunity to reduce risk without compromising the portfolio overall performance.

To complement their work, the authors provide solutions for unconditionally efficient portfolios in three different situations, with different combinations of one risky asset, one riskless asset and *n* risky assets present in the economy. In the end, their results prove to be robust and supports the importance of proper conditional modelling regarding asset prices, reflecting the dynamic behavior of the economy.

2.3. Board Composition and Firm's Performance

Up until now, we have studied the performance of companies without considering their internal factors, more specifically, their board composition. Several authors have addressed this subject and its impact. Adams & Ferreira (2007) focused on the advisory role that boards may play, besides their usual monitoring of CEO's activity.

This is due to the fact that the manager's decisions are not always the ones that maximize shareholder value hence requiring an active participation from the board on the firm's activity and decisions. Nonetheless, the authors argue that the board may use their members' expertise to improve their overall decision-making processes, and it will improve as much as the board's preferences are in line with the manager's preferences.

However, for the latter to happen, the board needs that the CEO properly discloses information, which constitutes a trade-off for the manager. On one hand, the better the disclosure, the better the advice. On the other hand, this also gives the board higher leverage to intervene in future CEO's decisions. Having said that, the authors conclude that the manager will not fully provide information about the company to a board that is highly independent.

Nevertheless, Adams & Ferreira (2007) also find that, when there are two entities performing the two roles (monitoring and advisory), such a trade-off no longer exists, which provides some support for the implementation of dual board systems. Notwithstanding, the authors show that, unless the benefits of control for the managers are too large, shareholders will be better off implementing a sole board system, although its effectiveness will be higher if the advisory role is also performed.

However, even the monitoring role isn't as straightforward and, as Campbell & Mínguez-Vera (2008) show, it has received higher attention in recent years with cases like the failure of Enron supporting the importance of this role. Among the factors that may influence its effectiveness, gender is the one addressed by these authors, who wanted to evaluate whether the presence of women on company boards would improve the firm's performance, taking into account the Spanish economy, since it was an European economy with civil law, contrasting with other existing studies with common law.

On the one hand, boards with women may increase shareholder value if they bring new ideas to improve decision-making within the company. On the other hand, their impact may be undesirable if the selection of women to be part of the board is based on pressure to have equality between sexes, instead of their qualifications. In a way, arguments around women's share in boards can be divided between ethical and economic. The first refers to the immorality of excluding women from firm's boards and supports the achievement of a more equitable representation, while the second is focused on the proper selection of candidates to the board of directors in order to maximize firm's performance which, as mentioned above, may or may not be a woman.

The authors consider diverse explanations for the impact of gender on firm's performance, from the greater diversity leading to a competitive advantage due to the company's increasing ability of penetrating markets, to the full use of the diverse characteristics that are distributed among demographics and are influenced by gender. Nevertheless, their findings suggest that the presence of women on the board does not have an impact on firm's performance. Notwithstanding, they also find that the balance between men and women on board has a positive impact in performance, which leads them to conclude that the presence of women will not destroy value to shareholders.

In line with these conclusions, we have seen market regulators putting gender diversity on the topics to debate, surrounding corporate governance. However, as Chapple & Humphrey (2014) report, the approaches of each regulator may differ, with some imposing quotas and others only recommending or requiring higher disclosure regarding firm's gender composition. With this in mind, the authors investigated the impact of gender diversity in the financial performance, at an aggregate level, building portfolios of firms with gender diversity and comparing their returns with firms without any women on the board.

The outcome of their analysis corroborated the findings of Campbell & Mínguez-Vera (2008), having no suggestion of an impact in financial performance from the presence of women on boards and supporting the difficulty that seems to exist in finding economic arguments in favor of gender diversity, while using market data. Nevertheless, there seems to be a justification for large firms to increase their diversity in order to promote their image before society, although that does not imply that they are looking for quantified economic and financial improvements.

Nonetheless, Chapple & Humphrey (2014) also highlight that their analysis of the Australian market had a constrain that could hinder the validity of the conclusion. More specifically, they found that the percentage of boards with two women on them was around 10% and, above that level (three or more) the sample considered had even lower percentages. In a way, this shows why the Australian market regulator decided to address this issue but also shows that it may not be possible to properly evaluate the impact of gender diversity on performance, since there may be required a large number of women on boards in order to have a visible impact.

At an European level, this issue was also addressed, with Ahern & Dittmar (2012) reporting a new law, written in 2003, which imposed quotas on Norway. This situation was seen by the authors as a good event to study the connection between firm value and board features, since there were changes in the board composition due to an exogenous factor. Their interest in it also came from the possibility of improving our understanding of such a connection since it was not easy to know whether the existence of experienced board members increased firm's value or if the high firm value was responsible for the addition of more experienced members to their board, which was a similar problem to the one described by Campbell & Mínguez-Vera (2008), regarding the selection of women to the board.

In Norway's case, all public limited firms were forced to comply with the quota of at least 40% female directors and, obviously, the firms that already had a big percentage of women on their boards faced less constraints and more easily reached the objective. Nevertheless, there was a negative impact on firm's value across all the listed firms analyzed and, while the firms mentioned above had a lower impact, it showed that the huge reorganization imposed by the law affected shareholders' returns.

Boards are built to maximize those returns and, according to Ahern & Dittmar (2012), each member's characteristics were important to determine their capacity to perform the roles suggested in Adams & Ferreira (2007): monitoring and advisory. The authors find that female directors are substantially different than male directors and that it is reflected in the firm's behavior afterwards. Notwithstanding, firms reacted differently to the imposed quota, with some maintaining board size and reshuffling their members, and others avoiding the quota by relocating or changing their status (i.e. from public to private).

Norway wasn't the only case where initiatives and policies were taken, in order to promote gender diversity. However, such measures cannot be based on improvements to average firm performance, as Adams & Ferreira (2009) show. In their work, the authors try to answer a group of questions regarding women's role on boards and their impact on its roles and, ultimately, on the firm's performance.

