https://lib.uliege.be https://matheo.uliege.be

The financial leverage in the pricing of eurozone stocks. A study of a leverage augmented fama & french three-factor model

Auteur : Larose, Thomas
Promoteur(s) : Santi, Caterina
Faculté : HEC-Ecole de gestion de l'Université de Liège
Diplôme : Master en sciences de gestion, à finalité spécialisée en Banking and Asset Management
Année académique : 2021-2022
URI/URL : http://hdl.handle.net/2268.2/14201

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.

THE FINANCIAL LEVERAGE IN THE PRICING OF EUROZONE STOCKS.

A STUDY OF A LEVERAGE AUGMENTED FAMA & FRENCH THREE-FACTOR MODEL

Jury :
Supervisor :
Caterina SANTI
Reader(s) :
Boris FAYS

Master thesis by
Thomas LAROSE
For a Master's Degree in Management
Sciences with specialization in
Banking & Asset Management
Academic year 2021/2022

ACKNOWLEDGEMENTS

The realisation of this master thesis was possible thanks to the help of several people to whom I would like to express my gratitude.

I would like to thank my supervisor Ms Caterina Santi, for her time and her answers to my questions, without which I would not have been able to complete this work. She also provided me with valuable advice that deepened my understanding of certain concepts in finance and statistical methods.

I would like to acknowledge Mr Thierry Larose and Ms Chloé Kim who agreed to proofread this work and share with me their opinion on the content and the language. Their advice was very helpful.

I also want to thank Mr Boris Fays, for having given some of his time to read and assess this thesis.

Finally, I would like to thank all the anonymous internet users who freely share clear programming and quantitative analysis insights. With their help, I was able to build my computer program for the empirical part of this study.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

A/E	Asset-to-equity ratio
B/M	Book-to-market ratio
CAPM	Capital Asset Pricing Model
D/A	Debt-to-asset ratio
D/E	Debt-to-equity ratio
HML	High minus Low
LHS	Left-hand side
LMU	Levered minus Unlevered
OECD	Organisation for Economic Co-operation and Development
RHS	Right-hand side
SMB	Small minus Big
VW	Value-weighted

1. INTRODUCTION

The object of this thesis is to investigate whether the degree of financial leverage in firms can explain part of their stock returns. To do so, we will use a Fama and French (1992, 1993) three-factor model augmented with a financial leverage risk factor, and see if it is a good model to describe returns on eurozone stocks from 1989 to 2021. We chose the euro area because it characterises the overall economy in Europe well, and we found that this geographical area is under-represented in the asset pricing literature. The period covers all the data available for the companies in that region.

A plethora of papers boast of having discovered new factors describing the movement of stock returns. Harvey et al. (2016) drew up an inventory of all the published factors since 1964; the total number of factors reached 316. The purpose of their article was to question the reliability of certain findings in the asset pricing literature. Indeed, such research can often lead to flawed data mining results that are statistically significant by chance. It is often caused by lack of a solid theoretical basis. Another reason for the emergence of this bias is stated in their paper: “the cost of data mining has dramatically decreased. In the past, data collection and estimation were time intensive, so it was more likely that only factors with the highest priors – potentially based on economic first principles – were tried” (Harvey et al., 2016, pp. 36–37). They also argued that replication studies are often costly and fastidious, therefore “there is a bias towards publishing “new” factors rather than rigorously verifying the existence of discovered factors” (Harvey et al., 2016, p. 11). In their book, Berkin and Swedroe (2016) stated the criteria needed for a factor premium to be considered for an investment. It must be persistent (over time), pervasive (holding across regions), robust (holding for various definitions), investable (feasible in reality) and intuitive (strong theoretical base). With this in mind, one can question the interest and contribution of our research question. Nevertheless, we believe that we have a sound theoretical basis that explains the existence of the factor we are about to address. This factor is not new in the published literature and, as we shall see, is rather debated. We will try to set out the objectives of this work and how it can contribute to the existing literature.

Firstly, there have been few studies made on our data in the scientific literature, either geographically or temporally. In fact, most studies focus on US stock market returns. In order to validate certain models that work particularly well in the United States, it is necessary to apply them to foreign data sets. Moreover, the majority of the papers found focus on a period prior to the 2008 global financial crisis. Our sample includes the latter, as well as the economic crisis related to the Covid-19 pandemic. It may be of interest to know whether these regime changes have had an impact on the performance of models that seek to understand the variability of stock returns.

Secondly, the role of leverage in asset pricing is quite controversial in the scientific literature. As we will see in our research, several authors contradict each other on the subject. We do not claim to be able to settle the debate. However, we hope to make a contribution with this new study, in order to advance research on the topic.

Thirdly, we wanted to have a tool that would be able to assist finance professionals in their work. We found that our model performs better than the Fama and French (1992, 1993) three-factor model on our sample. The model could be used to estimate the cost of equity of a project or a stock, help asset managers in their performance attribution, determining abnormal return in event studies, and understand the sources of risk in a portfolio of equities. Moreover, the permanent quest for an alpha is becoming more and more an unreachable and unattainable goal. The era of the Capital Asset Pricing Model [CAPM] is over. With the arrival of multi-factor asset pricing models in the world of empirical finance, it is now more relevant to focus on the exposure to different risk factors and to take advantage of risk premiums. Smart beta strategies are overshadowing alpha-seeking strategies (Singer, 2018).

A few sub-questions came to mind after we understood the *raison d'être* of our work. These will be explained more in details in the chapter on hypotheses (see Chapter 3). However, we can already draw some guidelines: Are the expected common stock returns negatively related to the financial leverage? Does the Fama and French (1992, 1993) three-factor model already capture financial risk in its factors? Is the importance of financial leverage dependent on the business cycle as a risk factor?

Our research has led us to discover that highly leveraged stocks have lower returns. The presence of the value effect is confirmed. However we did not find the existence of the size effect in our sample; this could be the subject of a new study using the same data we used. In addition, we found that the value factor and the financial leverage factor share significant explanatory power for the stock returns in our sample. We also have reason to believe that a leverage augmented Fama and French (1992, 1993) three-factor model does a better job than the original model. Moreover, the model remains better regardless of the period used for our analysis. However, our model seems to work less well on small-cap stocks whose market value is much lower than their book value (i.e. small deep value stocks); this problem could also be the subject of further research.

Much of the scientific literature we found seems to agree that leverage and financial distress are two factors that can explain some of the returns on stocks. However, there is not always consensus on all the points discussed: some believe that these factors have a negative relationship with returns while others believe the opposite; we have tried to contribute to this debate. Other researchers disagree on the origin of these factors, some think they are rationally explainable, and others argue that they are due to a behavioural bias of stock market participants, such as overreaction to market events. Furthermore, we have not found any articles that address the link between this research question and the various events that have affected the financial markets since the 1990s.

We will begin by presenting the concepts addressed in our presentation and attempt to define them in order to circumscribe our research. We will then draw up a timeline of all the publications on the subject, and present the theoretical foundations of our investigations. Afterwards, we will outline our hypotheses and briefly summarise the conclusions. Subsequently, we will explain how we handled our data, from cleaning to preprocessing. After that, we will explain in detail the methodology chosen to answer our research question. Next, we will objectively present our results in detail with the help of tables and graphs. Finally, we will discuss the implications of our results for financial market participants and try to conclude our study succinctly and propose some ideas for future research on the same issue.

2. LITERATURE REVIEW

2.1. Definitions of the concepts covered

We believed that it was necessary to define certain terms that will often come up in this document. In this way, we will avoid any misunderstanding of the use of these terms. These definitions were taken from the CFA Program Glossary provided by the CFA Institute on their website, which offers clear and concise entries for definitions of financial concepts.

Let's start with the notion of capital structure. The capital structure is defined as such: "The mix of debt and equity that a company uses to finance its business; a company's specific mixture of long-term financing" (CFA Institute, 2022). Therefore, the capital structure does not take into account short-term financing such as credit lines or commercial papers. All companies have had to face this dilemma at some point namely, how much debt should we take on? What should be our target capital structure? Several researchers have tried to find out which factors influence the decision of managers in choosing the capital structure, they are not short of theories rationalising the choice of decision makers. Modigliani and Miller (1958) stated that in a world without transaction and bankruptcy costs, taxes, and everyone has the same information, the capital structure choice is irrelevant;

> "To understand this proposition, we can think about two companies with the same expected, perpetual cash flows and uncertainty and, hence, the same discount rate applied to value these cash flows. Even if the companies have different capital structures, these two companies must have the same present value using discounted cash flow models. If capital structure changes were to have any effect on a company's value, there would be an arbitrage opportunity to make riskless profits" (Aggarwal et al., 2020, p. 8).

Modigliani and Miller (1958) were aware of the unrealistic nature of their assumptions. Therefore, they extended their theory. When we account for taxes, increasing the leverage increase the value of the firm "because interest paid is deductible from income for tax purposes in most countries, the use of debt provides a tax shield that translates into savings that enhance the value of a company" (Aggarwal et al., 2020, p. 11). New theories then appeared, bringing in their own set of explanatory variables in addition to the "tax" factor. For example, we can mention the cost of financial distress, which we will discuss later; agency costs, which are the "costs associated with the conflict of interest present when a company is managed by non-owners. […] result from the inherent conflicts of interest between managers and equity owners" (CFA Institute, 2022); and asymmetric information, "the differential of information between corporate insiders and outsiders regarding the company's performance and prospects" (CFA Institute, 2022). Moreover, Titman and Wessels (1988) listed no less than seven determinants of capital structure; they emphasised that businesses that can impose high costs on their customers, employees, and suppliers in the event of liquidation have lower debt ratios. Companies will therefore use a cost-benefit optimisation formula, incorporating the above-mentioned costs into their inputs, to determine their optimal debt level.

Now that we have analysed the cause of indebtedness, let's refocus on the consequences of it, through the notion of leverage for a company. We defined it as follows:

> "In the context of corporate finance, leverage refers to the use of fixed costs within a company's cost structure. Fixed costs that are operating costs (such as depreciation or rent) create operating leverage. Fixed costs that are financial costs (such as interest expense) create financial leverage" (CFA Institute, 2022).

A firm's leverage is therefore divided into two distinct parts. Operating leverage is "the use of fixed costs in operation" (CFA Institute, 2022); this concept will be used very little in our research. As for financial leverage, it is at the core of our research. It is defined as

> "The extent to which a company can effect, through the use of debt, a proportional change in the return on common equity that is greater than a given proportional change in operating income; also, short for the financial leverage ratio" (CFA Institute, 2022).

Thus, high leverage would imply greater earnings volatility in theory, and therefore increase risk. However, In times of growth a highly leveraged firm should, all other things being equal, see its earnings rise more strongly than a less leveraged competitor. Financial leverage is often measured using different easy calculated balance sheet ratios: debt-to-equity [D/E], debt-to-asset [D/A], asset-to-equity [A/E], etc. Some of them are harder to quantify but we will not use them in our research to keep it straightforward.

The literature we are about to discuss also refers a lot to a somewhat more abstract concept namely, financial distress. It is defined as follows: "Heightened uncertainty regarding a company's ability to meet its various obligations because of lower or negative earnings" (CFA Institute, 2022). Leverage and financial distress are therefore closely linked, but not synonymous; a high degree of leverage may increase probability of financial distress because of increased earnings uncertainty. However a higher cost of financial distress does not imply a higher leverage. On the contrary, companies with high cost of financial distress should optimally use less leverage to enhance their ability to meet their debts (George & Hwang, 2010). When we speak about "cost of financial distress", we are referring to the principle of optimising the capital structure, which considers it as an expected cost to be taken into account.

The cost of financial distress is more difficult to measure than financial leverage. Some researchers have tried to develop models to estimate the cost of financial distress of a company. Altman (1968) built a regression model predicting the chances of a company going bankrupt, based on five different financial ratios of solvency, liquidity and profitability. He found that using multiple ratio analysis predicts better than using one with a single ratio. Ohlson (1980) took Altman's model and refined it by adding factors and testing it on a larger panel of data; its predictive power improved significantly, but there are many pitfalls with this type of model: it is based on accounting data that does not necessarily reflect a current situation and it incorporates the going-concern assumption. It is nowadays more common to evaluate the cost of financial distress with models using market values as inputs, such as CDS spreads, or using an option pricing model. It is even advisable to combine market and book values to improve the effectiveness of the prediction model (Ciesielski et al., 2019).

In our research, we used financial leverage as a proxy of company's financial risk that is to say "the risk arising from a company's obligation to meet required payments under its financing agreements" (CFA Institute, 2022). As demonstrated earlier, we believe that leverage embodies a multitude of information, including the cost of financial distress, that has a significant influence on a company's earnings, and thus on the performance of its stock. This is the focus of our study.

2.2. A brief history of financial leverage and distress in asset pricing

In order to understand the evolution of research on the subject that we will address throughout this thesis, we will try to draw up a chronology of financial leverage and distress in asset pricing in finance literature. As we shall see, some studies contradict each other on specific issues, which makes the subject all the more interesting. The chronology is non-exhaustive; we have kept only those papers published in a reputable peer-reviewed journal, and those that have contributed something new to the literature. This timeline took his inspiration from the one used in Mirza et al. (2013).

Modigliani and Miller (1958) developed a theoretical model for calculating the cost of equity of a project or a company. At the time of publication, it was often assumed that the cost of debt was equal to the cost of equity. Their main propositions are as follows: "The market value of any firm is independent of its capital structure" (Modigliani & Miller, 1958, p. 268) and "the expected yield of share of stock is equal to the appropriate capitalisation rate […] plus a premium related to financial risk equal to the debt-to-equity ratio" (Modigliani & Miller, 1958, p. 271). They have also shown that when we include taxes in our equation, it is more advantageous for a company to increase its financial leverage, all other things being equal. Their theory therefore implies a positive risk premium for anyone investing in highly leveraged stocks, the market beta of equity being equal to the firm's asset beta plus a factor proportional to business's leverage ratio. Although we did not obtain the same results, we agree that leverage is a factor in equity returns.

Bhandari (1988) found that expected common stock returns are positively related to financial leverage, controlling for size and beta, as well as including and excluding January. They suggested that debt-to-equity is a good proxy for the risk of common equity and that this ratio is not acting as a substitute for beta. According to them, this risk factor should therefore be independent.

Fama and French (1992, 1993) gave birth of the three-factor model; they added a size and a book-to-market factors to the CAPM's beta (Sharpe, 1964), and claimed that it captured better variation in stock returns. They stated that a high book-to-market ratio tells us that a firm's market leverage is high relative to its book leverage; the value factor already encompasses the financial leverage factor. Therefore, stocks with high market leverage should perform better than stocks with high book leverage; the way leverage is calculated – whether it is from book or market value – is crucial. Fama and French (1995) tackled the leverage risk factor by concluding that it was entirely included in the value factor, and therefore should not be included independently in a multi-factor model.

Dichev (1998) showed that financial distress risk does not give higher returns; and therefore it should not be included in the value and size factors. Griffin and Lemon (2002) argued that those firms with high distress risk underperformed because of small low book-to-market companies included in the sample; i.e. small companies with low analyst coverage. They stated:

> "Firms with high distress risk have characteristics that make them more likely to be mispriced by investors. Consistent with mispricing arguments, firms with high distress risk exhibit the largest return reversals around earnings announcements, and the book-to-market return premium is largest in small firms with low analyst coverage" (Griffin & Lemon, 2002, p. 2335).

Ferguson and Shockley (2003) distinguished between leverage (measured by debt-to-equity) and financial distress (measured by z-score) and that these two factors in addition to market risk were sufficient to explain equity returns. The distinction they make between the two concepts is of particular interest, and they describe it with a clear example:

> "[…] a firm with substantial cash flow and few growth opportunities might find high debt levels attractive and could appear in the highest debt-to-equity categories without risking bankruptcy. Conversely, a firm in the middle leverage portfolios but with highly volatile cash flow could face substantial risk of distress" (Ferguson & Shockley, 2003, p. 2561).

By adding variables related to debt, and thus to the debt market, they hope that their model will overcome the problem of the CAPM (Sharpe, 1964), which is limited to the equity market as a proxy for the overall market.

Vassalou and Xing (2004) argued that default risk is a systematic risk factor. Size and value factors contain some default risk, but that is not their rationale. They have shown that size and value factors only exist among companies with a high risk of bankruptcy. They also advocate the use of market value instead of book value when estimating the financial risk of a company.

Dhaliwal et al. (2006) extended the theory of the Modigliani and Miller (1958) model by stating that corporate taxes reduce the financial risk premium, while personal taxes increase it. Therefore, the relevance of leverage depends on the balance between personal taxes and corporate taxes in the investor's country.

Penman et al. (2007) thought that the book-to-market [B/M] ratio could be divided into two distinct parts as to say, into an operating risk (measured by the book value of net operating assets divided by the market value of net operating assets) and a financial risk (measured by the market value of net debt divided by the market value of equity). According to their research, the operating risk is positively related to expected returns while financial risk is not.

Campbell et al. (2008) results brought a noteworthy dare to the assumption that the value and size risk premiums are proxies for financial distress premium; even though financially distress stocks have high loadings on Small minus Big [SMB] and High minus Low [HML] factors, they have anomalously low returns. They suggested that this anomaly existed because of behavioural bias in the investor's mind.

George and Hwang (2010) tried to solve a puzzle that was debated for some time among scholars. They wanted to know why returns and measures of leverage have a negative cross-sectional relation. Indeed, they discovered that firms with low leverage and high distress risk have significant positive risk-adjusted returns. What is important for them is the assumptions of market frictions. This is their answers to solve the puzzle:

> "The idea that equity risk is increasing in leverage relies on the frictionless markets assumption that makes investment and financing decisions separable; […]. Market frictions could lead low-leverage firms to have greater exposures to systematic risk, which dominates the amplification effect of leverage on equity risk. […]. Using a very simple model, we show that if financial distress is costly and firms make optimal capital structure decisions, then low-leverage firms are exposed to greater systematic risk than high-leverage firms" (George & Hwang, 2010, p. 57).

The idea behind this is that, financial distress cost is a significant factor when it comes to determining the capital structure for companies; firms with high costs select low leverage and have the greatest exposure to systematic risk, and vice versa. Furthermore, they uncovered that: "Book-to-market is not a measure of financial distress risk but instead captures exposure to priced risk that is unrelated to capital structure" (George & Hwang, 2010, p. 76). Their proposal is the one we found most plausible and consistent with our research findings; we will explore this link further later.

Chou et al. (2010) concluded that leverage and financial distress risk encompass the value premium. They combined the Fama and French (1992, 1993) three-factor model with the two factors highlighted by Ferguson and Shockley (2003) – i.e. financial distress and financial leverage – , and found that their model was able to describe almost all anomalies, none of which are captured by either of the two models individually.

Ozdagli (2012) did not deny the thesis of Modigliani and Miller (1958), they preferred to contain it by adding the concept of investment irreversibility; they said: "Financial leverage directly affects stock returns through its effect on equity risk as in Modigliani and Miller (1958) and indirectly through its effect on business risk, by influencing investment decisions" (Ozdagli, 2012, p. 1035). They further argued that the market value of leverage explains a large part of the value premium; however they found that "tax deductibility of interest payments increases effective investment irreversibility and that investment irreversibility weakens the relation between book-to-market values and returns" (Ozdagli, 2012, p. 1033). They also emphasised the difference between market leverage and book leverage, these two giving different interpretation.

Obreja (2013) paper stated that both the value premium and the book leverage premium are linked to operating leverage; operating leverage is the driving force, and without it, the value premium is negative. Furthermore, equity risk premium and book-to-market most of the time have a positive relationship. However, equity risk premium and book leverage relationship is not steady, and is potentially negative. He also found that the book-to-market ratio helps identify firms that have low productivity and with a high financial and/or operating leverage; in other words, companies with low book-to-market ratio face either high operating leverage or financial leverage. On the other hand, he claimed that book leverage is not efficient in explaining stock returns because low and high leverage firms can both have high equity risk premiums. Obreja also argued that market leverage is better in explaining stock returns than book leverage.

Gao et al. (2018) used a panel of world data to analyse the relation between financial distress and equity returns. They found a strong negative relationship between financial distress and equity returns, especially for small capitalisation stocks in developed countries. They objected to the rational explanation of this anomaly and looked instead at a behavioural justification; in their view, stocks with a high probability of bankruptcy are temporarily overpriced.