Furthermore, Adams & Ferreira (2009) were cautious regarding one of the main problems on these types of analysis: endogeneity. Such a problem may bias the data estimations or predictions, not allowing for a casual interpretation to be given. With this sorted out, the analysis would be able to provide reliable information about the effectiveness of boards and, consequently, the plausibility, or not, of success for the policies mentioned above.

As Ahern & Dittmar (2012), Adams & Ferreira (2009) find that women on boards behave very differently from men, leading to repercussions on the firm's behavior. Also, they corroborate the argument they presented before in Adams & Ferreira (2007) regarding the negative impact on shareholder value that an excess of monitoring may have. Since the study finds that women are more prone to have monitoring functions in the board, such a finding may be a possible explanation for the value destruction in firms with higher governance structures that start to diversify their board's gender composition.

Nonetheless, this work highlights the complexity of the relationship between gender diversity and firm's performance, in contrast with the more linear appearance that some press may have given it. To further understand the issue, Gul et al. (2011) decided to check whether the role of female directors would have an impact on stock prices, due to a possible increase in disclosure of information and higher incentive to gather private information.

More precisely, the authors argue that the former comes from a higher transparency of the CEOs due to an increase in the monitoring of their activities. Function which the women are more prone to have, according to Adams & Ferreira (2009). Regarding the latter, women introduce changes in the usual way the board makes its deliberations since their behavior is different from men, as stated by Ahern & Dittmar (2012) and Adams & Ferreira (2009), making board members to have higher consideration of the consequences that derive from their decisions.

The authors find that there is a positive relation between gender diversity and stock price informativeness, while taking into account the endogeneity problem already mentioned in Adams & Ferreira (2009) and controlling for other variables like governance and earnings quality. Furthermore, they also state that the public disclosure effect is more likely in larger firms, while the higher procurement of private information is more easily found in small firms. The first is directly reflected in the stock price and the second is seen through the activity in trading, due to the more informed investors.

In the end, these findings show the association between earnings quality and trading activity with stock price informativeness, supporting the opinion of the authors that regulation being applied through different approaches, like the ones mentioned in Chapple & Humphrey (2014), help the advantages of gender diversity to be realized in a faster way.

Notwithstanding, theses outcomes were not particular of developed markets and Nguyen et al. (2015) showed exactly that. Since Adams & Ferreira (2009) stated that the impact of a gender diverse board would depend on the structure of corporate governance previously present in the firm, the authors thought that the country they selected for their study, Vietnam, would benefit highly due to the overall poor corporate governance present on it.

Despite the usual structure in Vietnam being composed by two boards (one of directors and another of supervisors), similar to what Adams & Ferreira (2007) study, the functions of monitoring and advisory are concentrated on the first, while the second is only concerned with evaluating the annual financial statements and the performance of both, the board of directors and the CEO. Nevertheless, its role is quite inefficient since it does not have any clear guidelines on how to perform it, leaving each company to decide its approach.

In the end, the authors gave support to the works of Adams & Ferreira (2009) and Gul et al. (2011) by showing that, after controlling for several corporate governance factors, gender diversity seems to have a positive influence on firm's performance. Such an outcome would be possible by the greater benefit of adding women to the board since, as mentioned above, the corporate governance structure in Vietnam was underdeveloped, and the increase on firm's transparency would be reflected on both trading activity and stock price.

Chapter 3 - Methodology and Data

3.1. Research Method

This work has the objective of evaluating whether building a portfolio with companies, while taking into account their Gender Diversity Index (GDI) score, will impact its performance.

With that in mind, two types of portfolio will be built based on the level of gender diversity present in the companies. In order to evaluate their performance, while on an unconditional level, the models used will range from the three-factor model proposed in Fama & French (1993) to the four-factor model of Carhart (1997), using the previous model and adding the momentum factor from Jegadeesh & Titman (1993), and the five-factor model of Fama & French (2015), while taking into account (Jensen, 1968) alpha.

Having said that, we will estimate, for each type of portfolio, nine linear regressions, making use, in all of them, of Stoxx 600 and MSCI World as market proxies, to compare regressions. The impact on performance caused by gender diversity, will be observed on the monthly returns of the portfolios and, in the end, the objective is to verify which model better explains such performance.

3.2. Portfolio Composition

We based our selection on EWOB (2019) from the organization *European Women on Boards (EWOB)*, where they gathered information from firms in the Stoxx 600 and made an hierarchy based on their boards' gender composition and women's roles in the company. Such analysis was compiled into a Gender Diversity Index (where a score of 1 means a board with 50/50 composition, and values below or above 1 reveal a domination of males or females, respectively).

Afterwards, *EWOB* organized a Top 20 and a Bottom 20 based on the GDI score of the companies, being the former composed by firms with a GDI score at or above 0.8 and the latter composed by firms with a GDI score at or below 0.2. Our two portfolios were built to mimic the Top 20 and Bottom 20 tables, and we collected data on each firm using Datastream, for the period between 2000 and 2019, with a monthly tenor.

We gathered information on the MV (Market Value) of each company and used Datastream's datatype RI (Return Index), which includes the reinvestment of dividends, to allow us to properly calculate the monthly returns of the firms. Both of these indicators were adapted to be displayed in EUR (Euro) since, although all of the companies were part of *Stoxx 600*, there were some who belonged to countries with a currency other than Euro.

Furthermore, we also found the need to withdraw information on Stoxx 600 and MSCI World indexes, using the datatype RI for the first and MSRI for the latter, which did the same as RI but was specific for this index. The reason behind this was to allow us to compare regressions using the models mentioned above, replacing the market factors used in the regressions, with each index.

3.3. Models to estimate

As mentioned above, we will estimate regressions based on different models in the literature. Our main point is to evaluate if there are abnormal returns on our portfolios, using alpha to measure it, as first employed by Jensen (1968), and see whether our portfolios had an outperformance (positive and significant alpha), underperformance (negative and significant alpha) or neutral performance (alpha not significantly different from zero).