To conclude, it is not easy to find a consensus on the influence of leverage on stock returns. Modigliani and Miller (1958), and Bhandari (1988) are convinced that the relationship between leverage and expected return is positive. On the other hand, Penman et al. (2007), Campbell et al. (2008), and George and Hwang (2010) found the opposite. It is therefore difficult to draw definitive conclusions on this subject. Also, financial leverage relationship with the value premium is not steady across literature: Fama and French (1995) argued that the leverage factor was already embedded in the value factor, while Penman et al. (2007) and Chou et al. (2010) found that value factor did not explain entirely the leverage factor. However, most researchers agree on some points: there is a distinction between market value and book value of leverage, and the use of the former should be preferred; leverage and financial distress are factors that have explanatory power on stock returns. Fama and French (1992, 1993, 1995) believed that adding a leverage factor to their model would be redundant, however, using their methodology, we found evidence that it is not the case. We also noted that some researchers attribute the origin of the leverage factor to behavioural reasons, and thus that it has no real rational basis.

2.3. Theoretical basis of our research

As mentioned earlier, Modigliani and Miller (1958) demonstrated that the level of debt of a company is a significant factor if one wishes to estimate the required return of a stock. They also showed that a more leveraged company should deliver a higher though riskier return. Although their thesis is crucial in the theory of capital structure, it is based on too many unrealistic assumptions. Numerous research studies that have improved their model have been published since then.

The static trade-off theory of capital structure took the parameters of the Modigliani and Miller (1958) model, the level of indebtedness and the tax rate, and added a new one which is the cost of financial distress. With the first theory, it was rational to have indebtedness tending towards infinity given the absence of constraints in the model. This new theory took the cost of financial distress as a brake on financial leverage: from a certain point onwards, the cost of leverage outweighs its benefits. The level of leverage is therefore chosen using an optimisation function that takes these parameters into account (Aggarwal et al., 2020).

This new approach still implies a positive relationship between leverage and the expected return on a share; investors should get a positive premium for investing in levered companies. George and Hwang (2010) refuted this statement by arguing that their relationship is negative. They supported the fact that this idea is based on the frictionless markets assumption; when we take market friction into account in our model, we expect companies to consider investment and financing decisions together. Thus, their choices would take into consideration the asset risk. In their view, this implies that companies with a high cost of financial distress need to keep their leverage low, as they are much more vulnerable. An increase in the cost of financial distress accentuates the exposure to systematic risk and such a risk would be rewarded with higher expected return. This can be explained because "distress costs depress asset payoffs in low states. Since the occurrence of low states is at least partly systematic, distress costs heighten exposure to systematic risk" (George and Hwang, 2010, p. 57). This systematic risk takes precedence over the risk of increasing debt levels.

What followed from their research is that "a firm with high costs of financial distress optimally chooses lower leverage, has a lower probability of distress, and has a greater expected return than an otherwise identical firm with low distress costs" (George and Hwang, 2010, p. 60). The results we obtained from our research support what they stated.

We believe that the value factor and the leverage factor have something in common, but each of them has a distinct explanatory part. It could be that the value factor explains the increase in equity risk associated with higher debt, as mentioned by Modigliani and Miller (1958). While the leverage factor explains the financial distress cost theory put forward by George and Hwang (2010). We discuss this further in Chapter 7, using the results we have obtained.

This is a sensitive issue because the nature of the value factor is rather arcane; we know that it exists empirically and that it is not an anomaly, but it is difficult to interpret its existence. It is difficult to know exactly why some companies have a market value much lower than their book value. The implication is that their estimated future cash flows are discounted at a higher rate, and therefore reflect a greater uncertainty in their realisation. Fama and French (1992, 1993, 1995) argued that leverage and financial distress are integrated into the value factor with other variables. However, other researchers (Chou et al., 2010; Penman et al., 2007) showed the opposite: part of capital structure factor is not explained by the value factor. We wanted to further investigate this issue and we found that financial leverage retains a specific explanatory power.

3. HYPOTHESIS

Based on our theoretical research and our initial research question, we drew up the hypothesis surrounding the leverage augmented Fama and French (1992, 1993) three-factor model that we felt were most appropriate to add to the body of research related to this topic.

H1: *Expected common stock returns are negatively related to the financial leverage.*

As shown by George and Hwang (2010), firms with low leverage have a high cost of financial distress; therefore they have a higher expected return because of their higher exposure to systematic risk induced by financial distress. We believe that the systematic risk induced by firms with a high cost of financial distress takes precedence over the risk of increasing debt levels. Our results showed that low leverage portfolios give a higher value-weighted [VW] average excess return than those with high leverage. Furthermore, The Levered minus Unlevered [LMU] factor constructed as the difference, each month, between the simple average of the returns on high-leverage portfolios and the simple average of the returns on low-leverage portfolios, delivers strong negative returns on a risk-adjusted basis.

H2: *A factor model including financial leverage is a better proxy for common risk factors in returns than the original Fama and French three-factor model.*

There is evidence that financial leverage is a cogent factor in describing stock returns. We adopted the methodology of Fama and French (2015) to carry out our research on the subject. We found that including a financial leverage factor measured as debt-to-equity ratio is a better proxy for common risk factors in returns than the original three-factor model. Indeed, the statistical measures obtained from our regressions show a decrease in the intercept, whether we take all the dependent variable's portfolios analysed at the same time, or one by one.

H3: *The Fama and French three-factor model does not capture entirely financial leverage risk in its factors.*

We discovered a significant positive relationship between the value factor (HML) and the financial leverage factor (LMU). This has been confirmed by examining their correlation and by regressing them on each other with the other factors. However, they both seem to have a distinct explanatory capacity. We found that a multi-factor model without the value factor and another without the financial leverage factor underperform a model combining both.

H4: *The importance of financial leverage as a risk factor does not depend on the business cycle.*

This hypothesis is slightly more original than the others and is intended to extend a somewhat different field of research. The period of our sample is punctuated with events that have shaken the financial markets, in particular the equity markets. We wondered whether asset pricing models become less effective in the face of market turmoil. We tested the robustness of our model by testing it over several time scales in our sample, using rolling windows. We found that the performance of our model fluctuates well over time. However it remains strong. We tried to relate the performance of the model to some macroeconomic variables in order to find links that could explain these fluctuations, but we did not find anything conclusive in this respect.

4. DATA

4.1. Data collection

The data was mainly collected using the Thompson Reuters DataStream software. This software is accessible from the teacher's computers in the trading room at HEC Liège. The program can be used from Excel, via the DataStream tab. The risk-free rate was retrieved from the Organisation for Economic Co-operation and Development [OECD] (2022) website, as the German 3-month Bubill[1]. This choice is explained by the fact that Germany is the European country with the most stable economy, and we did not find any relevant data for the one-month rates. We created the following three Excel workbooks:

- Benchmark
- Risk-free rate
- Main data

We created a table showing the panel data extracted for each workbook (see Appendix 1). The data in the last workbook is extracted from ten countries: Austria, Belgium, Finland, France, Germany, Ireland, Italy, Netherlands, Portugal and Spain. We chose these countries because they are the largest markets in the eurozone and they correspond to the components of our benchmark (S&P Euro). Furthermore, these countries represent a significant part of the European Union GDP (Eurostat, 2022b) so they are good proxies for European financial markets. We preferred to limit ourselves to the euro as a currency, to avoid exchange rate problems. For each country, we selected those stocks that met the following criteria:

- Equity
- Denominated in euro
- Not a financial company (bank, financial services, non-life insurance and life insurance)
- Primary quote (in order to avoid counting companies listed on several stock exchanges more than once)

We made a detailed description, according to Thompson Reuters DataStream, of each of the variables we retrieved to perform our analysis (see Appendix 2). Thus, we obtained an Excel file containing five separate columns for each company, with the rows corresponding to the dates.

4.2. Data cleansing

We built a sample representing the sample size before cleaning (see Appendix 3). The figures shown in this sample seem unrealistic and therefore need some cleaning up.

The next steps were carried out using the Python programming language (Python Software Foundation, 2021) including pandas (McKinney, 2010), NumPy (Harris et al., 2020), Matplotlib (Hunter, 2007), seaborn (Waskom, 2021) and statsmodels (Seabold et al., 2010) libraries. We chose this programming language because it is a practical and powerful language for handling large amounts of data. Still, we will mainly use R (R Core Team, 2021), including GRS.test (Kim, 2017), readxl (Bryan &

[1] Bubills (officially known as *Unverzinsliche Schatzanweisungen des Bundes*) are Federal Treasury discount papers with a maturity of up to 12 months.

Wickham, 2019) and xlsx (Arendt & Dragulescu, 2020) packages, for our regressions because this language is very useful in statistical analysis.

We first imported the three workbooks from Excel and reduced the time frame from December 1988 to May 2021. The risk-free rate was first divided by one hundred so that it is in the same format as returns. It was then restated as a one month period using a compounding formula, these being basically annualised.

The first major clean-up was to remove equities for which any of the five columns return an '#ERROR' value (see Appendix 4). Then we deleted all German stocks labelled '(XET)'. We noticed that these are duplicates, so leaving them in the sample would result in identical values appearing twice (see Appendix 5). Next, we also deleted the stocks that include a 'MARKET VALUE.1' column. The '.1' means that this value is already somewhere in the sample, so these should be deleted. Afterwards, we removed equities for which any of the five columns is full of 'NA'. It is not useful to include them in the sample because the lack of data for one of the five columns makes it impossible to use them. Finally, we calculated price returns for each stock, and replace every zero value by 'NA' value. Zero returns are usually eliminated when working with DataStream because it maintains the last available price when a stock ceases to be quoted.

After data cleansing, we had the size of our sample for each country (see Appendix 6). We see that the variation between the number of observations before and after the cleaning grows strongly, from -13% to -74%. This is due to the fact that we have discarded all non-relevant period data for stocks that have been delisted, bankrupt or merged. Moreover, France is the most represented country on average, followed by Germany and Italy. The sample size is the largest in the early 2000s.

4.3. Data preprocessing

The returns for each stock and for the benchmark were simply calculated using this formula:

$$R_{it} = \frac{P_{it}}{P_{it-1}} - 1$$

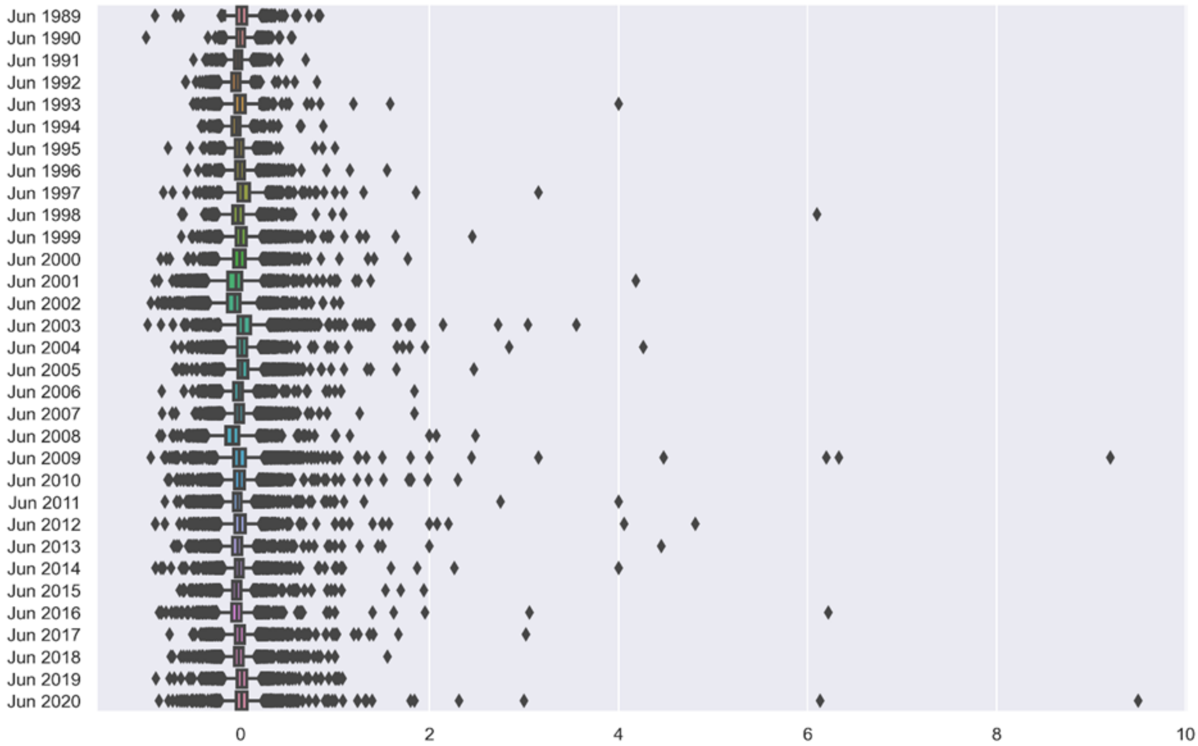
Where R_{it} is return on stock i for the month t
P_{it} is the price of stock i at month t
$P_{it-1}$is the price of stock i at month t-1

Monthly common shareholder's equity, total debt and total asset were divided by one thousand, in order to be in the same format as market value (i.e. in millions).

We then looked at summary statistics for our five variables (return, size, book value of equity, book value of debt and book value of asset). We first deleted all the returns above ten, as this seems unrealistic.

We built a boxplot for the price return variable and saw if there were too many outliers (see Figure 1). We have only taken the month of June of each year, as taking every month of the year would have created an unclear graph. However, the same could have been done for any month.

Figure 1: Price return boxplots from June 1989 to June 2020



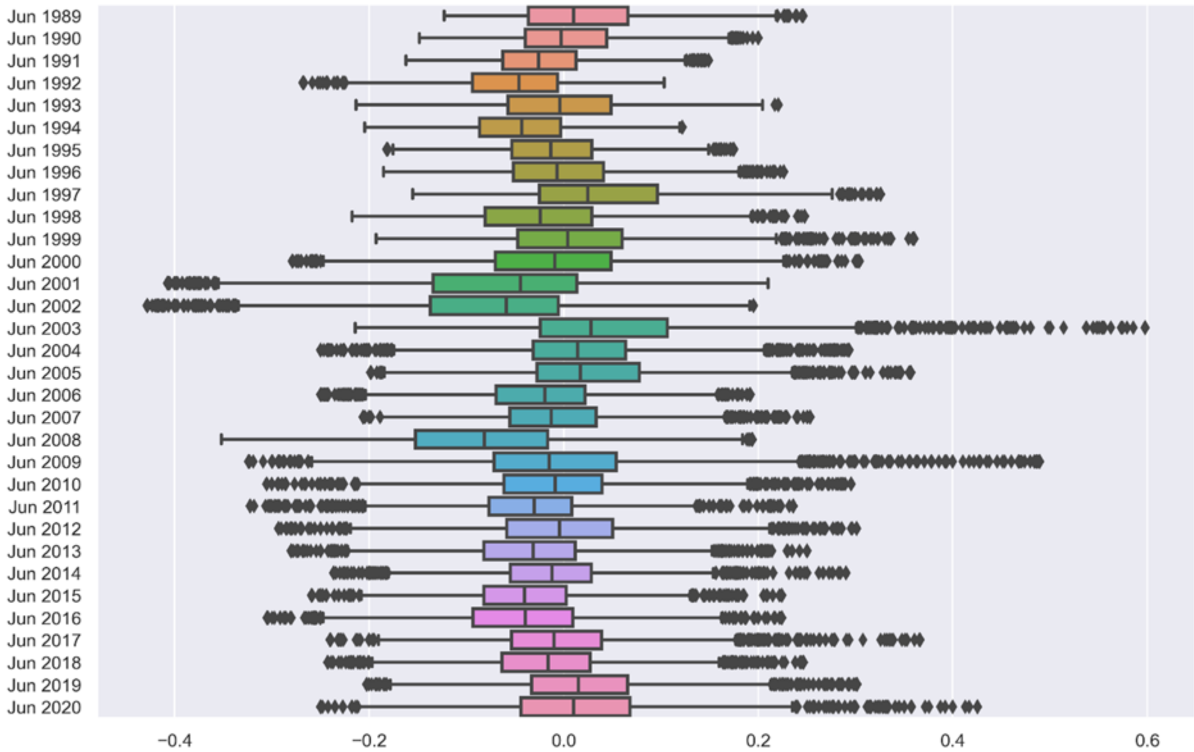
Source: Based on the author's calculation run in Python program using data from DataStream.

Even after eliminating returns higher than ten, some observations remained with an implausibly high value. It is unlikely that a company would have increased in value eightfold in a single month.

Adams et al. (2013), in their framework for handling outliers in finance, recommended identifying extreme values of all variables by computing the minimum and maximum value observations and those in the 1st, 5th, 50th, 95th and 99th percentiles. Therefore, we built a summary table of these statistics for each of the variables (see Appendix 7–Appendix 11). By looking at those statistics, we concluded that many outliers existed in each variable, especially for high values. The difference between the maximum and 99th centile can be quite high for all factors. As a consequence, winsorizing the data may be considered to exclude these outliers.

We can see what we got after winsorizing the observations at the top and bottom 2.5% for price returns (see Figure 2).

Figure 2: 5% winsorized price return boxplots from June 1989 to June 2020



Source: Based on the author's calculation run in Python program using data from DataStream.

And for the rest of variables, see Appendix 12–Appendix 16.

The sample then appeared to be free of extreme values. Since the dataset has been cleared of its outliers, we computed the book-to-market and the debt-to-equity ratios.

$$B/M = \frac{Common\ shareholders'\ equity_{it}}{Market\ value_{it}}$$

$$D/E = \frac{Book\ value\ of\ debt_{it}}{Common\ shareholders'\ equity_{it}}$$

These ratios will be used to construct the portfolios we will use in our statistical analysis. We chose the D/E ratio because it is the one most represented in the scientific literature. Also, it is very simple to calculate and is easily available. This ratio represents the leverage of a company in the strict sense, which is why we preferred it to a ratio including income statement or cash flow data such as the Net debt/EBITDA or Net debt/Free cash flow multiples. Leverage could also have been defined using the A/E and D/A ratios. However, they are corollaries of the D/E ratio. Their result gives us the same interpretation with a few slight differences.

5. METHODOLOGY

5.1. Portfolio construction

5.1.1. 5x5 sort

These portfolios will serve as dependent variables for the regression, also called left-hand side portfolios [LHS]. Portfolio construction is largely inspired by Fama and French (2015) methodology outlined in their paper.

We first calculated the breaking values that will allow us to classify all the stocks in the appropriate portfolio according to size, value or leverage criteria.

At the end of June, breaking values are computed for Size, B/M and D/E as follows:

- Size breaking values are the four quintiles for market cap at the end of June t
- B/M breaking values are the four quintiles for B/M at the end of December t-1
- D/E breaking values are the four quintiles for D/E at the end of December t-1

Afterwards, at the end of June, stocks are allocated independently to five Size, B/M and D/E groups according to their breaking values computed before. The intersection of the sorts produces 25 Size-B/M and 25 Size-D/E portfolios. Portfolios are rebalanced and reconstituted each year at the end of June. Specifically, constituents and their weights remain constant for twelve months. Weights correspond to market cap. Excess returns are computed by subtracting value-weighted monthly portfolio returns by the monthly risk-free rate.

5.1.2. 2x4x4 sort

These portfolios will serve as dependent variables for the regression or LHS portfolios. Portfolio construction is largely inspired by Fama and French (2015) methodology outlined in their paper.

At the end of June, breaking values are computed for Size, B/M and D/E as follows:

- Size breaking value is the median for market cap at the end of June t
- B/M breaking values are the three quartiles for B/M at the end of December t-1
- D/E breaking values are the three quartiles for D/E at the end of December t-1

Afterwards, at the end of June, stocks are allocated independently to two Size, four B/M and D/E groups according to their breaking values computed before. The intersection of the sorts produces 32 Size-B/M-D/E portfolios. Portfolios are rebalanced and reconstituted each year at the end of June, specifically, constituents and their weights remain constant for twelve months. Weights correspond to market cap. Excess returns are computed by subtracting value-weighted monthly portfolio returns by the monthly risk-free rate.

5.1.3. 2x3 sort

These portfolios will serve to build the four factors that will be used as independent variables for the regression, also called right-hand side portfolios [RHS]. Portfolio construction is largely inspired by Fama and French (2015) methodology outlined in their paper.

At the end of June, breaking values are computed for Size, B/M and D/E as follows:

- Size breaking value is the median for market cap at the end of June t
- B/M breaking values are the 30th and 70th quantiles for B/M at the end of December t-1
- D/E breaking values are the 30th and 70th quantiles for D/E at the end of December t-1

Afterwards, at the end of June, stocks are allocated independently to two Size, three B/M and D/E groups according to their breaking values computed before. The intersection of the sorts produces 6 Size-B/M and 6 Size-D/E portfolios. Portfolios are rebalanced and reconstituted each year at the end of June, specifically, constituents and their weights remain constant for twelve months. Weights correspond to market cap.