Our first model will be Fama & French (1993) three-factor model:

$$R_{it} - R_{ft} = \alpha_{it} + \beta_i \cdot [R_{mt} - R_{ft}] + s_i \cdot SMB_t + h_i \cdot HML_t + \varepsilon_{it}$$

Where R_{it} represents portfolio i return, on period t, R_{ft} represents the risk-free asset's return, on period t, α_{it} represents the abnormal return of portfolio i, in period t, β_i represents portfolio i sensibility to general market variations (i.e. the systematic risk), R_{mt} represents the return of the market portfolio, on period t, SMB_t (Small minus Big) represents the difference in return between a diversified portfolio of small and large cap stocks, in period t, HML_t (High minus Low) represents the difference in return between a diversified portfolio of high and low book-to-market companies, in period t and ε_{it} represents the error term of portfolio i, in period t.

When we consider Carhart (1997) model:

$$R_{it} - R_{ft} = \alpha_{it} + \beta_i \cdot [R_{mt} - R_{ft}] + s_i \cdot SMB_t + h_i \cdot HML_t + w_i \cdot WML_t + \varepsilon_{it}$$

WML_t (Winners minus Losers) represents the difference between diversified portfolios of stocks that performed well and bad in the recent past, in period t. The remaining variables were already explained above.

On unconditional performance, the last model is Fama & French (2015) five-factor model, which we can state as follows:

$$R_{it} - R_{ft} = \alpha_{it} + \beta_i \cdot [R_{mt} - R_{ft}] + s_i \cdot SMB_t + h_i \cdot HML_t + r_i \cdot RMW_t + c_i \cdot CMA_t + \varepsilon_{it}$$

Where, RMW_t (Robust minus Weak) represents the difference between a diversified portfolio of stocks with robust and weak profitability, in period t and CMA_t (Conservative minus Aggressive) represents the difference between a diversified portfolio of stocks with a conservative and an aggressive investment profile, in period t. The remaining variables were already explained in the previous models.

3.4. Data

In order to estimate these models, we decided to build two portfolios using 40 companies present in the Stoxx 600 index, gathering data for the period between January 2000 and December 2019. Each portfolio had 20 companies and they were divided using the Gender Diversity Index (GDI) provided by the organization *EWOB* (*European Women On Boards*), who created this score for all the companies present in the index. Therefore, the 20 that had the highest GDI were joined to form one portfolio (Top 20) while the 20 that had the lowest GDI formed the other one (Bottom 20).

Additionally, we used Thomson-Reuters' Datastream Database to obtain data for the period mentioned above, on a monthly basis, for their share price and their market value (MV), in order to calculate their monthly returns and adequately create the equal-weighted and the value-weighted portfolios.

The index where such companies are present, Stoxx 600, is a dynamic index, i.e. its composition may change along the time, according to the way its components behave and are impacted by the market conditions. Furthermore, it is an index that focuses on European countries and has companies with diverse characteristics, from large to mid and small capitalization. This index data, along with the MSCI World Index, which is an index that covers large and mid-cap firms in all developed countries, were also obtained from Thomson-Reuters' Datastream Database, on a monthly basis.

To complement our analysis, and apply the models required, we needed to gather data on the factors that help explaining the difference on our portfolios' returns. In order to achieve that, we used the data present on Ken French's website, while taking into account that all our companies are European.

Lastly, we built the portfolios, as mentioned above, and then created two subdivisions on each of them: one equally weighted and another that based each company's impact on its percentage of the total MV of the portfolio (value-weighted).

Table 1 - Descriptive statistics for the monthly returns of Top 20 and Bottom 20 Portfolios equally weighted (EW)and value weighted (VW), from a total of 239 observations.

Portfolio	Mean	Median	Standard Deviation	Min	Max
Top 20 (VW)	0.0116	0.0132	0.0409	-0.1190	0.1420
Top 20 (EW)	0.0118	0.0148	0.0467	-0.1490	0.2060
Bot 20 (VW)	0.0147	0.0191	0.0568	-0.2580	0.2410
Bot 20 (EW)	0.0162	0.0216	0.0510	-0.1690	0.2110

Table 1 shows the descriptive statistics for the portfolios we built. At a first glance, we can see that all portfolios had a positive average return, which is also supported by their median showing they kept that level, at least 50% of the horizon time.

Nevertheless, it is important to remark that the equally weighted portfolio with the Bottom 20 companies (Bot 20 EW) had a median above 0.02, which highlights a possible better performance from this portfolio when joined with the fact that it presents the highest average.

Furthermore, we also observe that the amplitude of the portfolios' returns is quite big, having values that range from -0.258 to 0.241, both being achieved on the value-weighted Bottom 20 (Bot 20 VW) portfolio.

All in all, we can see that the Bottom 20 portfolios performed better, albeit having higher volatility, showed on their higher standard deviation as well as on the difference between their minimum and maximum values attained.

Table 2 - Descriptive Statistics of the variables used in our analysis, from a total of 239 observations;

** Stoxx 600 and Exc_Stoxx600 statistics are derived from 228 observations due to the availability data.*

Variable	Mean	Median	Standard Deviation	Min	Max
Exc_RPewTop	0.0104	0.0133	0.0468	-0.1490	0.2060
Exc_RPvwTop	0.0103	0.0109	0.0409	-0.1210	0.1410
Exc_RPewBot	0.0149	0.0216	0.0511	-0.1720	0.2100
Exc_RPvwBot	0.0133	0.0179	0.0569	-0.2600	0.2360
MktRF	0.0042	0.0059	0.0518	-0.2200	0.1370
SMB_1	0.0012	0.0016	0.0201	-0.0689	0.0490
SMB_2	0.0019	0.0022	0.0197	-0.0741	0.0479
HML	0.0042	0.0038	0.0259	-0.0945	0.1120
RMW	0.0031	0.0036	0.0166	-0.0484	0.0608
CMA	0.0032	0.0012	0.0184	-0.0733	0.0875
WML	0.0075	0.0098	0.0427	-0.2610	0.1370
MSCI	0.0046	0.0093	0.0397	-0.1170	0.1150
Exc_MSCI	0.0032	0.0071	0.0399	-0.1180	0.1150
Stoxx600*	0.0043	0.0107	0.0430	-0.1410	0.1450
Exc_Stoxx600*	0.0031	0.0103	0.0431	-0.1420	0.1450