The SMB factor (Small minus Big) means to mimic the risk factor in returns related to size is the difference, each month, between the simple average of the returns on six small-stock portfolios and the simple average of the returns on six big-stock portfolios (Fama & French, 2015) (see Appendix 17–Appendix 18).

$$SMB_{2x3} = \frac{(SHV + SMV + SLV + SHL + SML + SLL)}{6} - \frac{(BHV + BMV + BLV + BHL + BML + BLL)}{6}$$

The HML factor (High minus Low) means to mimic the risk factor in returns related to book-to-market equity is the difference, each month, between the simple average of the returns on two high-book-to-market portfolios and the simple average of the returns on two low-book-to-market portfolios (Fama & French, 2015) (see Appendix 17–Appendix 18).

$$HML_{2x3} = \frac{(SHV + BHV)}{2} - \frac{(SLV + BLV)}{2}$$

The LMU factor (Levered minus Unlevered) means to mimic the risk factor in returns related to company financial leverage is the difference, each month, between the simple average of the returns on two high-leverage portfolios and the simple average of the returns on two low-leverage portfolios (Fama & French, 2015) (see Appendix 17–Appendix 18).

$$LMU_{2x3} = \frac{(SHL + BHL)}{2} - \frac{(SLL + BLL)}{2}$$

It is important to note that the HML and LMU factors derived with the 2x3 method do not take into account 40% of the stocks in our sample. Indeed, the middle portfolios are omitted in their construction.

5.1.4. 2x2 sort

These portfolios will serve to build the four factors that will be used as independent variables for the regression or the RHS portfolios. Portfolio construction is largely inspired by Fama and French (2015) methodology outlined in their paper.

At the end of June, breaking values are computed for Size, B/M and D/E as follows:

- Size breaking value is the median for market cap at the end of June t
- B/M breaking value is the median for B/M at the end of December t-1
- D/E breaking value is the median for D/E at the end of December t-1

Afterwards, at the end of June, stocks are allocated independently to two Size, B/M and D/E groups according to their breaking values computed before. The intersection of the sorts produces 4 Size-B/M and 4 Size-D/E portfolios. Portfolios are rebalanced and reconstituted each year at the end of June, specifically, constituents and their weights remain constant for twelve months. Weights correspond to market cap.

The SMB factor means to mimic the risk factor in returns related to size is the difference, each month, between the simple average of the returns on four small-stock portfolios and the simple average of the returns on four big-stock portfolios (Fama & French, 2015) (see Appendix 19–Appendix 20).

$$SMB_{2x2} = \frac{(SHV + SLV + SHL + SLL)}{4} - \frac{(BHV + BLV + BHL + BLL)}{4}$$

The HML factor means to mimic the risk factor in returns related to book-to-market equity is the difference, each month, between the simple average of the returns on two high-book-to-market portfolios and the simple average of the returns on two low-book-to-market portfolios (Fama & French, 2015) (see Appendix 19–Appendix 20).

$$HML_{2x2} = \frac{(SHV + BHV)}{2} - \frac{(SLV + BLV)}{2}$$

The LMU factor means to mimic the risk factor in returns related to company financial leverage is the difference, each month, between the simple average of the returns on two high-leverage portfolios and the simple average of the returns on two low-leverage portfolios (Fama & French, 2015) (see Appendix 19–Appendix 20).

$$LMU_{2x2} = \frac{(SHL + BHL)}{2} - \frac{(SLL + BLL)}{2}$$

5.1.5. 2x2x2 sort

These portfolios will serve to build the four factors that will be used as independent variables for the regression or RHS portfolios. Portfolio construction is largely inspired by Fama and French (2015) methodology outlined in their paper.

At the end of June, breaking values are computed for Size, B/M and D/E as follows:

- Size breaking value is the median for market cap at the end of June t
- B/M breaking value is the median for B/M at the end of December t-1
- D/E breaking value is the median for D/E at the end of December t-1

Afterwards, at the end of June, stocks are allocated independently to two Size, B/M and D/E groups according to their breaking values computed before. The intersection of the sorts produces 8 Size-B/M-D/E portfolios. Portfolios are rebalanced and reconstituted each year at the end of June, specifically, constituents and their weights remain constant for twelve months. Weights correspond to market cap.

The SMB factor means to mimic the risk factor in returns related to size is the difference, each month, between the simple average of the returns on four small-stock portfolios and the simple average of the returns on four big-stock portfolios (Fama & French, 2015) (see Appendix 21).

$$SMB_{2x2x2} = \frac{(SHVHL + SHVLL + SLVHL + SLVLL)}{4} - \frac{(BHVHL + BHVLL + BLVHL + BLVLL)}{4}$$

The HML factor means to mimic the risk factor in returns related to book-to-market equity is the difference, each month, between the simple average of the returns on four high-book-to-market portfolios and the simple average of the returns on four low-book-to-market portfolios (Fama & French, 2015) (see Appendix 21).

$$HML_{2x2x2} = \frac{(BHVHL + BHVLL + SHVHL + SHVLL)}{4} - \frac{(BLVHL + BLVLL + SLVHL + SLVLL)}{4}$$

The LMU factor means to mimic the risk factor in returns related to company financial leverage is the difference, each month, between the simple average of the returns on four high-leverage portfolios and the simple average of the returns on four low-leverage portfolios (Fama & French, 2015) (see Appendix 21).

$$LMU_{2x2x2} = \frac{(BHVHL + BLVHL + SHVHL + SLVHL)}{4} - \frac{(BHVLL + BLVLL + SHVLL + SLVLL)}{4}$$

In the end we have five models (CAPM, SMB, three-factor model, four-factor model and SMB-LMU), five sets of LHS portfolios and three sets of RHS factors.

5.1.6. Summary statistics

All of these steps will produce a plethora of portfolios with return data sets to process. We will employ various statistical indicators to assist us in this task. In order to test the H1 hypothesis: *Expected common stock returns are negatively related to the financial leverage*, we will calculate the average monthly excess returns for all the portfolios. We will then classify them to check the relationship between the average excess return and the different ratios used to build the portfolios. We will also use a t-test to check if the factors expected returns are statistically different from zero.

5.2. Regressions

Our model consists in regressing a multitude of portfolios on the different factors constructed. This is the regression equation from the leverage augmented Fama and French three-factor model:

$$R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + l_i LMU_t + e_{it}$$

Where R_{it} is the return on security or portfolio i for period t
R_{Ft} is the risk-free return
a_i is the 'alpha' or return in excess of that expected given the portfolio's level of systematic risk (assuming the four factors capture all systematic risk)
b_i is the sensitivity to the market factor or market beta
R_{Mt} is the return on the VW market portfolio
s_i is the sensitivity to the SMB factor or size beta
SMB_t is the return on a diversified portfolio of small stocks minus the return on a diversified portfolio of big stocks
h_i is the sensitivity to the HML factor or value beta
HML_t is the difference between the returns on diversified portfolios of high and low book-to-market stocks
l_i is the sensitivity to the LMU factor or leverage beta
LMU_t is the difference between the returns on diversified portfolios of high and low debt-to-equity stocks

If the model is accurate, thus if it proxies for common risk factors in returns, the factor loadings, b_i, s_i, h_i, and l_i, capture all variation in expected returns, the intercept a_i is zero for all securities and portfolios (Fama & French, 2015). The intercept is, by definition, the excess return for a portfolio or stock if all the slopes in the regression are equal to zero; therefore, if it is zero, it means that the factors explain all the returns and that there is no residual return, also called alpha.

To evaluate the model performance, we test how well the different sets of factors, constructed using all RHS sorts, explain excess return on the LHS portfolios. This test corresponds to the GRS statistic of Gibbson et al. (1989) that tests if the intercept of all combinations of LHS portfolios is different from zero. Other techniques are outlined in the Fama and French (2015) article such as comparing average absolute value of all the regressions' intercepts $A|\alpha_i|$, and average absolute value of the intercepts over the average absolute value of the average on portfolio i minus the average of the portfolio returns $A|\alpha_i| / A|\bar{r}_i|$, between the different models. We will use these same indicators in the analysis of results, to support the interpretation of the GRS test. The latter equation is used to determine the percentage of returns non-explained by the model. The outcomes will allow us to test the H2 hypothesis: *A factor model including financial leverage is a better proxy for common risk factors in returns than the original Fama & French three-factor model.* We next examine regression details such as intercepts, slopes and test statistic that will help us to draw conclusions from the regressions. Having analysed the overall performance of our model, this step will lead us to focus on its performance in detail. It will be interesting to see if our model struggles to work on different types of portfolios; putting our results in tabular form will make it easier.

Then, to see if there is any relationship between the four different factors, we will regress each factor on the other three and analyse the data produced. This process will help us to test H3 hypothesis: *The Fama and French three-factor model does not capture entirely financial leverage risk in its factors*, that is to say, to determine whether the leverage factor is redundant in the Fama and French three-factor model. Significant coefficient slopes between two factors may indicate a linkage that needs to be further analysed; two variables fluctuating in the same way could point to a redundancy.

Finally, we will focus on the performance of the model over time, in order to answer the final hypothesis H4: *The importance of financial leverage as a risk factor does not depend on the business cycle.* To do this, we will calculate the GRS statistic (Gibbson et al., 1989) and the different intercept test statistics from a sample period of five years, rolled every month. We will apply this method to both models, the Fama and French original three-factor model, and the one augmented with a leverage factor, and compare the data obtained from both models to see if their performance is not affected by the economic conjuncture. We will also relate these statistics to two macroeconomic indicators: equity market volatility measured as the CBOE volatility Index (VIX), and euro area GDP growth rate. We will calculate the cross-correlation of the performance of our model with these two indicators to find a significant pattern.

6. RESULTS

Here we describe objectively the full results of our research and we deal with the previously mentioned hypothesis. The approach to analysis is largely inspired by that used by Fama and French (2015). As our methodology is very similar, it is reasonable to assume that we should follow the structure of their reasoning in order to compare our findings.

6.1. Summary statistics

Let's examine how the performance of our portfolios reacts when we sort them according to the variables that will be the focus of our analysis (see Figure 3).

The first thing that comes to attention is the absence of a size effect in our study sample; both panels show that bigger portfolios imply higher excess return on average. However, Panel A (see Figure 3) confirms that higher B/M portfolios gives higher excess return on average, the value effect, with small portfolios having more sensitivity to B/M changes; and lower B/M portfolios having more sensitivity to size changes. The big cap deep value portfolio outperforms other portfolios with an average monthly excess return of 0.76%.

Consistent with George and Hwang (2010), Panel B (see Figure 3) show that low leverage portfolios give a higher value-weighted average excess return than those with high leverage – but decrease a little bit with the lowest leverage category in our case. Our results contradict Modigliani and Miller (1958), and Bhandari (1988), who claimed the existence of a positive risk premium for equities with high financial leverage. We also see that smaller portfolios excess return is more sensitive to D/E changes. The big cap unlevered portfolio outperforms other portfolios with an average monthly excess return of 0.54%.

Figure 4 is in line with what we have seen in the previous table: big cap portfolios outperform small cap portfolios on average. Furthermore, for small cap portfolios, lower D/E portfolios and higher B/M implies higher excess return on average. For big cap portfolios, higher B/M portfolios imply higher excess return on average and lower B/M portfolios are more sensitive to D/E changes.

We have calculated the value-weighted average ratios obtained for each portfolio; these will allow us to better understand their structure (see Figure 5). We found interesting irregularities in our sample: small high value portfolios have an extreme average B/M value of 23, while others are much lower. Furthermore, the same anomaly appears with small highly leveraged portfolios, with an average B/M value of 29. We checked our sample for errors but could not find anything. Besides, smaller stocks have a higher D/E and B/M ratio on average, and growth portfolios (low B/M) have a higher leverage. This is consistent with what Obreja (2013) found in his paper, that is, companies with low book-to-market ratio face either high operating leverage or financial leverage.

Figure 3: Average monthly percent excess returns for portfolios formed on Size and B/M, Size and D/E

June 1989 – May 2021, 354 months. At the end of each June, stocks are allocated to five Size groups (Small to Big) using market cap breakpoints. Stocks are allocated independently to five B/M groups (Low to High), again using breakpoints. The intersections of the two sorts produce 25 value-weighted Size-B/M portfolios. In the sort for June of year t, B is book equity at the end of December ending in year t-1 and M is market cap at the end of December of year t-1. The Size-D/E portfolios are formed in the same way, except that the second sort variable is debt-to-equity. For D/E, in the sort for June of year t, D is book debt at the end of December ending in year t-1. The table shows averages of monthly returns in excess of the three-month Bubill rate.

	Low	2	3	4	High
Panel A: Size-B/M portfolios					
Small	-0,57	-0,10	-0,13	0,14	0,63
2	-0,42	-0,22	-0,01	0,10	0,52
3	-0,41	-0,14	0,00	0,25	0,49
4	-0,09	0,21	0,36	0,56	0,53
Big	0,28	0,52	0,43	0,52	0,76
Panel B: Size-D/E portfolios					
Small	0,32	0,35	0,17	0,11	-0,44
2	0,19	0,09	0,08	-0,03	-0,23
3	0,00	0,21	0,17	0,06	-0,26
4	0,31	0,39	0,31	0,30	0,13
Big	0,54	0,50	0,44	0,44	0,29

Source: Based on the author's calculation run in Python and R program using data from DataStream, and Fama and French (2015) for the figure template.

Figure 4: Average monthly percent excess returns for portfolios formed on Size, B/M and D/E

June 1989 – May 2021, 384 months. At the end of each June, stocks are allocated to two Size groups (Small to Big) using market cap median breakpoints. Stocks in each Size group are allocated independently to four B/M groups (Low to High) and four D/E groups (Low to High), again using breakpoints. The intersections of the three sorts produce 32 value-weighted Size-B/M-D/E portfolios. In the sort for June of year t, B is book equity at the end of December ending in year t-1, D is book debt at the end of December ending in year t-1 and M is market cap at the end of December of year t-1. The table shows averages of monthly returns in excess of the three-month Bubill rate.

	Small				Big			
Panel A: Size-B/M-D/E								
B/M →	Low	2	3	High	Low	2	3	High
Low D/E	-0,21	-0,09	0,26	0,65	0,35	0,54	0,46	0,71
2	-0,28	-0,02	0,37	0,47	0,44	0,61	0,59	0,69
3	-0,45	-0,15	0,06	0,46	0,31	0,41	0,48	0,58
High D/E	-0,72	-0,36	-0,11	0,29	0,10	0,30	0,34	0,66

Source: Based on the author's calculation run in Python and R program using data from DataStream, and Fama and French (2015) for the figure template.

Figure 5: Time-series averages of the ratios

June 1989 – May 2021, 384 months. Time-series averages of book-to-market ratios (B/M), and debt-to-equity (D/E) for portfolios formed on (i) Size and B/M, (ii) Size and D/E, and (iii) Size, B/M and D/E. In the sort for June of year t, B is book equity at the end December ending of year t-1 and M is market cap at the end of December of year t-1. For D/E, in the sort for June of year t, D is book debt at the end of December ending in year t-1. Each of the ratios for a portfolio for a given year is the value-weighted average (market cap weights) of the ratios for the firms in the portfolio.

	B/M					D/E				
25 Size-B/M portfolios										
B/M →	Low	2	3	4	High	Low	2	3	4	High
Small	0,22	0,46	0,70	1,06	23,02	2,78	1,49	1,11	0,99	1,03
2	0,21	0,46	0,69	1,04	5,04	3,18	1,05	0,87	0,84	0,89
3	0,22	0,46	0,69	1,04	3,71	2,28	0,90	0,83	0,86	1,06
4	0,22	0,45	0,69	1,03	3,15	1,55	0,86	0,84	0,99	1,15
Big	0,22	0,46	0,69	1,00	2,07	1,68	0,94	0,95	0,91	1,05
25 Size-D/E portfolios										
D/E →	Low	2	3	4	High	Low	2	3	4	High
Small	2,61	5,03	11,09	28,57	10,99	0,07	0,30	0,65	1,15	5,32
2	1,14	1,53	1,99	2,45	1,94	0,07	0,31	0,65	1,14	5,20
3	0,88	1,09	1,48	1,32	1,18	0,07	0,31	0,64	1,14	4,39
4	0,79	0,80	0,96	1,18	0,84	0,07	0,31	0,65	1,14	3,87
Big	0,43	0,56	0,66	0,74	0,56	0,07	0,32	0,66	1,13	3,66

	B/M					D/E		
32 Size-B/M-D/E portfolios								
				Small				
B/M →	Low	2	3	High	Low	2	3	High
Low D/E	0,26	0,54	0,89	2,99	0,09	0,10	0,09	0,10
2	0,26	0,54	0,89	3,85	0,42	0,42	0,42	0,43
3	0,27	0,55	0,90	6,74	0,93	0,93	0,93	0,91
High D/E	0,21	0,54	0,89	6,57	8,53	3,10	2,71	2,81
				Big				
B/M →	Low	2	3	High	Low	2	3	High
Low D/E	0,24	0,52	0,87	2,08	0,08	0,10	0,11	0,11
2	0,27	0,54	0,87	1,99	0,42	0,45	0,46	0,45
3	0,27	0,54	0,87	2,12	0,93	0,95	0,98	0,98
High D/E	0,25	0,53	0,86	2,00	5,12	2,02	2,12	2,23

Source: Based on the author’s calculation run in Python and R program using data from DataStream, and Fama and French (2015) for the figure template.

Now let’s look at the summary statistics for the different factor returns studied, from the standpoint of each different sorting (see Figure 6).

Panel A (see Figure 6) shows that for all sorts, the market factor is positive, but the most volatile. Moreover, SMB factor is the worst on a risk-adjusted basis, while HML factor is the best. LMU factor is negative and the most affected by changes in the sort. The 2x3 sort focus more on the outliers than the other sorts, and thus gives more extreme returns (Fama & French, 2015). The t-statistics help us to validate our hypothesis on the relationship between leverage and expected return. It appears that for the three different types, the expected return of the LMU factor is negatively significant with a high degree of confidence. However, we also obtained a surprising result. The market factor is not statistically significant. This is caused by a rather high standard error, and thus a high variance in our sample. This result encourages us to reflect on the choice of the benchmark used to construct this factor.

Panel B (see Figure 6) tells us optimistically that different versions of the factors do not change significantly the series of data. Nonetheless, the 2x2x2 sort have a greater impact on the LMU factor compared to the 2x2 sort, with the lowest correlation, 0.88. Not surprisingly, the SMB factor has the highest correlations among factors as all three forms of the factor use all the stocks in the sample (Fama & French, 2015).

Panel C (see Figure 6) shows us something of interest: HML and LMU factors have a positive correlation with an average of 0.60 among the different sorts. This is evidence that these two factors have components in common, although one does not fully explain the other. In addition, SMB factor has a negative correlation with the market factor, -0.45.

Figure 6: Summary statistics for monthly factor percent returns

June 1989 - May 2021, 384 months. Rm-Rf is the value-weighted return on the market portfolio of all sample stocks minus the three-month Bubill rate. Panel A of the table shows average monthly returns (Mean), the standard deviations of monthly returns (Std dev.), and the test for the null hypothesis that the expected value of the sample is equal to zero (t-Statistic). Panel B shows the correlations of the same factor from different sorts and Panel C shows the correlations for each set of factors.

Panel A: Averages and standard deviations for monthly returns

	2 x 3 Factors				2 x 2 Factors				2 x 2 x 2 Factors			
	$R_m - R_f$	*SMB*	*HML*	*LMU*	$R_m - R_f$	*SMB*	*HML*	*LMU*	$R_m - R_f$	*SMB*	*HML*	*LMU*
Mean	0,30	-0,40	0,53	-0,27	0,30	-0,40	0,35	-0,18	0,30	-0,46	0,37	-0,20
Std dev.	5,16	2,14	2,29	1,82	5,16	2,21	1,50	1,21	5,16	2,19	1,48	1,15
t-Statistic	1,16	-3,63	4,56	-2,87	1,16	-3,51	4,56	-2,86	1,16	-4,15	4,85	-3,35

Panel B: Correlations between different versions of the same factor

	SMB			HML			LMU		
	2 x 3	2 x 2	2 x 2 x 2	2 x 3	2 x 2	2 x 2 x 2	2 x 3	2 x 2	2 x 2 x 2
2 x 3	1,00	0,99	0,98	1,00	0,91	0,90	1,00	0,91	0,88
2 x 2	0,99	1,00	0,99	0,91	1,00	0,99	0,91	1,00	0,97
2 x 2 x 2	0,98	0,99	1,00	0,90	0,99	1,00	0,88	0,97	1,00

Panel C: Correlations between different factors

	2 x 3 Factors				2 x 2 Factors				2 x 2 x 2 Factors			
	$R_m - R_f$	*SMB*	*HML*	*LMU*	$R_m - R_f$	*SMB*	*HML*	*LMU*	$R_m - R_f$	*SMB*	*HML*	*LMU*
Rm - Rf	1,00	-0,46	0,08	0,13	1,00	-0,45	0,03	0,14	1,00	-0,43	0,01	0,11
SMB	-0,46	1,00	0,01	-0,09	-0,45	1,00	-0,04	-0,10	-0,43	1,00	-0,10	-0,07
HML	0,08	0,01	1,00	0,63	0,03	-0,04	1,00	0,63	0,01	-0,10	1,00	0,55
LMU	0,13	-0,09	0,63	1,00	0,14	-0,10	0,63	1,00	0,11	-0,07	0,55	1,00

Source: Based on the author's calculation run in Python and R program using data from DataStream, and Fama and French (2015) for the figure template.

6.2. Model performance

Now that we have examined our sample and factors in detail, we can move on to analyse the overall performance of our model. GRS statistic is the central performance indicator; however, we will add two others mentioned above (see Chapter 5.2 for how these indicators are constructed), which will support our findings (see Figure 7). Let us recall that GRS statistic is used to test if the intercepts of all LHS portfolio excess returns are different from zero; if it is the case, the model does not explain very well the variation in returns (Gibbson et al., 1989). Here we will not find the definitive asset pricing model that will explain clearly the return that can be expected from a stock; what we are particularly interested in is whether adding the leverage factor gives us a more accurate model.

In Panel A of Figure 7, all tests are easily rejected; thus all models are incomplete descriptions of expected returns. However, the four-factor model produces lower GRS statistics and higher p-value than the other models for every sort; we noticed the same for the average absolute value intercepts. Also, for all sorts, the proportion of expected returns left unexplained is the lowest for the new models – i.e. going from an average of 68% left unexplained for the Fama and French three-factor model (1992, 1993), to an average of 49%. The SMB LMU model does not help to proxy better the expected return, which suggests that the HML and LMU factors are essential and complementary.

Panel B of Figure 7 shows the performance results for LHS portfolios controlled for size and D/E. Not surprisingly, the GRS tests are not rejected for all four-factor models with a 99% degree of confidence. Yet, the proportion of expected returns left unexplained is larger when using these portfolios, compared to these used in Panel A (size-B/M portfolios).

Panel C of Figure 7 shows that the GRS test rejects all models, except for the four-factor model again. For every sort, the average absolute intercept, and the proportions of expected returns left unexplained are the lowest when we add the LMU factor.

The outcomes we have obtained are quite reassuring. We have shown that our new model does perform better than the original one, regardless of how the portfolios are formed on both sides of the equation; and thus, the Fama and French (1992, 1993) three factors are not sufficient to explain returns in euro area stocks. The greatest improvement between the models comes from the LHS 25 Size-D/E portfolios, with a difference in the GRS statistic of 2.35, 2.08, and 1.82. Then we have the 32 Size-B/M-D/E portfolios that give us an improvement of 1.80, 1.80, and 1.58. Finally, the slightest difference in performance but still a noticeable difference: the 25 Size-B/M portfolios with 0.63, 0.23, and the lowest improvement of them all 0.18. Figure 7 also tells us that simply adding the LMU factor and omitting the HML factor is not enough to obtain a good model. Moreover, remember that we found that these same two factors are positively correlated to a significant extent (see Chapter 6.1), so we believe that these two variables would share some of their explanatory power. It is therefore appropriate to investigate this problem in depth. Let us observe the results obtained by using three factors in regressions to explain average returns on the fourth (see Figure 8).

It looks like LMU and HML factors have a strong positive relationship in every sort; when we regress the variable HML on the others, we obtain a slope coefficient of 0.77 for LMU. Conversely, when we regress the variable LMU on the others, we obtain an average slope coefficient of 0.48 for HML. These statistics support our view that these two variables have something in common. We also found that the market and SMB factors have a significant negative relationship in every sort, which is consistent with the negative correlation we got in Figure 6; this could be explained because the benchmark used to construct the first factor is heavily represented by big capitalisation stocks. Furthermore, we noticed that the 2x2x2 sort weakens the variance in returns explained by the factors, represented by their R^2. It seems that this way of separating the portfolios in order to construct the factors reduces the explanatory power of the factors on the others. Furthermore, the intercept is statistically significant for all regressions, except when the market factor is used as the dependent variable. It may be evidence that the three other factors explain well the market factor. Indeed, if the other factors return was zero, the market factor would not give a statistically significant return. In addition, with the same regression, the LMU factor is significant for both 2x2 and 2x2x2 sorts, but not for the 2x3 sort; it has maybe something to do with the fact that the 2x3 sort does not include all stocks from the sample in the HML and LMU factors. Besides, we obtained a negative alpha for the SMB and LMU factors, which strengthens our beliefs concerning the hypothesis that financial leverage and expected return are negatively linked.

We now come to our fourth hypothesis: is the performance of our model affected by economic cycles? To test this conjecture, we compared the performance over time of the four-factor model with the three-factor model, using the GRS statistic. Each data point is calculated over a period of five years, rolled successively until the end. We have replicated this method for the three types of LHS portfolios, and plotted those using graphs (see Figure 9–Figure 11).

The results are rather positive; for all three sorts, the four-factor model almost always outperform the three-factor model no matter the macroeconomic conditions (see Figure 9–Figure 11). Some periods are marked by a significant increase in the GRS statistic; until 2002, the performance of both models is rather stable, then falls until 2008. The following period seems to be characterised by volatility in performance. We wanted to understand what could lead these curves to fluctuate in such a way, so we related them to periods of recessions and economic downturns in the euro area; they are represented by grey bars on the charts. Yet, these do not allow us to draw any conclusions. To go further, we also calculated the cross-correlation of the GRS statistic with macroeconomic indicators, such as volatility on the financial markets (CBOE Volatility Index) and GDP growth on the territory under study: we did not find anything conclusive on this issue, so we saw little need to include these results in our work. However, no results do not mean any findings: identifying the factors that influence the performance of asset pricing models is another topic that could be the subject of another study.

Figure 7: Summary statistics for tests of three-, four-factor models

June 1989 – May 2021, 384 months. The table tests the ability of three-, four-factor models to explain monthly excess returns on 25 Size-B/M portfolios (Panel A), 25 Size-D/E portfolios (Panel B), and 32 Size-B/M-D/E portfolios (Panel C). For each set of 25 or 32 regressions, the table shows the factors that augment Rm-Rf in the regression model, the GRS statistic testing whether the expected values of all 25 or 32 intercept estimates are zero, the average absolute value of the intercepts $A|\alpha_i|$, the average absolute value of the intercept over the average absolute value of r_i , which is the average return on portfolio i minus the average of the portfolio returns, $A|\alpha_i| / A|\bar{r}_i|$.

	2 x 3 Factors				2 x 2 Factors				2 x 2 x 2 Factors			
	GRS	*p-value*	$A\|\alpha_i\|$	$A\|\alpha_i\| / A\|\bar{r}_i\|$	*GRS*	*p-value*	$A\|\alpha_i\|$	$A\|\alpha_i\| / A\|\bar{r}_i\|$	*GRS*	*p-value*	$A\|\alpha_i\|$	$A\|\alpha_i\| / A\|\bar{r}_i\|$
Panel A: 25 Size-B/M portfolios												
SMB HML	2,70	0,000	0,186	0,60	2,88	0,000	0,214	0,69	2,89	0,000	0,228	0,74
SMB HML LMU	2,07	0,002	0,137	0,44	2,65	0,000	0,157	0,51	2,71	0,000	0,160	0,51
SMB LMU	4,96	0,000	0,372	1,20	5,51	0,000	0,367	1,18	5,13	0,000	0,380	1,22
Panel B: 25 Size-D/E portfolios												
SMB HML	4,05	0,000	0,271	1,49	3,74	0,000	0,287	1,57	3,65	0,000	0,301	1,65
SMB HML LMU	1,70	0,020	0,152	0,83	1,66	0,026	0,155	0,85	1,83	0,010	0,172	0,94
SMB LMU	2,05	0,002	0,276	1,51	2,14	0,001	0,256	1,40	2,24	0,001	0,303	1,66
Panel C: 32 Size-B/M-D/E portfolios												
SMB HML	3,06	0,000	0,243	0,80	3,20	0,000	0,261	0,86	3,10	0,000	0,266	0,87
SMB HML LMU	1,26	0,162	0,135	0,44	1,40	0,078	0,139	0,45	1,51	0,041	0,153	0,50
SMB LMU	3,46	0,000	0,363	1,19	3,51	0,000	0,357	1,17	3,32	0,000	0,366	1,20

Source: Based on the author’s calculation run in Python and R program using data from DataStream, and Fama and French (2015) for the figure template.

Figure 8: Using three factors in regressions to explain average returns on the fourth

June 1989 – May 2021, 384 months. Rm-Rf is the value-weighted return on the market portfolio of all sample stocks minus the three-month Bubill rate; SMB (small minus big) is the size factor; HML (high minus low B/M) is the value factor, and LMU (high minus low leverage). The 2 x 3 factors are constructed using separate sorts of stocks into two Size groups and three B/M groups (HML), or three D/E groups (LMU). The 2 x 2 factors use the same approach except the second sort for each factor produces two rather than three portfolios. Each factor from the 2 x 3 and 2 x 2 sorts uses 2 x 3 = 6 or 2 x 2 = 4 portfolios to control for size and one other variable (B/M, D/E). The 2 x 2 x 2 factors use the 2 x 2 x 2 = 8 portfolios to jointly control for Size, B/M, and D/E. Int is the regression intercept

	Int	$R_m - R_f$	*SMB*	*HML*	*LMU*	*R²*
2 x 3 Factors						
$R_m - R_f$						
Coef	-0,15		-1,10	0,12	0,15	0,22
Std err.	0,30		0,11	0,13	0,17	
t-Statistic	-0,59		-9,99	0,89	0,91	
p-value	0,56		0,00	0,37	0,36	
SMB						
Coef	-0,43	-0,19		0,10	-0,12	0,22
Std err.	0,10	0,02		0,05	0,07	
t-Statistic	-4,03	-9,99		1,93	-1,80	
p-value	0,00	0,00		0,06	0,07	
HML						
Coef	0,78	0,02	0,09		0,79	0,40
Std err.	0,10	0,02	0,05		0,05	
t-Statistic	8,29	0,89	1,93		15,74	
p-value	0,00	0,37	0,06		0,00	
LMU						
Coef	-0,56	0,01	-0,07	0,50		0,41
Std err.	0,10	0,02	0,04	0,03		
t-Statistic	-7,49	0,91	-1,80	15,74		
p-value	0,00	0,36	0,07	0,00		

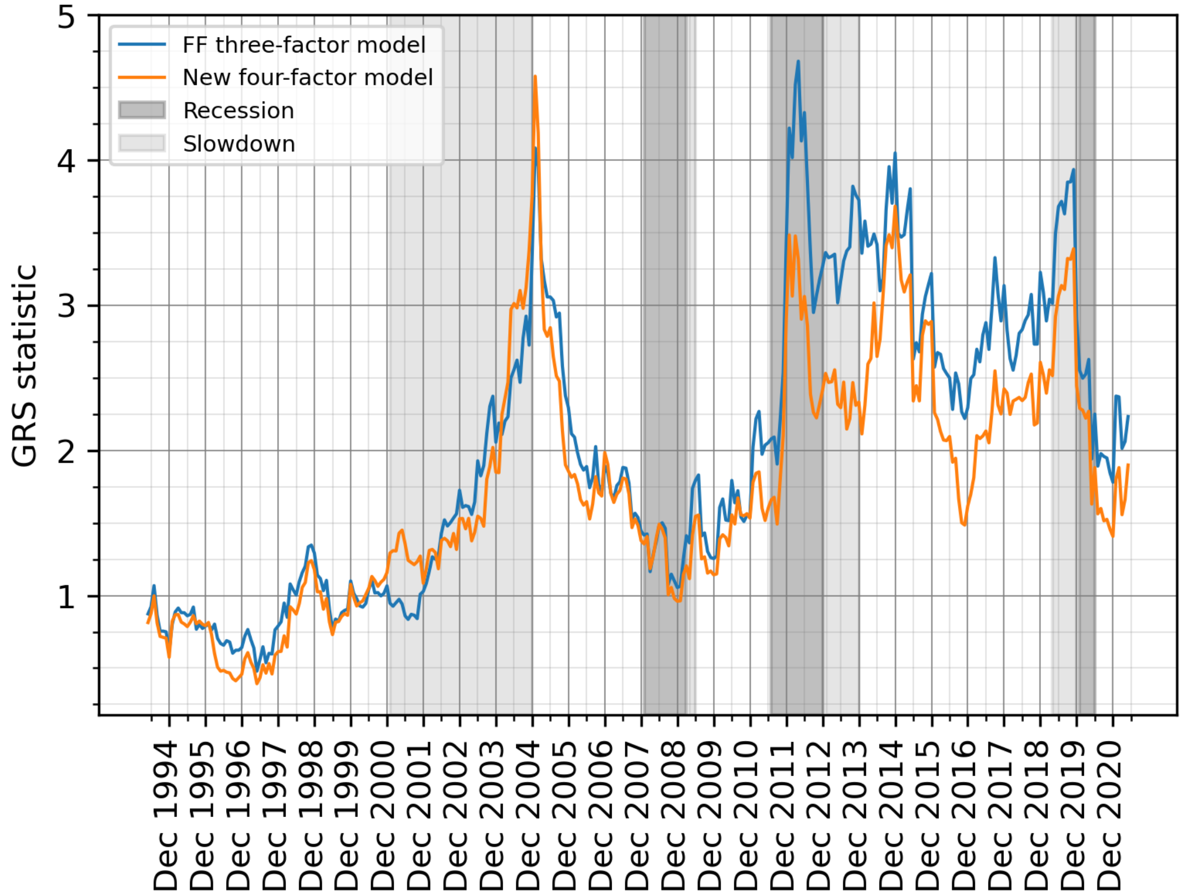
	Int	$R_m - R_f$	*SMB*	*HML*	*LMU*	*R²*
2 x 2 Factors						
$R_m - R_f$						
Coef	0,11		-1,03	-0,29	0,64	0,22
Std err.	0,30		0,11	0,20	0,25	
t-Statistic	0,43		-9,67	-1,42	2,58	
p-value	0,67		0,00	0,16	0,01	
SMB						
Coef	-0,35	-0,19		-0,01	-0,06	0,21
Std err.	0,10	0,02		0,09	0,11	
t-Statistic	-3,12	-9,67		-0,08	-0,58	
p-value	0,00	0,00		0,94	0,57	
HML						
Coef	0,50	-0,02	0,00		0,79	0,40
Std err.	0,10	0,01	0,03		0,05	
t-Statistic	8,10	-1,42	-0,08		15,97	
p-value	0,00	0,16	0,94		0,00	
LMU						
Coef	-0,37	0,03	-0,01	0,51		0,42
Std err.	0,00	0,01	0,02	0,03		
t-Statistic	-7,45	2,58	-0,58	15,97		
p-value	0,00	0,01	0,57	0,00		

	Int	$R_m - R_f$	*SMB*	*HML*	*LMU*	*R²*
2 x 2 x 2 Factors						
$R_m - R_f$						
Coef	0,08		-1,02	-0,35	0,59	0,20
Std err.	0,30		0,11	0,19	0,25	
t-Statistic	0,29		-9,35	-1,83	2,36	
p-value	0,77		0,00	0,07	0,02	
SMB						
Coef	-0,33	-0,18		-0,17	0,07	0,20
Std err.	0,10	0,02		0,08	0,11	
t-Statistic	-3,01	-9,35		-2,05	0,66	
p-value	0,00	0,00		0,04	0,51	
HML						
Coef	0,49	-0,02	-0,07		0,72	0,32
Std err.	0,10	0,01	0,03		0,06	
t-Statistic	7,44	-1,83	-2,05		13,04	
p-value	0,00	0,07	0,04		0,00	
LMU						
Coef	-0,35	0,02	0,02	0,43		0,32
Std err.	0,10	0,01	0,03	0,03		
t-Statistic	-6,96	2,36	0,66	13,04		
p-value	0,00	0,02	0,51	0,00		

Source: Based on the author's calculation run in Python and R program using data from DataStream, and Fama and French (2015) for the figure template.

Figure 9: Performance comparison between the three-, four-factor models over time using the 25 value-weighted Size-B/M portfolios

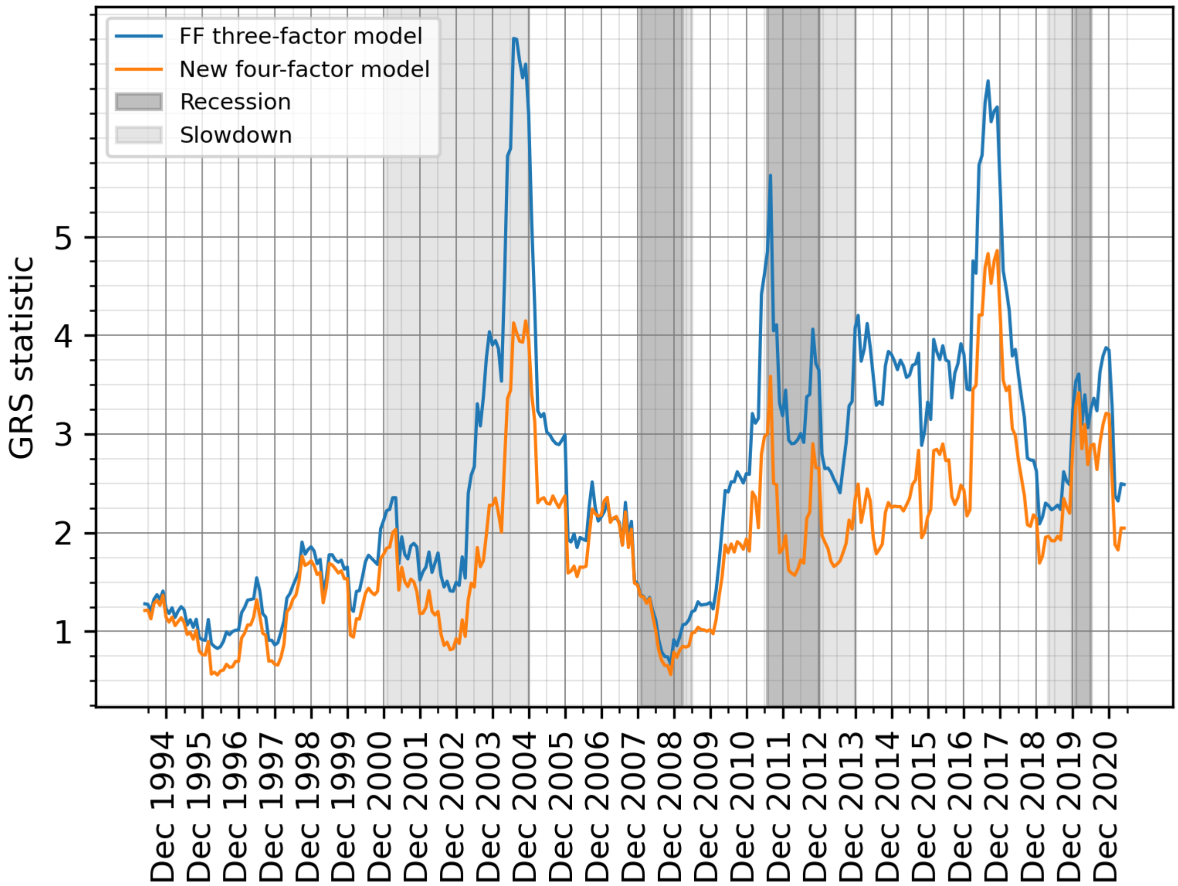
June 1989 – May 2021, 384 months. The GRS statistic is calculated from a sample period of five years, rolled every month until May 2021. Each date on the x-axis corresponds to the last observation of the rolling window. Regressions are made for 25 value-weighted Size-B/M portfolios. The RHS variables are constructed using independent 2 x 3 sorts on size and each of B/M and D/E. Recession and slowdown periods are for the euro area.