The first four variables: Exc_RPewTop, Exc_RPvwTop, Exc_RPewBot and Exc_RPvwBot represent the excess return of the equally weighted portfolio (in the case of Exc_RPewTop and Exc_RPewBot) over the risk free rate, and the same reasoning applies to the value weighted portfolios (Exc_RPvwTop and Exc_RPvwBot). Additionally, the first two variables represent the portfolios with the Top 20 companies, regarding GDI score, while the next two represent the portfolios with the Bottom 20 companies. Having said that, we can classify them as our explained variables, following the methodology already exposed in the previous chapter, $(R_{it} - R_{ft})$.

The variables MktRF, SMB and HML are present in both Fama & French (1993) and Fama & French (2015), but RMW and CMA are only present in the latter, while WML is a variable introduced in Carhart (1997). Nevertheless, all these variables were already explained in the previous chapter, therefore, in order to avoid redundancy, we will kindly ask to refer to that chapter for clarification on their meaning.

The bottom four variables: MSCI, Exc_MSCI, Stoxx600 and Exc_Stoxx600 are variables that we decided to use as a proxy for our market variable, in order to compare with the regressions obtained by using the market factor from Fama & French (1993): MktRF. MSCI and Stoxx 600 represent the return from the MSCI World Index and the Stoxx 600 Index, respectively. However, in order to have comparable regressions, we also calculated Exc_MSCI and Exc_Stoxx600, which is the excess return of both indexes over the risk-free rate ($R_{mt} - R_{ft}$).

Since the risk-free rate is a common factor to the first four variables, it is with no surprise that we conclude the same as in the previous section, where we analyzed the returns of each portfolio. The portfolios with the Bottom 20 companies seem to perform better in both situations, equally and value weighted. Nevertheless, it is worth to mention that all portfolios have positive average excess returns, which shows that their performance was superior to the one of the risk-free asset.

The MktRF shows a positive average excess return over the risk-free rate (0.0042) but, when compared with our explained variables, all of them have higher average excess returns, hinting at an average performance not only better than the risk-free asset but also better than the market factor itself.

Regarding the difference in return between a diversified portfolio of small and large cap stocks (SMB_1 and SMB_2)[1], we see that, on average, small cap stocks outperform large cap stocks (0.0012 and 0.0019).Additionally, we see that

[1] SMB_1 referes to the data used for SMB variable on the three and four-factor models, while SMB_2 refers to the data used for SMB variable on the five-factor model.

companies with a high book-to-market ratio have higher returns than the ones with a low book-to-market ratio (0.0042), given by HML. The amplitude of values on the first is rather low, with maximum and minimum never going above the absolute value of 0.1, while on the latter that threshold is surpassed when considering the maximum.

On the WML factor, we can see that, on average, the companies with a better past performance have a tendency to have higher returns than the ones with a worse past performance, presenting a difference that is close to 1%, on a monthly basis.

The last two factors: RMW and CMA, reveal that companies with strong profitability have, on average, higher returns than companies with weak profitability, while companies with a conservative investment profile are, on average, better performers than companies with an aggressive investment profile.

Finally, our two variables to replace the MktRF factor and work as a proxy for the market itself: Exc_MSCI and Exc_Stoxx600. These have lower average excess returns, when compared with MktRF (0.0032 and 0.0031, respectively), although very similar among themselves. Nevertheless, they present lower volatility than MktRF while having higher maximum values (0.1150 and 0.1450, respectively).

Chapter 4 – Regression Results

In this chapter we will analyze the outcome provided by the statistical software Gretl, used to estimate the models mentioned in the previous chapter, using the data and the portfolios already described.

Notwithstanding, it stand as important to mention that all the outcome posted here took into account the necessary adjustments to avoid issues related with autocorrelation or heteroskedasticity, in order to improve the value of the results obtained.

Last but not least, we will divide this chapter into subchapters concerning each of the models (three-factor, four-factor and five-factor). In all, both the outcome and the subsequent discussion on the results are going to be present.

4.1. Three-Factor Model

Table 3, 4 and 5 below highlights the outcome of our estimation following the three-factor model proposed in Fama & French (1993), while using alpha (α), as showed in Jensen (1968), to measure the existence and significance of abnormal returns for both portfolios. The main difference among the tables comes from the use of different variables for the market proxy.

Table 3 - OLS estimation of the three-factor model (Monthly data from January 2000 to December 2019);

	Top 20		**Bottom 20**	
	EW	**VW**	**EW**	**VW**
Constant	0.0076*** (0.0021)	0.0086*** (0.0018)	0.0122*** (0.0020)	0.0107*** (0.0028)
MktRF	0.6946*** (0.0481)	0.5731*** (0.0437)	0.7260*** (0.0621)	0.7545*** (0.0827)
SMB	−0.1067 (0.1442)	−0.3243*** (0.1096)	0.3552*** (0.128293)	0.3259 (0.1974)
HML	−0.0108 (0.1217)	−0.0903 (0.1036)	−0.1879 (0.1257)	−0.2254 (0.1703)
R^2	0.5904	0.5439	0.5291	0.4562
F-Test	72.0884	64.7161	47.8412	28.4075

Note: Based on 239 observations. HAC robust standard-errors are inserted between parentheses. *** represents a p-value < 0.01, ** represents a p-value <0.05 and * represents a p-value <0.10

Table 4 - OLS estimation of the three-factor model, with Stoxx 600 as a market proxy (Monthly data from December 2000 to December 2019);

	Top 20		Bottom 20	
	EW	VW	EW	VW
Constant	0.0066*** (0.0011)	0.0075*** (0.0011)	0.0112*** (0.0016)	0.0082*** (0.0017)
Exc_Stoxx600	1.0278*** (0.0367)	0.8543*** (0.0484)	1.0214*** (0.0454)	1.0906*** (0.0682)
SMB	0.2567*** (0.0776)	−0.0234 (0.0966)	0.6726*** (0.1263)	0.6797*** (0.1922)
HML	0.0497 (0.0766)	−0.0706 (0.0673)	−0.0355 (0.1018)	−0.1761 (0.1150)
R^2	0.8432	0.7919	0.7201	0.7176
F-Test	268.0947	108.6819	176.4831	98.11396

Note: Based on 228 observations. HAC robust standard-errors are inserted between parentheses. *** represents a p-value < 0.01, ** represents a p-value <0.05 and * represents a p-value <0.10.

Table 5 - OLS estimation of the three-factor model with MSCI as a market proxy (Monthly data from January 2000 to December 2019)

	Top 20		Bottom 20	
	EW	VW	EW	VW
Constant	0.0055*** (0.0016)	0.0069*** (0.0015)	0.0099*** (0.0019)	0.0083*** (0.0025)
Exc_MSCI	0.9697*** (0.0587)	0.7751*** (0.0624)	1.0238*** (0.0583)	1.0653*** (0.0858)
SMB	0.1871* (0.1014)	−0.0913 (0.0956)	0.6661*** (0.1214)	0.6496*** (0.1832)
HML	0.3665*** (0.0717)	0.2177*** (0.0669)	0.2079** (0.0930)	0.1861 (0.1196)
R^2	0.6758	0.5835	0.6192	0.5362
F-Test	91.0410	59.2567	105.5245	53.2562

Note: Based on 239 observations. HAC robust standard-errors are inserted between parentheses. *** represents a p-value < 0.01, ** represents a p-value <0.05 and * represents a p-value <0.10.

Jensen (1968)'s alpha provides us with the suggestion that all portfolios, on average, have an excess return $(R_{it} - R_{ft})$ which is statistically significant for all levels. Nevertheless, such finding does not help on answering our question since the regressions have different factors.

Regarding the SMB (Small Minus Big) factor we have mixed results. On the Bottom 20 portfolios we can see a positive and statically significant contribution to the excess return on most of them, being the exception the VW (Value-Weighted) portfolio in our first regression (using the market factor provided in Ken French's website as a proxy for the market). When taking into account the Top 20 portfolios, the outcome is not so clear. On the regression with the market factor from Ken French's website, we have a negative and statistically significant impact on the VW portfolio, while the EW does not have a statistically significant contribution. When we consider the Stoxx 600 as our market proxy, we have an opposite outcome, where the EW portfolio has a positive and statistically significant impact while the VW portfolio does not seem to have an impact that is statistically significant. When considering the MSCI index as our market proxy, we obtain a similar outcome to the Stoxx 600 proxy. However, our EW portfolio only has a positive and statistically significant contribution from SMB for the level of significance of 10%.

The HML factor only has statistical significance when used with MSCI as a market proxy, revealing a positive contribution on both Top 20 portfolios. However, the same significance is not reflected on the Bottom 20 portfolios since only the EW portfolio has a statistically significant impact, even though only for a 5% significance level.

Last, but not least, we have to address our measure of sensitivity to market variations, beta. In here, our results reveal different sensitivities for the different variables we used as proxies, which was expected. Nonetheless, independent of the proxy used, the market factor has a statistically significant impact on all portfolios with the Bottom 20 portfolios having, on average, a higher sensitivity than the Top

20 portfolios, with the only exception being the EW Portfolio which reveals a higher sensitivity on the Top 20.

Furthermore, the higher sensitivity is seen when the proxy used is the Stoxx 600, which may be due to the fact that all the companies present on the analysis are constituents of this index. Between each variable used as a proxy for the market, we can see that the difference between Bottom 20 and Top 20 portfolios is wider when we are considering the MSCI, being the former's portfolios the ones with the higher sensitivity. Notwithstanding, both proxies introduced in this study have higher sensitivities than the one used in Fama & French (1993).

As for R^2, the model which seems to better suit the data is the one with the Stoxx 600 as a proxy which, again, may come from the fact that every company in this study is a component of such index. Nevertheless, when we compare the MSCI proxy with the market factor from Ken French's website, we see that the regression from the former has a higher value than the latter, which suggests that the addition of these variables may have improved the model.

4.2. Four-Factor Model

Table 6, 7 and 8 below highlights the outcome of our estimation following the four-factor model proposed in Carhart (1997), based on the previous model but adding the momentum factor proposed by Jegadeesh & Titman (1993), while using alpha (α), as showed in Jensen (1968), to measure the existence and significance of abnormal returns for both portfolios.

Table 6 - OLS estimation of the four-factor model (Monthly data from January 2000 to December 2019);

	Top 20		Bottom 20	
	EW	**VW**	**EW**	**VW**
Constant	0.0094*** (0.0020)	0.0091*** (0.0018)	0.0144*** (0.0020)	0.0137*** (0.0027)
MktRF	0.6345*** (0.0423)	0.5546*** (0.0451)	0.6519*** (0.0516)	0.6549*** (0.0626)
SMB	−0.0499 (0.1296)	−0.3069*** (0.1083)	0.4253*** (0.1204)	0.4202** (0.1839)
HML	−0.0650 (0.1122)	−0.1069 (0.1014)	−0.2547** (0.1212)	−0.3151** (0.1520)
WML	−0.1855*** (0.0587)	−0.0570 (0.0496)	−0.2290*** (0.0661)	−0.3077*** (0.0865)
R^2	0.6109	0.5447	0.5554	0.4952
F-Test	67.9760	49.0932	48.8848	30.8135

Note: Based on 239 observations. HAC robust standard-errors are inserted between parentheses. *** represents a p-value < 0.01, ** represents a p-value <0.05 and * represents a p-value <0.10.