Source: Based on the author’s calculation run in Python and R program using data from DataStream, and Eurostat (2022a) for recession and slowdown periods.

Figure 10: Performance comparison between the three-, four-factor models over time using the 25 value-weighted Size-D/E portfolios

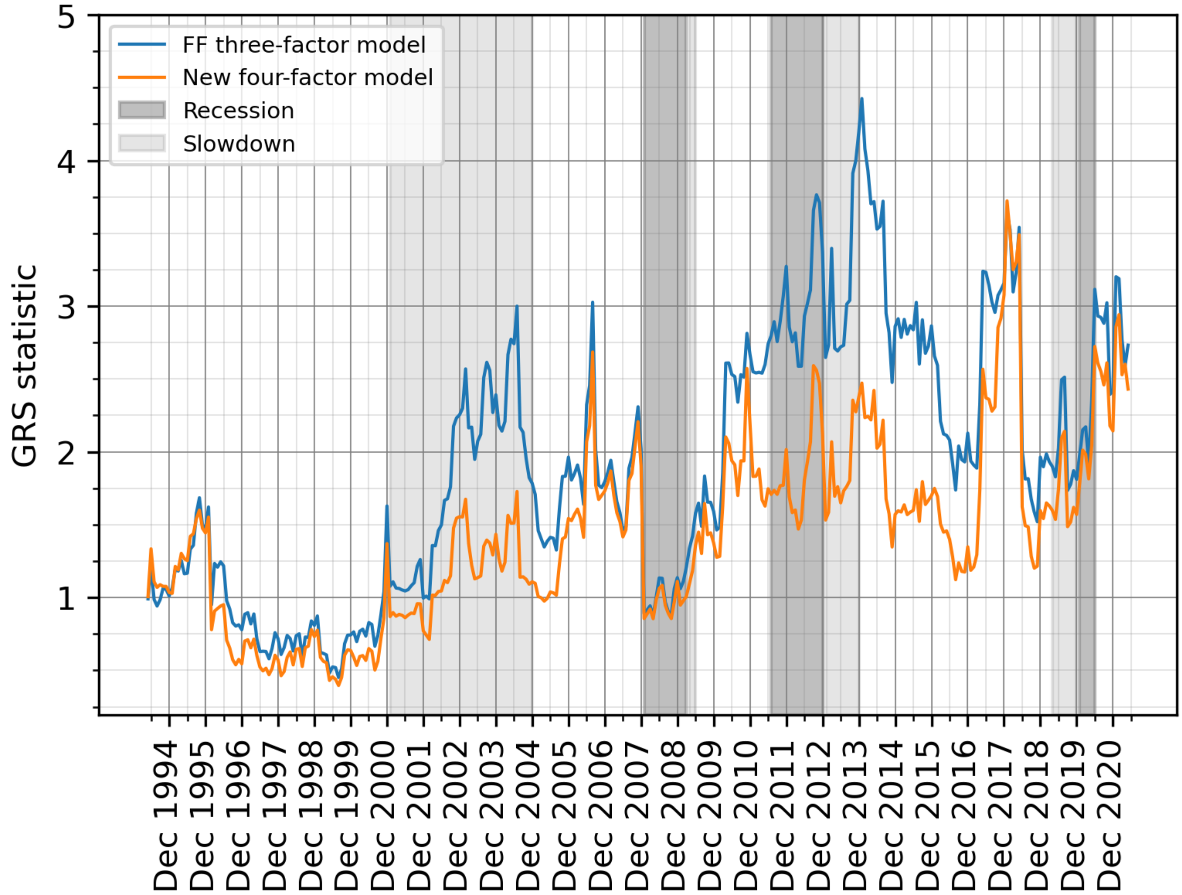
June 1989 – May 2021, 384 months. The GRS statistic is calculated from a sample period of five years, rolled every month until May 2021. Each date on the x-axis corresponds to the last observation of the rolling window. Regressions are made for 25 value-weighted Size-D/E portfolios. The RHS variables are constructed using independent 2 x 3 sorts on size and each of B/M and D/E. Recession and slowdown periods are for the euro area.



Source: Based on the author's calculation run in Python and R program using data from DataStream, and Eurostat (2022a) for recession and slowdown periods.

Figure 11: Performance comparison between the three-, four-factor models over time using the 32 value-weighted Size-B/M-D/E portfolios

June 1989 – May 2021, 384 months. The GRS statistic is calculated from a sample period of five years, rolled every month until May 2021. Each date on the x-axis corresponds to the last observation of the rolling window. Regressions are made for 32 value-weighted Size-B/M-D/E portfolios. The RHS variables are constructed using independent 2 x 3 sorts on size and each of B/M and D/E. Recession and slowdown periods are for the euro area.



Source: Based on the author's calculation run in Python and R program using data from DataStream, and Eurostat (2022a) for recession and slowdown periods.

6.3. Regression details

Having analysed the overall performance of the four-factor model, we now want to see how the model performs for each LHS portfolio (see Figure 12-Figure 14).

Panel A of Figure 12 focuses on the original three-factor model for 25 value-weighted size-B/M portfolios. We see that the portfolios of small deep value stocks produce significant positive intercepts. This result is puzzling because this issue usually comes from stocks with a low B/M (Fama & French, 1993). Comparing the t(a) of the two models (see Panel B), there is a clear improvement in the four-factor model; however, the same problem remains with the high B/M portfolios which have a statistically significant alpha. Not surprisingly, slopes for market factors are all close to one. Slopes for SMB factors are all positive and approach zero as the portfolio size is augmented; we expected to see negative slopes for big cap portfolios but it is unexpectedly not the case here; on the contrary, some big cap portfolio has a significant positive exposure to the SMB factor, with a test statistic for the intercept over 2. The slope coefficient for HML factor seems more consistent; even though small cap portfolios appear to have a less significant exposure. Lastly, we have an interesting result for LMU slope coefficients: they are all negative.

Figure 13 uses models for 25 value-weighted size-D/E portfolios. Panel A shows that the lower the leverage, the weaker is the three-factor model, and changes in size do not affect a lot the intercept significance. Looking at Panel B, we see that performance is enhanced by adding the LMU factor; some portfolio intercepts become insignificant. Low D/E portfolios have a less significant exposure to HML factor and a significant negative exposure to LMU on average. Also, micro-cap portfolios have a lower significant exposure to LMU factors on average. We still have some big cap portfolios with a significant positive exposure to the SMB factor.

The last table is Figure 14, it focuses on models for 32 value-weighted size-B/M-D/E portfolios. The interpretation of this table supports that of the previous ones; the model explains better the returns for almost all portfolios, and again, small cap high B/M portfolios have the highest intercept. Some big cap portfolios have a significant positive exposure to the SMB factor, especially those with high B/M.

In Appendix 22-Appendix 24, we have similarly represented the performance of the models over time as in Figure 9–Figure 11, replacing the GRS statistic with the t-test of the intercepts of each LHS portfolio. These graphs support our previous claim that the performance is not linear but fluctuates according to as yet unknown factors. We also wanted to highlight some LHS portfolios to see if any act in a way that might be interesting to note. However, we did not find anything that would advance our research.

Figure 12: Regressions for 25 value-weighted Size-B/M portfolios

June 1989 – May 2021, 384 months. At the end of June each year, stocks are allocated to five Size groups (Small to Big) using market cap breakpoints. Stocks are allocated independently to five B/M groups (Low B/M to High B/M), again using breakpoints. The intersections of the two sorts produce 25 Size-B/M portfolios. The LHS variables in each set of 25 regressions are the monthly excess returns on the 25 Size-B/M portfolios. The RHS variables are the excess market return, Rm-Rf, the Size factor, SMB, the value factor, HML, and the leverage factor, LMU, constructed using independent 2 x 3 sorts on Size and each of B/M, and D/E. Panel A of the table shows three-factor intercepts produced by the Mkt, SMB, and HML. Panel B shows four-factor intercepts, slopes, and t-statistics for these coefficients. The five-factor regression equation is

$$R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + l_i LMU_t + e_{it}$$

B/M →	Low	2	3	4	High	Low	2	3	4	High
Panel A: Three-factor intercepts: Rm-Rf, SMB, and HML										
			a					*t(a)*		
Small	-0,19	0,28	0,10	0,29	0,62	-0,87	1,40	0,54	1,79	4,49
2	0,02	0,04	0,17	0,09	0,41	0,19	0,40	1,60	0,82	4,09
3	0,08	-0,01	-0,01	0,13	0,29	0,66	-0,11	-0,11	1,34	2,63
4	0,26	0,21	0,21	0,32	0,11	1,97	1,86	1,83	2,88	0,85
Big	0,28	0,28	0,07	0,01	0,17	2,82	3,19	0,81	0,10	1,12
Panel B: Four-factor intercepts: Rm-Rf, SMB, HML, and LMU										
			a					*t(a)*		
Small	-0,25	0,15	-0,05	0,19	0,54	-1,08	0,71	-0,23	1,09	3,71
2	-0,06	-0,06	0,06	0,03	0,28	-0,49	-0,50	0,56	0,28	2,64
3	-0,03	-0,07	-0,05	0,13	0,25	-0,25	-0,59	-0,42	1,21	2,09
4	0,06	0,16	0,15	0,25	0,04	0,42	1,31	1,21	2,13	0,28
Big	0,14	0,25	0,06	-0,01	0,10	1,36	2,75	0,70	-0,06	0,61
			b					*t(b)*		
Small	0,86	0,90	0,80	0,73	0,81	18,78	22,00	20,43	21,54	28,05
2	0,88	0,85	0,83	0,81	0,87	35,16	36,94	36,74	36,71	41,93
3	1,00	0,91	0,83	0,83	0,89	38,18	39,80	38,57	39,90	37,84
4	1,04	0,86	0,86	0,84	0,99	38,96	36,12	35,36	36,52	36,59
Big	0,85	0,85	0,90	0,84	1,03	40,88	46,96	49,82	42,02	32,53
			s					*t(s)*		
Small	1,49	1,30	1,25	1,14	1,25	13,40	13,11	13,16	13,98	17,98
2	1,28	1,22	1,20	1,04	1,17	21,20	21,96	21,99	19,47	23,20
3	1,29	1,01	0,95	0,93	1,02	20,48	18,22	18,13	18,60	18,02
4	0,97	0,76	0,73	0,70	0,74	15,04	13,11	12,43	12,53	11,40
Big	0,13	0,06	0,14	0,03	0,46	2,64	1,33	3,11	0,61	5,97
			h					*t(h)*		
Small	-0,05	-0,13	0,17	0,24	0,55	-0,44	-1,26	1,65	2,72	7,41
2	-0,32	0,01	0,17	0,37	0,70	-4,91	0,10	2,93	6,53	12,99
3	-0,44	0,03	0,27	0,43	0,64	-6,55	0,52	4,87	8,11	10,65
4	-0,34	0,10	0,38	0,55	0,83	-4,92	1,60	5,98	9,22	11,88
Big	-0,27	0,03	0,26	0,50	0,91	-5,11	0,63	5,67	9,78	11,22
			l					*t(l)*		
Small	-0,11	-0,23	-0,26	-0,18	-0,12	-0,75	-1,70	-2,04	-1,63	-1,33
2	-0,15	-0,18	-0,19	-0,10	-0,23	-1,88	-2,45	-2,62	-1,35	-3,45
3	-0,21	-0,10	-0,06	-0,01	-0,08	-2,42	-1,36	-0,88	-0,11	-1,04
4	-0,35	-0,09	-0,11	-0,12	-0,13	-4,07	-1,20	-1,39	-1,56	-1,43
Big	-0,25	-0,04	-0,01	-0,03	-0,13	-3,66	-0,63	-0,17	-0,41	-1,23

Source: Based on the author's calculation run in Python and R program using data from DataStream, and Fama and French (2015) for the figure template.

Figure 13: Regressions for 25 value-weighted Size-D/E portfolios

June 1989 – May 2021, 384 months. At the end of June each year, stocks are allocated to five Size groups (Small to Big) using market cap breakpoints. Stocks are allocated independently to five D/E groups (Low D/E to High D/E), again using breakpoints. The intersections of the two sorts produce 25 Size-D/E portfolios. The LHS variables in each set of 25 regressions are the monthly excess returns on the 25 Size-D/E portfolios. The RHS variables are the excess market return, Rm-Rf, the Size factor, SMB, the value factor, HML, and the leverage factor, LMU, constructed using independent 2 x 3 sorts on Size and each of B/M, and D/E. Panel A of the table shows three-factor intercepts produced by the Mkt, SMB, and HML. Panel B shows four-factor intercepts, slopes, and t-statistics for these coefficients. The five-factor regression equation is

$$R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + l_i LMU_t + e_{it}$$

D/E →	Low	2	3	4	High	Low	2	3	4	High
Panel A: Three-factor intercepts: Rm-Rf, SMB, and HML										
			a					*t(a)*		
Small	0,60	0,54	0,30	0,15	-0,32	3,46	2,85	1,50	0,67	-1,72
2	0,49	0,27	0,16	0,05	-0,19	4,05	2,55	1,50	0,40	-1,60
3	0,39	0,31	0,16	0,07	-0,33	3,11	2,94	1,54	0,70	-2,83
4	0,53	0,42	0,19	0,10	-0,03	4,44	3,71	1,71	0,82	-0,26
Big	0,63	0,30	0,14	0,06	-0,05	5,29	3,00	1,52	0,94	-0,48
Panel B: Four-factor intercepts: Rm-Rf, SMB, HML, and LMU										
			a					*t(a)*		
Small	0,37	0,32	0,33	0,21	-0,25	2,03	1,61	1,57	0,88	-1,26
2	0,15	0,03	0,06	0,08	-0,02	1,26	0,28	0,54	0,67	-0,12
3	0,03	0,14	0,13	0,13	-0,14	0,26	1,30	1,21	1,14	-1,18
4	0,21	0,28	0,08	0,10	0,06	1,71	2,37	0,67	0,71	0,45
Big	0,22	0,12	0,08	0,09	0,15	1,97	1,14	0,84	1,23	1,34
			b					*t(b)*		
Small	0,81	0,85	0,82	0,89	0,83	22,54	21,43	19,66	18,90	20,89
2	0,89	0,89	0,86	0,84	0,87	38,03	41,86	39,06	34,88	35,28
3	1,01	0,88	0,89	0,90	0,93	41,27	40,71	40,81	40,40	39,29
4	0,93	0,90	0,92	0,94	1,00	39,38	38,45	39,36	35,78	37,80
Big	0,80	0,89	0,89	0,93	0,87	36,09	43,18	46,65	66,23	39,35
			s					*t(s)*		
Small	1,26	1,42	1,30	1,28	1,34	14,49	14,93	12,91	11,25	14,03
2	1,30	1,14	1,19	1,20	1,29	22,92	22,13	22,33	20,77	21,64
3	1,24	0,98	0,98	1,03	1,15	20,87	18,82	18,69	19,29	20,18
4	0,85	0,76	0,77	0,74	0,90	14,97	13,55	13,62	11,55	14,07
Big	0,17	0,02	0,07	0,02	0,23	3,20	0,38	1,46	0,54	4,26
			h					*t(h)*		
Small	0,16	0,41	0,21	0,30	0,22	1,72	4,09	1,94	2,50	2,18
2	0,22	0,23	0,32	0,22	0,21	3,58	4,16	5,72	3,54	3,28
3	-0,05	0,20	0,26	0,16	0,25	-0,81	3,69	4,71	2,74	4,06
4	0,00	0,13	0,36	0,37	0,30	0,05	2,26	6,06	5,39	4,39
Big	-0,11	0,05	0,14	0,16	0,09	-1,93	0,87	2,96	4,33	1,63
			l					*t(l)*		
Small	-0,41	-0,38	0,06	0,11	0,12	-3,50	-3,01	0,47	0,73	0,96
2	-0,60	-0,43	-0,18	0,07	0,31	-7,83	-6,18	-2,45	0,84	3,91
3	-0,64	-0,29	-0,04	0,10	0,33	-8,02	-4,20	-0,62	1,35	4,25
4	-0,58	-0,24	-0,20	-0,01	0,17	-7,62	-3,16	-2,62	-0,14	1,95
Big	-0,72	-0,33	-0,10	0,04	0,36	-9,96	-4,85	-1,61	0,98	5,03

Source: Based on the author's calculation run in Python and R program using data from DataStream, and Fama and French (2015) for the figure template

Figure 14: Regressions for 32 value-weighted Size-B/M-D/E portfolios

Regressions for 32 value-weighted Size-B/M-D/E portfolios; June 1989 – May 2021, 384 months. At the end of June each year, stocks are allocated to two Size groups (Small to Big) using market cap breakpoints. Stocks are allocated independently to four D/E and B/M groups, again using breakpoints. The intersections of the three sorts produce 32 Size-B/M-D/E portfolios. The LHS variables in each set of 32 regressions are the monthly excess returns on the 32 Size-B/M-D/E portfolios. The RHS variables are the excess market return, Rm-Rf, the Size factor, SMB, the value factor, HML, and the leverage factor, LMU, constructed using independent 2 x 3 sorts on Size and each of B/M, and D/E. Panel A of the table shows three-factor intercepts produced by the Mkt, SMB, and HML. Panel B shows four-factor intercepts, slopes, and t-statistics for these coefficients. The five-factor regression equation is

$$R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + l_i LMU_t + e_{it}$$

	Small								Big							
B/M →	Low	2	3	High	Low	2	3	High	Low	2	3	High	Low	2	3	High
Panel A: Three-factor intercepts: Rm-Rf, SMB, and HML																
	a				*t(a)*				*a*				*t(a)*			
Low D/E	0,49	0,24	0,38	0,63	3,00	1,69	3,05	4,77	0,54	0,38	0,26	0,44	4,16	2,94	1,79	2,56
2	0,28	0,11	0,39	0,32	1,63	0,87	2,97	3,02	0,40	0,37	0,22	0,20	3,10	3,18	1,83	1,28
3	-0,02	0,00	0,04	0,25	-0,14	0,01	0,31	2,10	0,23	0,13	-0,03	-0,07	1,73	1,32	-0,32	-0,46
High D/E	-0,49	-0,23	-0,21	0,15	-3,53	-1,51	-1,38	1,03	-0,10	-0,02	-0,06	0,08	-0,88	-0,13	-0,46	0,47
Panel B: Four-factor intercepts: Rm-Rf, SMB, HML, and LMU																
	a				*t(a)*				*a*				*t(a)*			
Low D/E	0,09	-0,12	0,16	0,26	0,54	-0,83	1,23	1,98	0,15	0,09	-0,01	0,23	1,15	0,67	-0,09	1,23
2	0,09	0,09	0,26	0,23	0,47	0,62	1,89	2,08	0,27	0,27	0,11	0,10	1,97	2,23	0,90	0,62
3	-0,01	-0,04	0,01	0,27	-0,06	-0,24	0,06	2,16	0,10	0,12	-0,01	-0,17	0,69	1,14	-0,07	-1,12
High D/E	-0,21	0,08	0,01	0,28	-1,44	0,47	0,09	1,84	0,11	0,12	0,07	0,20	0,85	0,92	0,54	1,03
	b				*t(b)*				*b*				*t(b)*			
Low D/E	1,05	0,92	0,84	0,86	32,13	32,82	33,43	32,96	0,82	0,77	0,83	0,71	32,91	29,38	28,02	19,74
2	0,99	0,86	0,84	0,87	27,31	31,84	30,89	39,14	0,86	0,90	0,88	0,95	31,99	37,42	35,25	28,60
3	0,92	0,81	0,80	0,84	25,85	26,49	27,24	34,00	0,91	0,93	0,87	1,01	32,72	44,55	41,59	33,45
High D/E	0,86	0,91	0,79	0,94	30,64	28,93	24,68	31,20	0,83	0,90	0,91	0,97	34,08	34,65	34,64	25,90
	s				*t(s)*				*s*				*t(s)*			
Low D/E	1,48	1,24	1,11	1,10	18,74	18,27	18,16	17,53	0,19	0,20	0,25	0,46	3,14	3,25	3,45	5,28
2	1,38	1,10	0,98	1,05	15,71	16,91	14,99	19,57	0,10	0,08	0,17	0,45	1,58	1,44	2,75	5,63
3	1,25	1,10	1,05	1,12	14,46	14,79	14,80	18,62	0,13	0,17	0,03	0,42	1,98	3,43	0,62	5,81
High D/E	1,23	1,35	1,07	1,31	18,17	17,84	13,83	17,93	0,21	0,22	0,44	0,54	3,51	3,57	6,95	6,00
	h				*t(h)*				*h*				*t(h)*			
Low D/E	-0,42	0,11	0,33	0,72	-5,06	1,54	5,09	10,75	-0,31	0,28	0,34	0,63	-4,90	4,20	4,46	6,76
2	-0,42	0,09	0,33	0,63	-4,55	1,29	4,80	11,04	-0,22	0,09	0,40	0,79	-3,14	1,43	6,22	9,20
3	-0,42	0,10	0,37	0,70	-4,60	1,24	4,96	10,96	-0,15	0,11	0,46	1,03	-2,05	2,13	8,51	13,39
High D/E	-0,30	-0,07	0,29	0,54	-4,11	-0,89	3,56	7,01	-0,16	0,11	0,42	0,80	-2,54	1,60	6,22	8,29
	l				*t(l)*				*l*				*t(l)*			
Low D/E	-0,71	-0,64	-0,39	-0,66	-6,74	-7,01	-4,75	-7,85	-0,70	-0,52	-0,49	-0,39	-8,65	-6,15	-5,06	-3,31
2	-0,35	-0,05	-0,22	-0,15	-2,96	-0,54	-2,54	-2,10	-0,23	-0,16	-0,19	-0,17	-2,63	-2,11	-2,28	-1,61
3	0,02	-0,07	-0,06	0,05	0,19	-0,68	-0,66	0,57	-0,23	-0,02	0,04	-0,19	-2,61	-0,25	0,62	-1,91
High D/E	0,50	0,55	0,41	0,24	5,53	5,39	3,92	2,43	0,37	0,25	0,23	0,20	4,70	2,92	2,71	1,64

Source: Based on the author's calculation run in Python and R program using data from DataStream, and Fama and French (2015) for the figure template.