Table 7 -OLS estimation of the four-factor model, with Stoxx 600 as a market proxy (Monthly data from December 2000 to December 2019);

	Top 20		**Bottom 20**	
	EW	**VW**	**EW**	**VW**
Constant	0.0064*** (0.0012)	0.0060*** (0.0013)	0.0121*** (0.0019)	0.0095*** (0.0020)
Exc_Stoxx600	1.0368*** (0.0368)	0.9221*** (0.0394)	0.9803*** (0.0459)	1.0337*** (0.0624)
SMB	0.2566*** (0.0776)	−0.0236 (0.0887)	0.6727*** (0.1266)	0.6799*** (0.1898)
HML	0.0524 (0.0748)	−0.0503 (0.0599)	−0.0478 (0.0994)	−0.1931* (0.1099)
WML	0.0177 (0.0426)	0.1333*** (0.0475)	−0.0808 (0.0666)	−0.1117 (0.0820)
R^2	0.8427	0.8037	0.7219	0.7216
F-Test	220.5939	139.5777	173.7442	108.7554

Note: Based on 228 observations. HAC robust standard-errors are inserted between parentheses. *** represents a p-value < 0.01, ** represents a p-value <0.05 and * represents a p-value <0.10.

Table 8 - OLS estimation of the four-factor model with MSCI as a market proxy (Monthly data from January 2000 to December 2019);

	Top 20		**Bottom 20**	
	EW	**VW**	**EW**	**VW**
Constant	0.0067*** (0.0016)	0.0070*** (0.0017)	0.0115*** (0.0021)	0.0108*** (0.0025)
Exc_MSCI	0.9181*** (0.0575)	0.7735*** (0.0707)	0.9551*** (0.0560)	0.9599*** (0.0667)
SMB	0.2050** (0.1001)	−0.0907 (0.0946)	0.6899*** (0.1210)	0.6861*** (0.1809)
HML	0.3148*** (0.0650)	0.2161*** (0.0644)	0.1392 (0.0972)	0.0807 (0.1087)
WML	−0.1010** (0.0454)	−0.0035 (0.0575)	−0.1463** (0.0707)	−0.2245*** (0.0726)
R^2	0.6819	0.5817	0.6286	0.5552
F-Test	82.1458	44.3602	106.3112	63.8748

Note: Based on 239 observations. HAC robust standard-errors are inserted between parentheses. *** represents a p-value < 0.01, ** represents a p-value <0.05 and * represents a p-value <0.10.

As a start, the introduction of the momentum factor seems to be a positive addition to the explanatory power of our model. Such a statement is supported by the increase on R^2 with only one portfolio (Top 20 VW on Table 8) having a decrease in it. Nevertheless, the difference is very small hence we can consider the overall impact of the introduction of the new factor as positive.

Furthermore, we also kept a statistically significant and positive alpha (constant) which shows that, on average, our portfolios have an excess return, even after introducing one more factor.

Looking at the SMB factor, on the Bottom 20 portfolios we have a similar impact as in the previous model, although one of them (Bottom 20 VW on Table 6) only has a statistically significant impact for a 5% level of significance, instead of the 1% of before. When observing the Top 20 values, again we find the same pattern of mixed outcomes as in the previous model, with an improvement on the impact when we compare table 8 with table 5. the former has a statistically significant impact for a 5% significance level, while the latter only has it for a 10% level.

On HML, again we find its impact more strongly and positively pronounced on Table 8 (with MSCI index as a market proxy) for the Top 20 portfolios while, unlike in the three-factor model, the Bottom 20 portfolios do not have any significant impact. However, the situation changes we consider Table 6 and 7, comparing them with Table 3 and 4. In here, the Bottom 20 portfolios have a significant impact, even though not at the 1% level of significance, that ranges from the VW portfolio on Table 7 (10% level of significance) to both on Table 6 (5% significance level). Furthermore, it seems worth to mention that the HML factor's impact on the Bottom 20 contributes negatively to the alpha constant (excess return), while contributing positively for the Top 20's alpha constant.

The new factor, WML, has a significant, and negative, impact on the first regression (Table 6), being the Bottom 20 EW portfolio the only one who is not affected by it. However, the impact is not as straightforward in the following two variations we tested. On Table 7, we can see that only Top 20 VW portfolio is affected, positively, by it while on Table 8 the broad impact is again seen, even though with a lower statistical significance (most cases are only for the 5% level) and a change on the portfolio that was unaffected, when compared with Table 6, since Top 20 VW is unaffected instead of the Bottom 20 EW portfolio.

Regarding market sensitivity, the addition of the new factor did not bring any interference at this level, since we have the same conclusions as in the previous model, seeing a positive and statistically significant impact on all portfolios and having Table 7 (with Stoxx 600 as a market proxy) as the regression where the sensitivity is higher.

4.3. Five-Factor Model

Table 9, 10 and 11 show the outcome of our estimation following the five-factor model proposed in Fama & French (2015), based on the three-factor model already studied, but adding two more factors related with profitability (RMW) and investment profile (CMA), while using alpha (α), as showed in Jensen (1968), to measure the existence and significance of abnormal returns for both portfolios.

Table 9 - OLS estimation of the five-factor model (Monthly data from January 2000 to December 2019);

	Top 20		**Bottom 20**	
	EW	**VW**	**EW**	**VW**
Constant	0.0090*** (0.0021)	0.0080*** (0.0020)	0.0150*** (0.0023)	0.0154*** (0.0033)
MktRF	0.6254*** (0.0470)	0.5545*** (0.0514)	0.6267*** (0.0667)	0.5940*** (0.0895)
SMB	−0.1507 (0.1329)	−0.3512*** (0.1014)	0.3178** (0.1248)	0.2567 (0.1875)
HML	0.1411 (0.1771)	0.1514 (0.1340)	−0.1736 (0.1640)	−0.1893 (0.2268)
RMW	−0.1147 (0.1775)	0.2813** (0.1373)	−0.4750*** (0.1590)	−0.7550*** (0.2295)
CMA	−0.3792** (0.1834)	−0.2842* (0.1473)	−0.3795* (0.1979)	−0.6031** (0.2676)
R^2	0.5990	0.5536	0.5500	0.5024
F-Test	50.79542	48.73416	41.0839	30.5220

Note: Based on 239 observations. HAC robust standard-errors are inserted between parentheses. *** represents a p-value < 0.01, ** represents a p-value <0.05 and * represents a p-value <0.10.

Table 10 - OLS estimation of the five-factor model, with Stoxx 600 as a market proxy (Monthly data from December 2000 to December 2019);

	Top 20		**Bottom 20**	
	EW	**VW**	**EW**	**VW**
Constant	0.0061*** (0.0012)	0.0052*** (0.0013)	0.0118*** (0.0019)	0.0101*** (0.0019)
Exc_Stoxx600	1.0373*** (0.0390)	0.9211*** (0.0409)	1.0057*** (0.0532)	1.0192*** (0.0695)
SMB	0.2632*** (0.0826)	0.0049 (0.0913)	0.6877*** (0.1303)	0.6518*** (0.2002)
HML	0.0676 (0.1097)	0.0391 (0.0892)	−0.1497 (0.1459)	−0.2741 (0.1874)
RMW	0.0726 (0.1155)	0.4157*** (0.1395)	−0.2174 (0.1616)	−0.3881** (0.1654)
CMA	−0.0093 (0.1153)	0.1115 (0.1133)	−0.0019 (0.1681)	−0.1946 (0.2280)
R^2	0.8425	0.8076	0.7215	0.7243
F-Test	196.9257	128.3673	120.7367	77.8028

Note: Based on 228 observations. HAC robust standard-errors are inserted between parentheses. *** represents a p-value < 0.01, ** represents a p-value <0.05 and * represents a p-value <0.10.

Table 11 - OLS estimation of the five-factor model with MSCI as a market proxy (Monthly data from January 2000 to December 2019);

	Top 20		**Bottom 20**	
	EW	**VW**	**EW**	**VW**
Constant	0.0068*** (0.0017)	0.0064*** (0.0016)	0.0126*** (0.0023)	0.0129*** (0.0030)
Exc_MSCI	0.8874*** (0.0524)	0.7433*** (0.0614)	0.9089*** (0.0643)	0.8821*** (0.0923)
SMB	0.1144 (0.0908)	−0.1378* (0.0790)	0.5933*** (0.1242)	0.5281*** (0.1812)
HML	0.4892*** (0.1372)	0.4658*** (0.1141)	0.1727 (0.1313)	0.1362 (0.1791)
RMW	−0.0924 (0.1577)	0.2731* (0.1423)	−0.4399** (0.1986)	−0.7085*** (0.2177)
CMA	−0.4095*** (0.1506)	−0.3591** (0.1506)	−0.3881** (0.1604)	−0.5886** (0.2276)
R^2	0.6867	0.5990	0.6389	0.5793
F-Test	73.1839	46.1297	85.5053	60.1439

Note: Based on 239 observations. HAC robust standard-errors are inserted between parentheses. *** represents a p-value < 0.01, ** represents a p-value <0.05 and * represents a p-value <0.10.

The addition of the new factors to the three-factor model seems to improve its explanatory power. However, when compared with the improvement realized by adding the momentum factor, we see that it is a very similar improvement. This situation may come from the characteristics of our subset of companies which may expose them more to one factor than another. Nevertheless, we still see a positive and statistically significant alpha (constant) which, as before, suggests the existence of an excess return on the portfolios chosen so we will, as in the previous models, analyze the impact of each factor.

The SMB factor in this regression lost some influence, since not all Bottom 20 portfolios have a statistically significant and positive impact as before. Table 10 and 11 keep such influence but Table 9 shows a different picture with only the EW portfolio having an impact from SMB and only at the 5% level of significance. Notwithstanding, our Top 20 portfolios show the same impact that we saw on the first model, and was little change on the previous model, having only a difference concerning the portfolio that has a negative impact, at the 10% level of significance, on Table 11 (MSCI proxy). On this situation, unlike the previous two models (Table 5 and 8), the portfolio that is impacted is the Top 20 VW and not the EW.

When we look into the impact from the HML factor we can see that, in this model, it has the least impact. In fact, only the Top 20 portfolios in Table 11 (MSCI proxy) have a significant and positive impact from the HML factor while all the other instances studied reveal no impact from this factor. This contrasts with the previous studies since the impact from this factor was felt in other situations, even though not for levels of significance of 1% but is in line with what Fama & French (2015) suggest regarding the redundancy of this factor.

The RMW factor provides us with different outcomes and impacts, being the Bottom 20 portfolios the only ones that have a statistically significant and negative impact for all levels of significance (Table 9). Besides those, we have various level of impact and significance for the Top 20 portfolios, although never going below the 5% level, except for the instance of the VW portfolio on Table 10. The same situation can be observed over the influence of the CMA factor, although in this case there is one regression where this factor has no influence, in whichever portfolio observed (Table 10, Stoxx 600 proxy). Apart from that, we observe several levels of statistical significance for this factor, for both portfolios, always contributing negatively and only achieving the level of 1% significance in one instance: Top 20 EW portfolio on Table 11.

The market sensitivity of this portfolios to the various variables used as proxies is in line with the previous two regressions, registering a higher sensitivity for the Stoxx 600 index, being followed by the MSCI index and the market factor available in Ken French’s website.

Chapter 5 - Conclusion

The present work aimed at adding to an existing debate on the finance world regarding the equality of genre in company's boards. Using the *European Women on Boards'* document, we attempted to analyze the impact of this factor on the financial performance of the companies to check whether there was a significant impact and, if such, in which direction (positive or negative).