7. DISCUSSION

On the whole, we are quite satisfied with our results: in our sample, the model we have built is indeed more accurate than the original. At the very beginning of our research, we anticipated a positive relationship between stocks expected return and financial leverage. We were therefore surprised to see that this was not the case when the portfolios had been properly formed. However, after extending our literary research, we realised the reason for this negative relationship. Reading George and Hwang (2010) article was a major enlightenment in this regard. Despite this discovery, we still believed that financial leverage has explanatory properties. This surprise may reflect some choices in our working method that should perhaps have been changed. If we were to start our research again, we would first become fully acquainted with the scientific literature before starting out empirical research. As this work was empirical in nature, care had to be taken not to neglect the theory. In fact, a major part of the time allocated to this study was devoted to the development of the software for processing the data and analysing it. We do not believe that the theoretical aspect has been overlooked. However another approach could have been considered. Yet it is through such mistakes that one improves one's methodology for future work.

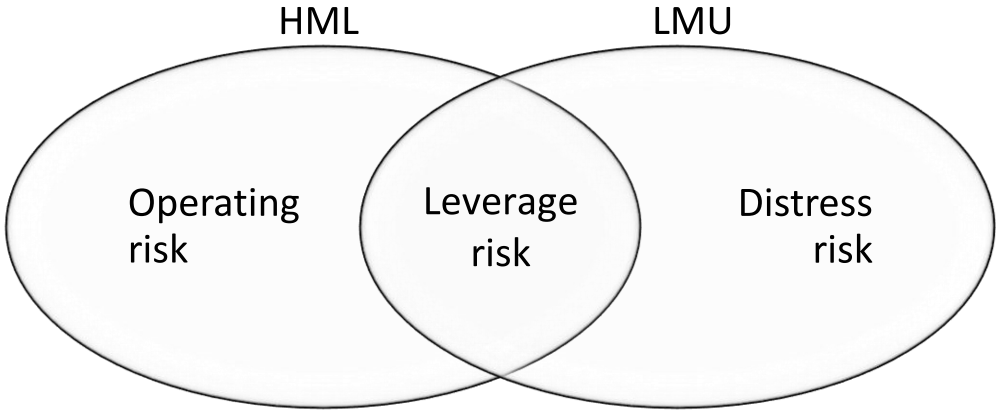
We found some rather surprising results concerning the size effect that deserve to be looked into. It appears from our sample that, on average, larger caps outperform smaller caps. We checked our sample to see if this was due to a coding error; there did not appear to be one. It is difficult to draw any hasty conclusions on this discrepancy; it may be specific to our sample. It might be helpful for future studies to take the exact same sample and analyse it in depth. Also, although the size effect is extremely well documented and its presence is proven in many historical data sets, there is no guarantee that it will last forever. In finance, it is quite common for forecasts based on the past to prove misleading. Thus, all our scientific beliefs must be constantly questioned. Whilst we do not make any inferences from this result, this one gives food for thought.

As for the value effect, it is indeed present in our sample. Stocks with a high book-to-market have a better return on average than those without. This result was not unexpected.

Looking at the average ratio obtained for each portfolio, we noticed that the portfolios with small-cap value stocks have an extreme number for their book-to-market. These results are difficult to interpret but one could think, for example, that these portfolios contain companies that are not well covered by analysts, and therefore their market price would tend to be farther away from their intrinsic value. It is not unreasonable to think that small-cap stocks are less well known to financial analysts. This lower market efficiency would lead to extreme values for valuation ratios. In addition, growth stocks happen to be more leveraged than average. The market sees these stocks as having good future growth potential, so they may be able to afford to increase their leverage. Growth stocks generally have an attractive business model that makes capital providers more confident in their investment decision. Finally, Titman and Wessels (1988) argued that the size of a firm was a determinant of its capital structure, and that they had a positive relationship; however, we found no such relationship in our sample.

Concerning the building of the factors, taking the same kind as Fama and French (2015), we have noticed that the differences in the method of construction do not have a significant impact. As mentioned earlier, the factors HML and LMU have a significant positive correlation. We can use the findings of Penman et al. (2007) to interpret this result: the value factor is divided into two parts representing operating and financial risk. We would therefore be tempted to think that the LMU factor is redundant and should be dropped. Yet we have found that our model loses performance when this is the case. We could hypothesise that the HML factor explains part of the equity risk obtained when a firm is overleveraged, the one explained by Modigliani and Miller (1958). By taking on debt, firms would increase their cost of equity and thus be undervalued relative to a peer firm with a lower degree of leverage; making the former suitable to enter the value factor. However, we have seen with George and Hwang (2010) the opposite regarding leveraged stocks: the expected return should be lower. Firms with a high risk of financial distress need to keep their debt levels low if they want to remain in business, as they are more exposed to systematic risk. This part, the financial distress cost, could therefore be explained by the LMU factor. Each of the factors would thus have their own interpretation of financial risk. We do not have much evidence to support this hypothesis, but it is one way to explain their correlation. Figure 15 shows the idea behind this hypothesis.

Figure 15: Rationale for the correlation between the HML and LMU factors



Regarding the performance of the model, it is clear that adding the leverage factor helps to improve it. It is not surprising to see that the intercepts (the amount not explained by the model) drop significantly when applying the new model on the 25 LHS Size-D/E portfolios, given the way they are constructed. The key point is that, the HML factor remains relevant with this way of sorting, which supports our hypothesis that each has a distinct explanatory power. We found that the proportion of unexplained expected return is around 48% for two types of LHS portfolios and 87% for the last type. These results are rather intriguing and suggest that there is still room for one or more factors to be added to the model. In particular, one could try to create a model by adding the leverage factor to the five-factor Fama and French (2015) model.

Unfortunately, we did not find a relationship between the performance of our model and any macroeconomic variable. However, performance does fluctuate over time and sometimes sharply; there should be a reason for this volatility. Future studies could relate the model to a wider range of variables to detect even a hint of a pattern. The prevailing monetary policy conditions (accommodative or restrictive) in the euro area could have been included as this can have a significant influence on risk appetite. We have used rolling windows and cross-correlations as methods, but other more sophisticated statistical approaches could be considered to find a significant relationship.

Looking individually at how portfolios react to the models, what is most striking is that some portfolios with large-cap stocks have a positively significant relationship with the SMB factor. We wanted to compare these results with those obtained by Fama and French (2015) and here is what they got in their case: "the SMB slopes are strongly positive for small stocks and slightly negative for big stocks. The market and SMB slopes are similar for different models, so they cannot account for changes in the intercepts observed when factors are added" (Fama & French, 2015, p. 13). We also obtained extremely positive slopes for small-cap stocks; but for large-caps, they are not slightly negative but rather slightly above zero. We also found an interesting result with small-cap value stocks: the models appeared to have difficulty explaining their performance. This result is difficult to evaluate, especially since this problem usually comes from small-cap growth stocks and not from small-cap value stocks. This anomaly could be the focus of future investigations to verify if its sample specific.

Our findings have several implications for financial market participants. Firstly, our model is a tool for investors to employ in order to estimate the required return on a stock, and to predict what they can expect by investing in it; and to screen stocks using low leverage as a criterion, in order to find stocks with a higher expected return. Secondly, asset managers will be able to use this model to attribute the performance of their portfolio to the four different risk factors, by evaluating the factor loadings derived from the regression; to look for any abnormal historical performance using event studies; to evaluate the exposure of their portfolio to the different risk factors; to assess the cost of equity of a stock for use as a discount rate in a stock valuation model, such as the discounted cash flow model; and to select the appropriate stocks if they want to calibrate their sensitivity to a particular risk factor. Thirdly, policy-makers will be able to use our model to understand what drives stock returns over different periods, and adjust their monetary policy accordingly to influence the behaviour of their desired risk factors.

8. CONCLUSION

We developed a leverage augmented Fama & French (1992, 1993) three-factor model by replicating their methodology so that we could compare our findings. Our sample consisted of listed companies in euro area countries excluding financial companies, whether still active or not, from 1989 to 2019. We have named our new factor LMU (Levered minus Unlevered), designed in the same way as the HML factor; i.e. by taking the difference, each month, between the simple average of the returns on high-leverage portfolios and the simple average of the returns on low-leverage portfolios. It turned out that expected common stock returns are negatively related to the financial leverage. We measured the performance of our model using the GRS statistical indicator (Gibbson et al., 1989), which tests the hypothesis that the intercept of the dependent variables is different from zero. The lower the value of this parameter, the better the model explains the variability of the stock returns in the sample. As a supplement, we have also introduced various indicators that take into account the absolute value of the intercepts (see Chapter 5.2). We found that, based on these indicators and in our sample, our model is a better proxy for common risk factors in returns than the original three-factor model. Furthermore, we found a significant positive relationship between the value factor and the financial leverage factor, each keeping a distinct explanatory part. Finally, we found that the performance of our model fluctuates over time, but still holds up well. We tried to relate this volatility to some macroeconomic variable related to business cycles, but we did not find anything conclusive.

In carrying out this work, many ideas came to mind about what might be the subject of future research.

We could use the market value of the leverage instead of the book value to calculate the financial leverage factor. There is a time bias when taking the book value; it is only available quarterly when the reports are published, at the very least. As a result, we end up with data that do not necessarily reflect the situation of the company at the time. The market value of an asset is quoted more frequently and more accurately, but suffers from a lack of availability for some assets, so the database would be reduced in size. Another measure of leverage could have been applied. For example, the Net debt/EBITDA ratio tells us a different picture of a company's solvency because it takes profitability into account in its calculation. This indicator is more interesting for financial analysis because it is interpreted as the number of years it takes to pay off all its debt.

Moreover, as we have already mentioned, our findings are distinguished by the absence of a size effect. We could take our sample and work on it again to discover the origin of this abnormality.

Also, we could extend our research by adding the factors RMW (robust minus weak) and CMA (conservative minus aggressive) to our model, thus obtaining a six-factor model. We would therefore see how these factors interact and perhaps learn more about the link between the value factor and the financial leverage factor. As we have seen in the literature review, financial leverage and financial distress are intrinsically linked. One could follow the methodology of Chou et al. (2010) by combining them into one model with the three original factors. This would also help us to confirm or refute George and Hwang (2010) theory on the relationship between financial distress and expected return.

Next, we chose to build the financial leverage factor by being long the levered ones and short the unlevered ones, i.e. levered minus unlevered (LMU). The factor could have been set up differently, i.e. unlevered minus levered (UML). This method would have the advantage of transforming the fourth factor into a risk premium, which is theoretically positively linked to the expected stock return.

Finally, the benchmark we selected to construct the market factor is quite limited; it represents only the large caps of the countries in our sample. Although we believe that this was not too detrimental to our study, it did result in a significantly negative relationship between the size factor (SMB) and the market factor, and non-significant positive expected return for the market factor. The choice of this index could therefore be reviewed for future work on the same topic and replaced by a more inclusive one.

APPENDICES

Appendix 1: Panel data extracted for each workbook

Workbook	Panel data
Benchmark	S&P Euro monthly price from 31/12/1969 to 31/12/2021
Risk-free rate	Three-month monthly Treasury bill of German interest rates from 31/12/1969 to 31/12/2021
Main data	Monthly price of Eurozone stocks from 31/12/1969 to 31/12/2021 (excluding financial stocks)
	Monthly market value of Eurozone stocks from 31/12/1969 to 31/12/2021 (excluding financial stocks)
	Monthly common shareholders' equity of Eurozone stocks from 31/12/1969 to 31/12/2021 (excluding financial stocks)
	Monthly total debt of Eurozone stocks from 31/12/1969 to 31/12/2021 (excluding financial stocks)
	Monthly total asset of Eurozone stocks from 31/12/1969 to 31/12/2021 (excluding financial stocks)

Appendix 2: Description of each variable according to Thompson Reuters DataStream

P – Price - Trade	Represent the official closing price adjusted for subsequent capital action.
MV – Market Value (Capital)	The share price multiplied by the number of ordinary shares in issue. The amount in issue is updated whenever new tranches of stock are issued or after a capital change.
WC03501 – Common Equity	Represent common shareholders' investment in a company.
WC03255 – Total Debt	Represent all interest bearing and capitalized lease obligations. It is the sum of long and short-term debt.
WC02999 – Total Assets	Represent the sum of total current assets, long-term receivables, investment in unconsolidated subsidiaries, other investments, net PPE and other assets.

Source: Thompson Reuters DataStream

Appendix 3: Number of stocks included in the sample for each country per year before data cleansing

For each end of June, we counted the number of monthly price returns that are not a 'NA' value to get an estimate of the number of stocks.

	Austria	Belgium	Finland	France	Germany	Ireland	Italy	Netherland	Portugal	Spain	TOTAL
1990	39	70	41	306	398	39	156	136	69	87	1341
1995	76	109	94	486	490	41	176	157	100	118	1847
2000	105	226	170	944	1338	59	271	245	121	191	3670
2005	141	262	199	1166	1748	62	361	267	135	220	4561
2010	167	340	220	1433	2436	72	437	293	143	261	5802
2015	178	336	248	1601	2602	77	511	304	149	311	6317
2020	195	354	301	1698	2825	82	634	324	150	422	6985

Source: Based on the author's calculation run in Python program using data from DataStream

Appendix 4: Example of stock excluded from the sample because of '#ERROR' columns

ALPE-ADRIA-IMMOBILIEN AG	ALPE-ADRIA-IMMOBILIEN AG	#ERROR	#ERROR	#ERROR
105,8	3,17	$$ER: E100,NO WORLDSCOPE	$$ER: E100,NO WORLDSCOPE	$$ER: E100,NO WORLDSCOPE
106,16	3,18			
107,36	3,22			
102,8	3,08			
102,94	3,09			

Source: Thompson Reuters DataStream

Appendix 5: Example of stock excluded from the German sample because of '(XET)' labelled columns

1&1 (XET)	1&1 (XET) - MARKET VAL	1&1 (XET) - COMMON SH	1&1 (XET) - TOTAL DEBT	1&1 (XET) - TOTAL ASSET
40,895	2239,6	283400	146067	582507
42,82	2345,02	3805080	267	4591122
42,8	2343,93	3805080	267	4591122
47,7	2612,27	3805080	267	4591122
49,205	2694,69	3805080	267	4591122

Source: Thompson Reuters DataStream

Appendix 6: Number of stocks included in the sample for each country per year after data cleansing

The last column corresponds to the relative variation of the cleaned total sample compared to the uncleaned one, for each year.

	Austria	Belgium	Finland	France	Germany	Ireland	Italy	Netherland	Portugal	Spain	TOTAL	Variation
1990	32	64	33	286	331	27	140	120	48	81	1162	-13%
1995	66	75	74	422	359	26	145	137	65	103	1472	-20%
2000	71	132	139	732	647	42	217	177	68	126	2351	-36%
2005	67	118	131	642	645	33	218	129	51	111	2145	-53%
2010	70	127	119	634	667	29	228	101	44	113	2132	-63%
2015	61	112	119	547	534	27	245	91	42	127	1905	-70%
2020	45	99	136	497	465	24	296	89	32	159	1842	-74%

Source: Based on the author's calculation run in Python program using data from DataStream.

Appendix 7: Descriptive statistics for price return from June 1989 to June 2020 (stated in decimal form)

	Jun 1989	Jun 1990	Jun 1991	Jun 1992	Jun 1993	Jun 1994	Jun 1995	Jun 1996	Jun 1997	Jun 1998	Jun 1999	Jun 2000	Jun 2001	Jun 2002	Jun 2003	Jun 2004
Count	1 091	1 162	1 266	1 267	1 292	1 384	1 472	1 551	1 720	1 927	2 128	2 351	2 457	2 405	2 201	2 178
Minimum	-0,90	-1,00	-0,50	-0,59	-0,50	-0,42	-0,77	-0,57	-0,82	-0,63	-0,63	-0,85	-0,91	-0,95	-0,98	-0,70
1st percentile	-0,15	-0,19	-0,23	-0,32	-0,29	-0,25	-0,23	-0,23	-0,23	-0,28	-0,26	-0,36	-0,53	-0,54	-0,34	-0,35
5st percentile	-0,10	-0,11	-0,13	-0,21	-0,15	-0,17	-0,14	-0,15	-0,11	-0,18	-0,14	-0,20	-0,33	-0,34	-0,14	-0,18
Median	0,01	-0,00	-0,03	-0,05	-0,00	-0,04	-0,01	-0,01	0,02	-0,02	0,00	-0,01	-0,04	-0,06	0,03	0,01
95st percentile	0,19	0,15	0,10	0,07	0,16	0,08	0,13	0,17	0,23	0,16	0,23	0,20	0,14	0,11	0,39	0,21
99st percentile	0,36	0,29	0,22	0,16	0,29	0,24	0,27	0,34	0,44	0,35	0,55	0,44	0,45	0,36	0,95	0,48
Maximum	0,84	0,55	0,69	0,81	4,00	0,88	1,00	1,55	3,15	6,10	2,46	1,77	4,19	1,05	3,56	4,26

	Jun 2005	Jun 2006	Jun 2007	Jun 2008	Jun 2009	Jun 2010	Jun 2011	Jun 2012	Jun 2013	Jun 2014	Jun 2015	Jun 2016	Jun 2017	Jun 2018	Jun 2019	Jun 2020
Count	2 145	2 189	2 346	2 309	2 184	2 132	2 086	1 984	1 877	1 894	1 905	1 895	1 874	1 874	1 842	1 842
Minimum	-0,69	-0,83	-0,83	-0,86	-0,95	-0,77	-0,80	-0,90	-0,71	-0,90	-0,64	-0,86	-0,75	-0,74	-0,89	-0,86
1st percentile	-0,28	-0,31	-0,30	-0,45	-0,50	-0,50	-0,48	-0,45	-0,40	-0,36	-0,41	-0,47	-0,33	-0,37	-0,33	-0,43
5st percentile	-0,14	-0,19	-0,15	-0,29	-0,22	-0,20	-0,22	-0,20	-0,20	-0,16	-0,18	-0,23	-0,15	-0,18	-0,13	-0,16
Median	0,02	-0,02	-0,01	-0,08	-0,01	-0,01	-0,03	-0,00	-0,03	-0,01	-0,04	-0,04	-0,01	-0,02	0,01	0,01
95st percentile	0,24	0,11	0,17	0,11	0,31	0,19	0,12	0,20	0,15	0,17	0,12	0,13	0,22	0,15	0,20	0,28
99st percentile	0,50	0,33	0,40	0,38	0,83	0,54	0,50	0,63	0,51	0,63	0,45	0,38	0,62	0,49	0,52	0,77
Maximum	2,47	1,84	1,84	2,49	9,20	2,30	4,00	4,81	4,45	4,00	1,94	6,22	3,02	1,56	1,08	9,50

Source: Based on the author's calculation run in Python program using data from DataStream.