In all our models, we could see significant constants (alpha) suggesting that both portfolios have abnormal and significant returns, with the Bottom 20 (portfolio of companies with GDI score below 0.2) having a slight edge over the Top 20 (portfolio of companies with GDI score above 0.8) and, more precisely, the EW (equally-weighted) portfolio, suggesting that the presence of women on the board does not benefit the companies in terms of the market perception regarding value creation. Regarding our proxies for the market variable, our outcome indicates that they have more influence on explaining the returns of our portfolio, even though one of them, Stoxx 600, was already suspected to give us such an outcome since all companies form the portfolios are present in it.

Regarding our models, we see that both the four-factor and the five-factor seems to fare better than the three-factor model even though, between them, it is not easy to establish one as superior. These statements are based on both the statistical significance of the factors and their fit, given by their R^2. The similar pattern of fitness between the portfolios may come from the characteristics of our data, which may make it more exposed to one factor than another, as well as our choice of market proxies (as an example we can see the impact of WML factor with Stoxx 600 as a proxy for our market variable is almost null, while in the model with Ken French's website factor is rather significant).

This also highlights some of the limitations of our analysis. Despite our usage of an official document from an European entity, we only got access to data already processed, and not being able to check how it was treated and organized may hinder our ability to withdraw conclusions from the outcome. Furthermore, regarding, what appears to be an advantage of the Bottom 20 EW portfolio over the Bottom 20 VW portfolio, we should remind that, in this analysis, it was not taken into account any transaction costs, which would have considerable impact on the EW portfolio since it always needs to keep the weight of each component constant, despite its variations.

Expanding on the limitation regarding data access, we can also find an opportunity for further research where we would organize the data in different ways, dividing by the roles each women had on the boards and the actual influence they had on the company's decisions, or allowing for the division between age groups to see which companies are more prone to have a gender diverse board, younger or older. Notwithstanding, our analysis would also benefit on the introduction of conditional performance, allowing for the time variation of betas and verifying whether the impact of gender diversity in boards is relevant on an ever-changing economic environment as the one we live on today.

References

Adams, R. B., & Ferreira, D. (2007). A theory of friendly boards. *Journal of Finance*, *62*(1), 217–250.

Adams, R. B., & Ferreira, D. (2009). Women in the boardroom and their impact on governance and performance. *Journal of Financial Economics*, *94*(2), 291–309.

Ahern, K. R., & Dittmar, A. K. (2012). The changing of the boards: The impact on firm valuation of mandated female board representation. *Quarterly Journal of Economics*, *127*(1), 137–197.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*, 3–18.

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

Black, F., Jensen, M. C., & Scholes, M. (1972). The capital asset pricing model: Some empirical tests. In M. C. Jensen (Ed.), *Studies in the Theory of Capital Markets*. Praeger Publishers Inc.

Campbell, K., & Mínguez-Vera, A. (2008). Gender diversity in the boardroom and firm financial performance. *Journal of Business Ethics*, *83*(3), 435–451.

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

Chapple, L., & Humphrey, J. (2014). Does board gender diversity have a financial impact? Evidence using a stock portfolio performance. *Journal of Business Ethics*, *122*(4), 709–723.

EWOB. (2019). *European women on boards gender diversity index*. https://europeanwomenonboards.eu/wp-content/uploads/2020/06/EWoB-GDI-final-report_20200108-incl.-EU-disclaimer.pdf

Fama, E. F. (1968). Risk, return and equilibrium: Some clarifying comments. *The Journal of Finance*, *23*(1), 29–40.

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

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

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of*

Financial Economics, *116*, 1–22.
Fama, E. F., & French, K. R. (2017). International tests of a five-factor asset pricing model. *Journal of Financial Economics*, *123*(3), 441–463.
Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, *81*(3), 607–636.
Ferson, W. E., & Harvey, C. R. (1991). The variation of economic risk premiums. *Journal of Political Economy*, *99*(2), 385–415.
Ferson, W. E., & Harvey, C. R. (1999). Conditioning variables and the cross section of stock returns. *The Journal of Finance*, *54*(4), 1325–1360.
Ferson, W. E., Sarkissian, S., & Simin, T. (2008). Asset pricing models with conditional betas and alphas: The effects of data snooping and spurious regression. *Journal of Financial and Quantitative Analysis*, *43*(2), 331–353.
Ferson, W. E., & Schadt, R. W. (1996). Measuring fund strategy and performance in changing economic conditions. *The Journal of Finance*, *51*(2), 425–461.
Ferson, W. E., & Siegel, A. F. (2001). The efficient use of conditioning information in portfolios. *Journal of Finance*, *56*(3), 967–982.
Grinblatt, M., & Titman, S. (1989). Portfolio performance evaluation: Old issues and new insights. *The Review of Financial Studies*, *2*(3), 393–421.
Gul, F. A., Srinidhi, B., & Ng, A. C. (2011). Does board gender diversity improve the informativeness of stock prices? *Journal of Accounting and Economics*, *51*, 314–338.
Jagannathan, R., & Wang, Z. (1996). The conditional CAPM and the cross-section of expected returns. *The Journal of Finance*, *51*(1), 3–53.
Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), 65–91.
Jensen, M. C. (1968). The performance of mutual funds in the period 1945-1964. *Journal of Finance*, *23*(2), 389–416.
Jensen, M. C. (1969). Risk , the pricing of capital assets , and the evaluation of investment portfolios. *The Journal of Business*, *42*(2), 167–247.
Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The Journal of Finance*, *20*(4), 587–615.
Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77–91.
Merton, R. C. (1973). An intertemporal Capital Asset Pricing Model. *Econometrica*,

41(5), 867–887.

Nguyen, T., Locke, S., & Reddy, K. (2015). Does boardroom gender diversity matter? Evidence from a transitional economy. *International Review of Economics and Finance*, *37*, 184–202.

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

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

Tobin, J. (1958). Liquidity preference as behavior towards risk. *The Review of Economic Studies*, *25*(2), 65–86. https://doi.org/10.2307/2296205

Treynor, J. L. (1961). *Toward a theory of market value of risky assets*.

Wang, K. Q. (2003). Asset pricing with conditioning information: A new test. *The Journal of Finance*, *58*(1), 161–196.