Appendix 8: Descriptive statistics for size from June 1989 to June 2020

	Jun 1989	Jun 1990	Jun 1991	Jun 1992	Jun 1993	Jun 1994	Jun 1995	Jun 1996	Jun 1997	Jun 1998	Jun 1999	Jun 2000	Jun 2001	Jun 2002	Jun 2003	Jun 2004
Count	1 155	1 253	1 376	1 420	1 465	1 542	1 658	1 746	1 926	2 151	2 479	2 837	3 075	3 156	3 189	3 235
Minimum	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1st percentile	2	2	2	1	1	1	1	1	1	1	1	1	0	0	0	0
5st percentile	6	7	6	4	4	4	4	3	4	4	4	4	3	1	1	1
Median	78	92	70	63	58	76	69	72	79	99	83	100	70	52	46	56
95st percentile	1 560	1 881	1 632	1 618	1 621	2 082	2 183	2 550	3 403	3 801	3 624	3 945	3 370	3 298	2 960	3 559
99st percentile	6 609	7 426	5 583	6 096	6 682	7 833	8 802	9 961	15 911	20 833	21 332	27 692	24 344	18 843	15 644	18 476
Maximum	14 520	19 046	16 950	17 743	16 497	18 186	19 326	31 583	58 919	69 300	121 844	250 186	125 799	116 514	90 434	102 256

	Jun 2005	Jun 2006	Jun 2007	Jun 2008	Jun 2009	Jun 2010	Jun 2011	Jun 2012	Jun 2013	Jun 2014	Jun 2015	Jun 2016	Jun 2017	Jun 2018	Jun 2019	Jun 2020
Count	3 300	3 476	3 709	3 815	3 842	3 895	3 951	3 992	4 029	4 111	4 206	4 291	4 368	4 464	4 544	4 610
Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1st percentile	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5st percentile	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Median	66	74	89	71	55	60	63	55	57	68	69	66	76	77	71	68
95st percentile	4 267	4 766	5 211	4 208	3 493	3 750	4 233	3 861	4 080	4 637	5 154	4 811	5 494	5 466	5 440	5 388
99st percentile	20 553	24 765	33 496	22 873	19 551	20 733	23 900	21 079	22 645	29 799	31 000	26 339	31 560	33 846	34 246	32 192
Maximum	123 383	127 180	146 284	130 102	91 267	86 809	94 059	98 467	109 917	134 932	172 886	189 129	163 753	146 466	189 183	197 213

Source: Based on the author's calculation run in Python program using data from DataStream.

Appendix 9: Descriptive statistics for common shareholder's equity from June 1989 to June 2020

	Jun 1989	Jun 1990	Jun 1991	Jun 1992	Jun 1993	Jun 1994	Jun 1995	Jun 1996	Jun 1997	Jun 1998	Jun 1999	Jun 2000	Jun 2001	Jun 2002	Jun 2003	Jun 2004
Count	1 165	1 257	1 303	1 355	1 433	1 488	1 546	2 158	2 527	2 718	2 730	2 698	2 548	2 437	2 349	2 369
Minimum	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1st percentile	2	2	2	2	1	1	1	0	0	0	0	0	0	0	0	0
5st percentile	5	5	6	5	5	5	5	1	1	1	2	2	2	2	1	1
Median	56	60	64	62	61	65	68	41	33	34	39	49	50	46	43	43
95st percentile	1 539	1 586	1 581	1 534	1 397	1 688	1 812	1 457	1 386	1 494	1 758	2 075	2 367	2 456	2 606	2 489
99st percentile	5 852	6 235	7 034	7 295	7 658	8 202	8 629	7 258	7 512	7 711	9 547	13 270	12 274	11 247	12 923	14 336
Maximum	10 326	11 678	18 206	18 449	19 108	20 837	21 388	23 215	28 413	30 367	36 060	42 409	60 987	34 914	34 481	34 950

	Jun 2005	Jun 2006	Jun 2007	Jun 2008	Jun 2009	Jun 2010	Jun 2011	Jun 2012	Jun 2013	Jun 2014	Jun 2015	Jun 2016	Jun 2017	Jun 2018	Jun 2019	Jun 2020
Count	2 419	2 397	2 359	2 255	2 185	2 125	2 069	2 029	2 027	2 057	2 074	2 089	2 111	2 066	1 984	1 870
Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1st percentile	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
5st percentile	1	1	2	2	2	2	2	2	2	1	2	1	2	2	2	3
Median	46	53	63	62	64	68	72	73	69	74	75	77	78	81	90	100
95st percentile	2 702	3 447	3 519	3 391	3 743	4 290	4 500	4 837	5 106	4 944	5 388	5 508	5 936	6 287	6 693	6 734
99st percentile	14 388	17 458	19 339	19 484	19 910	23 884	22 791	24 302	26 085	24 524	24 285	26 573	28 735	31 748	30 909	30 567
Maximum	46 637	47 845	49 374	57 748	60 285	62 204	68 037	77 515	157 990	162 692	158 266	154 884	173 472	182 501	109 118	111 336

Source: Based on the author's calculation run in Python program using data from DataStream.

Appendix 10: Descriptive statistics for total debt from June 1989 to June 2020

	Jun 1989	Jun 1990	Jun 1991	Jun 1992	Jun 1993	Jun 1994	Jun 1995	Jun 1996	Jun 1997	Jun 1998	Jun 1999	Jun 2000	Jun 2001	Jun 2002	Jun 2003	Jun 2004
Count	1 123	1 201	1 252	1 308	1 386	1 435	1 496	2 055	2 423	2 610	2 599	2 560	2 458	2 365	2 294	2 308
Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1st percentile	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5st percentile	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
Median	32	37	42	43	43	42	43	24	18	18	20	25	28	28	26	23
95st percentile	1 664	2 075	2 339	2 464	2 392	2 232	1 939	1 283	1 246	1 289	1 639	2 235	2 575	2 515	2 489	2 445
99st percentile	5 063	6 298	8 738	9 942	10 330	9 330	8 243	6 230	5 738	6 162	10 197	15 130	18 230	18 407	15 936	13 347
Maximum	42 744	38 168	44 584	52 457	53 779	59 917	56 440	51 072	44 529	40 430	64 488	84 783	90 908	79 112	75 690	76 620

	Jun 2005	Jun 2006	Jun 2007	Jun 2008	Jun 2009	Jun 2010	Jun 2011	Jun 2012	Jun 2013	Jun 2014	Jun 2015	Jun 2016	Jun 2017	Jun 2018	Jun 2019	Jun 2020
Count	2 308	2 285	2 244	2 177	2 086	2 036	2 008	1 995	2 003	1 997	1 995	2 024	2 045	2 024	2 002	1 913
Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1st percentile	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5st percentile	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
Median	26	28	35	41	40	42	45	43	42	46	45	44	44	46	58	63
95st percentile	2 850	3 109	3 427	4 080	4 321	4 608	4 818	4 966	4 794	5 492	5 306	5 217	5 166	5 530	6 830	7 608
99st percentile	14 285	19 424	20 837	21 987	26 745	22 526	26 772	26 061	25 827	24 616	29 661	30 110	30 870	29 209	32 280	32 750
Maximum	80 932	78 518	80 885	69 380	94 794	77 011	93 533	117 663	123 508	139 021	153 165	162 386	174 560	203 479	211 073	216 467

Source: Based on the author's calculation run in Python program using data from DataStream.

Appendix 11: Descriptive statistics for total asset from June 1989 to June 2020

	Jun 1989	Jun 1990	Jun 1991	Jun 1992	Jun 1993	Jun 1994	Jun 1995	Jun 1996	Jun 1997	Jun 1998	Jun 1999	Jun 2000	Jun 2001	Jun 2002	Jun 2003	Jun 2004
Count	1 170	1 264	1 316	1 370	1 458	1 510	1 571	2 205	2 602	2 798	2 781	2 741	2 619	2 542	2 492	2 522
Minimum	2	1	1	2	2	2	0	0	0	0	0	0	0	0	0	0
1st percentile	5	5	5	5	5	4	4	1	1	1	1	2	2	1	1	1
5st percentile	13	15	15	16	15	15	15	5	3	3	5	6	6	4	4	3
Median	160	171	194	199	193	196	200	112	95	93	103	124	129	121	110	103
95st percentile	6 070	6 752	6 887	7 752	7 035	8 183	8 210	5 268	5 054	5 308	6 243	7 301	8 098	7 889	7 944	7 877
99st percentile	21 409	24 506	32 056	35 805	33 922	34 156	35 325	24 778	26 295	26 555	41 852	51 048	51 191	52 066	52 893	55 469
Maximum	87 501	93 196	97 738	98 845	101 460	102 222	104 222	106 086	125 522	264 356	176 037	197 698	206 985	185 613	176 596	181 201

	Jun 2005	Jun 2006	Jun 2007	Jun 2008	Jun 2009	Jun 2010	Jun 2011	Jun 2012	Jun 2013	Jun 2014	Jun 2015	Jun 2016	Jun 2017	Jun 2018	Jun 2019	Jun 2020
Count	2 517	2 478	2 441	2 349	2 280	2 222	2 191	2 158	2 171	2 192	2 182	2 202	2 209	2 170	2 099	1 978
Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1st percentile	0	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1
5st percentile	3	4	4	5	4	4	5	5	4	4	4	5	4	6	7	8
Median	107	124	145	150	146	161	167	166	166	165	174	168	174	177	209	228
95st percentile	8 920	10 163	10 367	11 381	11 698	12 923	14 608	15 251	14 864	15 941	15 893	16 646	18 015	20 696	23 476	25 218
99st percentile	58 999	59 220	68 297	81 658	77 697	81 525	82 222	77 955	76 036	80 983	79 173	80 255	72 280	87 576	96 354	99 798
Maximum	198 904	217 265	234 676	262 215	238 815	238 434	247 293	301 729	362 914	375 831	373 909	399 976	412 383	448 025	474 965	483 628

Source: Based on the author's calculation run in Python program using data from DataStream.

Appendix 12: Descriptive statistics for 5% winsorized price return from June 1989 to June 2020 (stated in decimal form)

	Jun 1989	Jun 1990	Jun 1991	Jun 1992	Jun 1993	Jun 1994	Jun 1995	Jun 1996	Jun 1997	Jun 1998	Jun 1999	Jun 2000	Jun 2001	Jun 2002	Jun 2003	Jun 2004
Count	1 091	1 162	1 266	1 267	1 292	1 384	1 472	1 551	1 720	1 927	2 128	2 351	2 457	2 405	2 201	2 178
Minimum	- 0,12	- 0,15	- 0,16	- 0,27	- 0,21	- 0,20	- 0,18	- 0,19	- 0,16	- 0,22	- 0,19	- 0,28	- 0,41	- 0,43	- 0,21	- 0,25
1st percentile	- 0,12	- 0,15	- 0,16	- 0,27	- 0,21	- 0,20	- 0,18	- 0,19	- 0,16	- 0,22	- 0,19	- 0,28	- 0,41	- 0,43	- 0,21	- 0,25
5st percentile	- 0,10	- 0,11	- 0,13	- 0,21	- 0,15	- 0,17	- 0,14	- 0,15	- 0,11	- 0,18	- 0,14	- 0,20	- 0,33	- 0,34	- 0,14	- 0,18
Median	0,01	- 0,00	- 0,03	- 0,05	- 0,00	- 0,04	- 0,01	- 0,01	0,02	- 0,02	0,00	- 0,01	- 0,04	- 0,06	0,03	0,01
95st percentile	0,19	0,15	0,10	0,07	0,16	0,08	0,13	0,17	0,23	0,16	0,23	0,20	0,14	0,11	0,39	0,21
99st percentile	0,25	0,20	0,15	0,10	0,22	0,12	0,18	0,23	0,33	0,25	0,36	0,30	0,21	0,20	0,60	0,29
Maximum	0,25	0,20	0,15	0,10	0,22	0,12	0,18	0,23	0,33	0,25	0,36	0,30	0,21	0,20	0,60	0,29

	Jun 2005	Jun 2006	Jun 2007	Jun 2008	Jun 2009	Jun 2010	Jun 2011	Jun 2012	Jun 2013	Jun 2014	Jun 2015	Jun 2016	Jun 2017	Jun 2018	Jun 2019	Jun 2020
Count	2 145	2 189	2 346	2 309	2 184	2 132	2 086	1 984	1 877	1 894	1 905	1 895	1 874	1 874	1 842	1 842
Minimum	- 0,20	- 0,25	- 0,21	- 0,35	- 0,32	- 0,31	- 0,32	- 0,29	- 0,28	- 0,24	- 0,26	- 0,30	- 0,24	- 0,24	- 0,20	- 0,25
1st percentile	- 0,20	- 0,25	- 0,21	- 0,35	- 0,32	- 0,31	- 0,32	- 0,29	- 0,28	- 0,24	- 0,26	- 0,30	- 0,24	- 0,24	- 0,20	- 0,25
5st percentile	- 0,14	- 0,19	- 0,15	- 0,29	- 0,22	- 0,20	- 0,22	- 0,20	- 0,20	- 0,16	- 0,18	- 0,23	- 0,15	- 0,18	- 0,13	- 0,16
Median	0,02	- 0,02	- 0,01	- 0,08	- 0,01	- 0,01	- 0,03	- 0,00	- 0,03	- 0,01	- 0,04	- 0,04	- 0,01	- 0,02	0,01	0,01
95st percentile	0,24	0,11	0,17	0,11	0,31	0,19	0,12	0,20	0,15	0,17	0,12	0,13	0,22	0,15	0,20	0,28
99st percentile	0,36	0,19	0,25	0,19	0,49	0,30	0,24	0,30	0,25	0,29	0,22	0,22	0,37	0,25	0,30	0,43
Maximum	0,36	0,19	0,25	0,19	0,49	0,30	0,24	0,30	0,25	0,29	0,22	0,22	0,37	0,25	0,30	0,43

Source: Based on the author's calculation run in Python program using data from DataStream.

Appendix 13: Descriptive statistics for 5% winsorized size from June 1989 to June 2020

	Jun 1989	Jun 1990	Jun 1991	Jun 1992	Jun 1993	Jun 1994	Jun 1995	Jun 1996	Jun 1997	Jun 1998	Jun 1999	Jun 2000	Jun 2001	Jun 2002	Jun 2003	Jun 2004
Count	1 155	1 253	1 376	1 420	1 465	1 542	1 658	1 746	1 926	2 151	2 479	2 837	3 075	3 156	3 189	3 235
Minimum	3	4	3	2	2	2	2	2	2	2	2	2	1	1	0	0
1st percentile	3	4	3	2	2	2	2	2	2	2	2	2	1	1	0	0
5st percentile	6	7	6	4	4	4	4	3	4	4	4	4	3	1	1	1
Median	78	92	70	63	58	76	69	72	79	99	83	100	70	52	46	56
95st percentile	1 560	1 881	1 632	1 618	1 621	2 082	2 183	2 550	3 403	3 801	3 624	3 945	3 370	3 298	2 960	3 559
99st percentile	3 083	3 521	3 276	3 269	3 329	3 913	4 042	5 232	7 165	9 287	8 812	10 961	9 833	8 632	6 835	8 104
Maximum	3 083	3 521	3 276	3 269	3 329	3 913	4 042	5 232	7 165	9 287	8 812	10 961	9 833	8 632	6 835	8 104

	Jun 2005	Jun 2006	Jun 2007	Jun 2008	Jun 2009	Jun 2010	Jun 2011	Jun 2012	Jun 2013	Jun 2014	Jun 2015	Jun 2016	Jun 2017	Jun 2018	Jun 2019	Jun 2020
Count	3 300	3 476	3 709	3 815	3 842	3 895	3 951	3 992	4 029	4 111	4 206	4 291	4 368	4 464	4 544	4 610
Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1st percentile	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5st percentile	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Median	66	74	89	71	55	60	63	55	57	68	69	66	76	77	71	68
95st percentile	4 267	4 766	5 211	4 208	3 493	3 750	4 233	3 861	4 080	4 637	5 154	4 811	5 494	5 466	5 440	5 388
99st percentile	9 266	10 571	11 708	9 785	7 597	8 753	9 755	8 169	9 217	10 894	11 847	10 938	12 652	12 637	12 013	11 765
Maximum	9 266	10 571	11 708	9 785	7 597	8 753	9 755	8 169	9 217	10 894	11 847	10 938	12 652	12 637	12 013	11 765

Source: Based on the author's calculation run in Python program using data from DataStream.

Appendix 14: Descriptive statistics for 5% winsorized common shareholders' equity from June 1989 to June 2020

	Jun 1989	Jun 1990	Jun 1991	Jun 1992	Jun 1993	Jun 1994	Jun 1995	Jun 1996	Jun 1997	Jun 1998	Jun 1999	Jun 2000	Jun 2001	Jun 2002	Jun 2003	Jun 2004
Count	1 165	1 257	1 303	1 355	1 433	1 488	1 546	2 158	2 527	2 718	2 730	2 698	2 548	2 437	2 349	2 369
Minimum	3	3	3	3	3	3	2	1	0	0	1	1	1	1	0	0
1st percentile	3	3	3	3	3	3	2	1	0	0	1	1	1	1	0	0
5st percentile	5	5	6	5	5	5	5	1	1	1	2	2	2	2	1	1
Median	56	60	64	62	61	65	68	41	33	34	39	49	50	46	43	43
95st percentile	1 539	1 586	1 581	1 534	1 397	1 688	1 812	1 457	1 386	1 494	1 758	2 075	2 367	2 456	2 606	2 489
99st percentile	2 604	2 815	3 118	3 006	3 174	3 344	3 451	2 797	3 007	3 059	3 915	4 765	5 417	5 088	5 183	5 122
Maximum	2 604	2 815	3 118	3 006	3 174	3 344	3 451	2 797	3 007	3 059	3 915	4 765	5 417	5 088	5 183	5 122

	Jun 2005	Jun 2006	Jun 2007	Jun 2008	Jun 2009	Jun 2010	Jun 2011	Jun 2012	Jun 2013	Jun 2014	Jun 2015	Jun 2016	Jun 2017	Jun 2018	Jun 2019	Jun 2020
Count	2 419	2 397	2 359	2 255	2 185	2 125	2 069	2 029	2 027	2 057	2 074	2 089	2 111	2 066	1 984	1 870
Minimum	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1st percentile	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5st percentile	1	1	2	2	2	2	2	2	2	1	2	1	2	2	2	3
Median	46	53	63	62	64	68	72	73	69	74	75	77	78	81	90	100
95st percentile	2 702	3 447	3 519	3 391	3 743	4 290	4 500	4 837	5 106	4 944	5 388	5 508	5 936	6 287	6 693	6 734
99st percentile	5 865	6 767	7 956	7 732	8 501	9 557	9 671	10 207	10 624	10 904	10 963	12 475	13 116	14 284	14 455	13 830
Maximum	5 865	6 767	7 956	7 732	8 501	9 557	9 671	10 207	10 624	10 904	10 963	12 475	13 116	14 284	14 455	13 830

Source: Based on the author's calculation run in Python program using data from DataStream.

Appendix 15: Descriptive statistics for 5% winsorized total debt from June 1989 to June 2020

	Jun 1989	Jun 1990	Jun 1991	Jun 1992	Jun 1993	Jun 1994	Jun 1995	Jun 1996	Jun 1997	Jun 1998	Jun 1999	Jun 2000	Jun 2001	Jun 2002	Jun 2003	Jun 2004
Count	1 123	1 201	1 252	1 308	1 386	1 435	1 496	2 055	2 423	2 610	2 599	2 560	2 458	2 365	2 294	2 308
Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1st percentile	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5st percentile	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
Median	32	37	42	43	43	42	43	24	18	18	20	25	28	28	26	23
95st percentile	1 664	2 075	2 339	2 464	2 392	2 232	1 939	1 283	1 246	1 289	1 639	2 235	2 575	2 515	2 489	2 445
99st percentile	3 637	3 760	4 066	4 049	3 990	3 951	3 702	3 025	2 745	3 469	4 184	5 643	6 410	6 269	6 029	5 911
Maximum	3 637	3 760	4 066	4 049	3 990	3 951	3 702	3 025	2 745	3 469	4 184	5 643	6 410	6 269	6 029	5 911

	Jun 2005	Jun 2006	Jun 2007	Jun 2008	Jun 2009	Jun 2010	Jun 2011	Jun 2012	Jun 2013	Jun 2014	Jun 2015	Jun 2016	Jun 2017	Jun 2018	Jun 2019	Jun 2020
Count	2 308	2 285	2 244	2 177	2 086	2 036	2 008	1 995	2 003	1 997	1 995	2 024	2 045	2 024	2 002	1 913
Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1st percentile	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5st percentile	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
Median	26	28	35	41	40	42	45	43	42	46	45	44	44	46	58	63
95st percentile	2 850	3 109	3 427	4 080	4 321	4 608	4 818	4 966	4 794	5 492	5 306	5 217	5 166	5 530	6 830	7 608
99st percentile	6 481	7 034	7 656	8 849	9 967	9 793	10 747	11 924	11 975	12 551	12 852	12 679	12 388	12 623	15 658	16 001
Maximum	6 481	7 034	7 656	8 849	9 967	9 793	10 747	11 924	11 975	12 551	12 852	12 679	12 388	12 623	15 658	16 001

Source: Based on the author's calculation run in Python program using data from DataStream.

Appendix 16: Descriptive statistics for 5% winsorized total asset from June 1989 to June 2020

	Jun 1989	Jun 1990	Jun 1991	Jun 1992	Jun 1993	Jun 1994	Jun 1995	Jun 1996	Jun 1997	Jun 1998	Jun 1999	Jun 2000	Jun 2001	Jun 2002	Jun 2003	Jun 2004
Count	1 170	1 264	1 316	1 370	1 458	1 510	1 571	2 205	2 602	2 798	2 781	2 741	2 619	2 542	2 492	2 522
Minimum	8	9	8	9	9	10	9	3	2	1	3	3	3	2	2	1
1st percentile	8	9	8	9	9	10	9	3	2	1	3	3	3	2	2	1
5st percentile	13	15	15	16	15	15	15	5	3	3	5	6	6	4	4	3
Median	160	171	194	199	193	196	200	112	95	93	103	124	129	121	110	103
95st percentile	6 070	6 752	6 887	7 752	7 035	8 183	8 210	5 268	5 054	5 308	6 243	7 301	8 098	7 889	7 944	7 877
99st percentile	10 637	12 871	14 170	14 574	14 066	15 778	14 907	11 350	10 420	11 186	14 629	17 496	18 371	18 166	17 949	18 455
Maximum	10 637	12 871	14 170	14 574	14 066	15 778	14 907	11 350	10 420	11 186	14 629	17 496	18 371	18 166	17 949	18 455

	Jun 2005	Jun 2006	Jun 2007	Jun 2008	Jun 2009	Jun 2010	Jun 2011	Jun 2012	Jun 2013	Jun 2014	Jun 2015	Jun 2016	Jun 2017	Jun 2018	Jun 2019	Jun 2020
Count	2 517	2 478	2 441	2 349	2 280	2 222	2 191	2 158	2 171	2 192	2 182	2 202	2 209	2 170	2 099	1 978
Minimum	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	4
1st percentile	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	4
5st percentile	3	4	4	5	4	4	5	5	4	4	4	5	4	6	7	8
Median	107	124	145	150	146	161	167	166	166	165	174	168	174	177	209	228
95st percentile	8 920	10 163	10 367	11 381	11 698	12 923	14 608	15 251	14 864	15 941	15 893	16 646	18 015	20 696	23 476	25 218
99st percentile	21 480	22 119	26 967	28 356	29 297	30 426	32 193	36 482	34 486	35 281	36 838	40 585	40 266	43 369	44 449	43 334
Maximum	21 480	22 119	26 967	28 356	29 297	30 426	32 193	36 482	34 486	35 281	36 838	40 585	40 266	43 369	44 449	43 334

Source: Based on the author's calculation run in Python program using data from DataStream.

Appendix 17: Sorting process for all 6 Size-B/M portfolios with the 2x3 sort

Market capitalisation	Book-to-market	Group
Big MV	High B/M	BHV
	Medium B/M	BMV
	Low B/M	BLV
Small MV	High B/M	SHV
	Medium B/M	SMV
	Low B/M	SLV

Appendix 18: Sorting process for all 6 Size-D/E portfolios with the 2x3 sort

Market capitalisation	Debt/Equity	Group
Big MV	High D/E	BHL
	Medium D/E	BML
	Low D/E	BLL
Small MV	High D/E	SHL
	Medium D/E	SML
	Low D/E	SLL

Appendix 19: Sorting process for all 4 Size-B/M portfolios with the 2x2 sort

Market capitalisation	Book-to-market	Group
Big MV	High B/M	BHV
	Low B/M	BLV
Small MV	High B/M	SHV
	Low B/M	SLV

Appendix 20: Sorting process for all 4 Size-D/E portfolios with the 2x2 sort

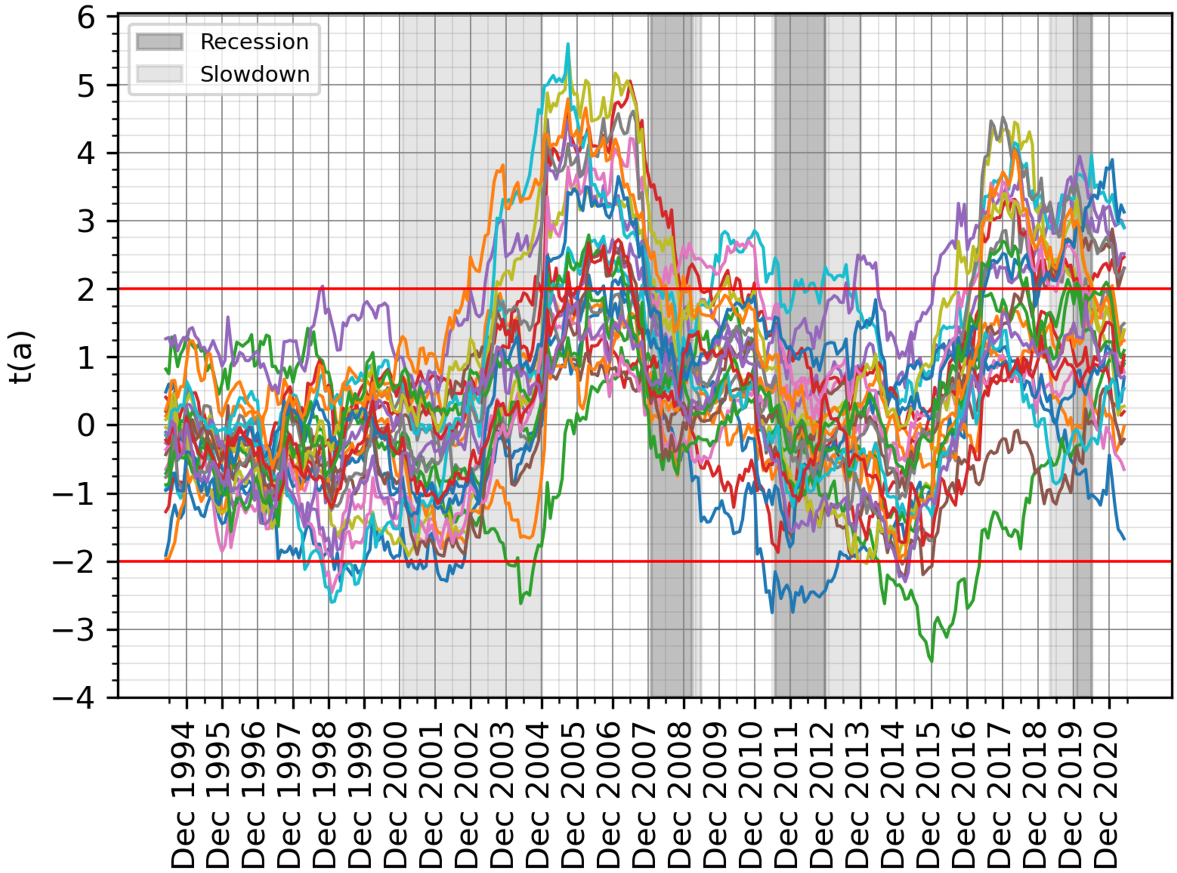
Market capitalisation	Debt/Equity	Group
Big MV	High D/E	BHL
	Low D/E	BLL
Small MV	High D/E	SHL
	Low D/E	SLL

Appendix 21: Sorting process for all 8 Size-B/M-D/E portfolios with the 2x2x2 sort

Market capitalisation	Book-to-market	Debt/Equity	Group
Big MV	High B/M	High D/E	BHVHL
		Low D/E	BHVLL
	Low B/M	High D/E	BLVHL
		Low D/E	BLVLL
Big MV	High B/M	High D/E	SHVHL
		Low D/E	SHVLL
	Low B/M	High D/E	SLVHL
		Low D/E	SLVLL

Appendix 22: Performance of the four-factor model over time, measured by the 25 intercept t-statistic of the value-weighted Size-B/M portfolios

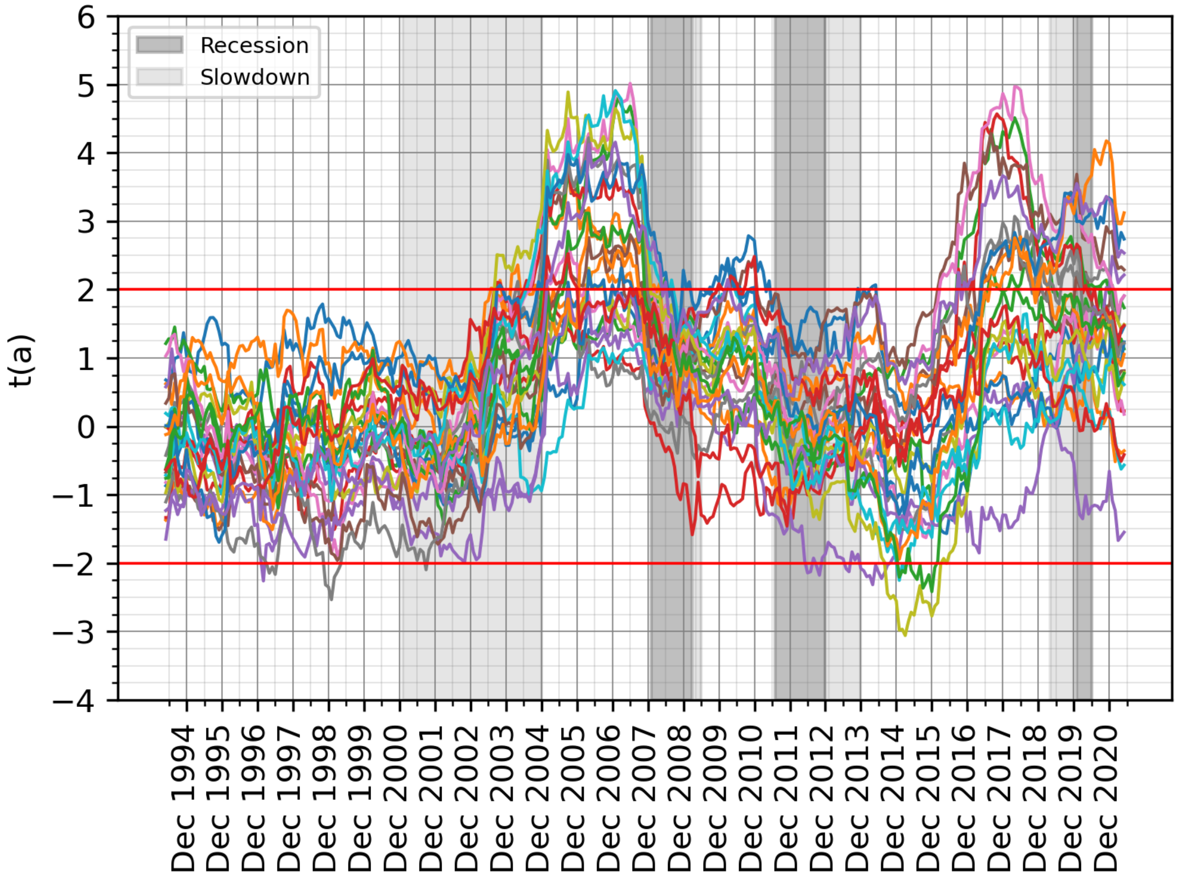
June 1989 – May 2021, 384 months. The intercept t-statistic is calculated from a sample period of five years, rolled every month until May 2021. Each date on the x-axis corresponds to the last observation of the rolling window. The RHS variables are constructed using independent 2x3 sorts on Size and each of B/M and D/E. The horizontal red lines represent the critical value for the t-test with a significance level of 5% (i.e. 2.00 and -2.00). Recession and slowdown periods are for the euro area.



Source: Based on the author's calculation run in Python and R program using data from DataStream, and Eurostat (2022a) for recession and slowdown periods.

Appendix 23: Performance of the four-factor model over time, measured by the 25 intercept t-statistic of the value-weighted Size-D/E portfolios

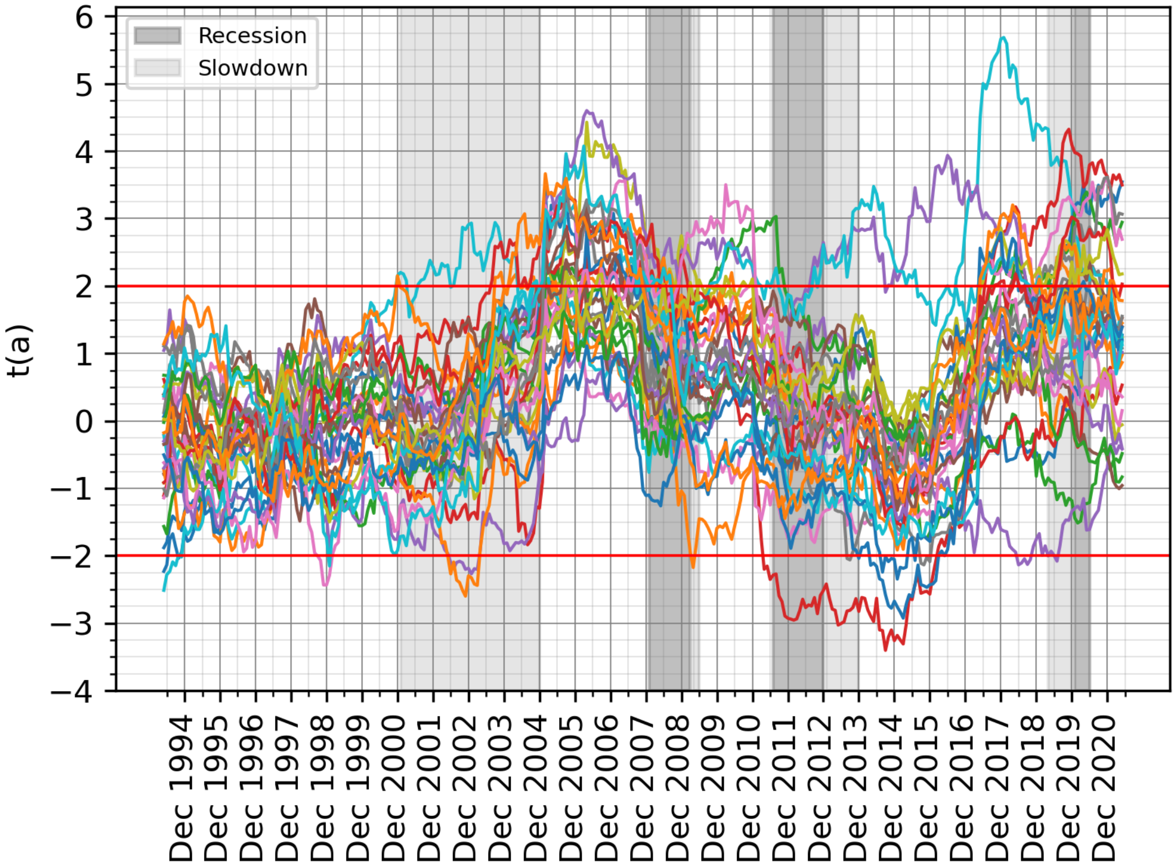
June 1989 – May 2021, 384 months. The intercept t-statistic is calculated from a sample period of five years, rolled every month until May 2021. Each date on the x-axis corresponds to the last observation of the rolling window. The RHS variables are constructed using independent 2 x 3 sorts on Size and each of B/M and D/E. The horizontal red lines represent the critical value for the t-test with a significance level of 5% (i.e. 2.00 and -2.00). Recession and slowdown periods are for the euro area.



Source: Based on the author's calculation run in Python and R program using data from DataStream, and Eurostat (2022a) for recession and slowdown periods.

Appendix 24: Performance of the four-factor model over time, measured by the 32 intercept t-statistic of the value-weighted Size-B/M-D/E portfolios

June 1989 – May 2021, 384 months. The intercept t-statistic is calculated from a sample period of five years, rolled every month until May 2021. Each date on the x-axis corresponds to the last observation of the rolling window. The RHS variables are constructed using independent 2 x 3 sorts on Size and each of B/M and D/E. The horizontal red lines represent the critical value for the t-test with a significance level of 5% (i.e. 2.00 and -2.00). Recession and slowdown periods are for the euro area.



Source: Based on the author's calculation run in Python and R program using data from DataStream, and Eurostat (2022a) for recession and slowdown periods.

REFERENCES

Adams, J., Hayunga, D., Mansi, S., Reeb, D., & Verardi, V. (2019). Identifying and treating outliers in finance. *Financial Management, 48*(2), 345–384. https://doi.org/10.1111/fima.12269

Aggarwal, R., Drake, P. P., Kobor, A., & Noronha, G. (2020). Capital structure. In *CFA Program Curriculum 2022: Level 2* (Vol. 3, pp. 5-31). CFA Institute.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance (New York)*, *23*(4), 589–609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

Arendt, C., & Dragulescu, A. (2020). xlsx: Read, Write, Format Excel 2007 and Excel 97/2000/XP/2003 Files. R package version 0.6.5. https://CRAN.R-project.org/package=xlsx

Berkin, A. L., & Swedroe, L. E. (2016). *Your complete guide to factor based investing: The way smart money invest today*. BAM Alliance Press.

Bhandari, L. C. (1988). Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence. *The Journal of Finance (New York)*, *43*(2), 507–528. https://doi.org/10.1111/j.1540-6261.1988.tb03952.x

Bryan, J., & Wickham, H. (2019). readxl: Read Excel Files. R package version 1.3.1. https://CRAN.R-project.org/package=readxl

Campbell, J. Y., Hilscher, J., & Szilagyi, J. (2008). In Search of Distress Risk. *The Journal of Finance (New York)*, *63*(6), 2899–2939. https://doi.org/10.1111/j.1540-6261.2008.01416.x

CFA Institute. (2022). *CFA Program glossary*. Retrieved April 8, 2022, https://www.cfainstitute.org/en/programs/cfa/curriculum/glossary#sort=%40glossaryz32xterm%20ascending

Ciesielski, J. T., Henry, E., & Selling, T. I. (2019). Evaluating Quality of Financial Reports. In *CFA Program Curriculum 2022: Level 2* (Vol. 2, pp. 295-377). CFA Institute.

Chou, P.-H., Ko, K.-C., & Lin, S.-J. (2010). Do relative leverage and relative distress really explain size and book-to-market anomalies? *Journal of Financial Markets (Amsterdam, Netherlands)*, *13*(1), 77–100. https://doi.org/10.1016/j.finmar.2009.07.007

Dhaliwal, D., Heitzman, S., & Zhen Li, O. (2006). Taxes, Leverage, and the Cost of Equity Capital. *Journal of Accounting Research*, *44*(4), 691–723. https://doi.org/10.1111/j.1475-679X.2006.00214.x

Eurostat (2022a). *Business Cycle Clock*. Retrieved March 17, 2022, https://ec.europa.eu/eurostat/cache/bcc/bcc.html

Eurostat (2022b). *GDP and main components (output, expenditure and income)*. Retrieved May 7, 2022, https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nama_10_gdp&lang=en

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance (New York)*, *47*(2), 427–465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3–56. https://doi.org/10.1016/0304-405X (93)90023-5

Fama, E. F., & French, K. R. (1995). Size and Book-to-Market Factors in Earnings and Returns. *The Journal of Finance (New York)*, *50*(1), 131–155. https://doi.org/10.1111/j.1540-6261.1995.tb05169.x

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010

Ferguson, M. F., & Shockley, R. L. (2003). Equilibrium “Anomalies.” *The Journal of Finance (New York)*, *58*(6), 2549–2580. https://doi.org/10.1046/j.1540-6261.2003.00615.x

Gao, P., Parsons, C. A., & Shen, J. (2018). Global Relation between Financial Distress and Equity Returns. *The Review of Financial Studies*, *31*(1), 239–277. https://doi.org/10.1093/rfs/hhx060

George, T. J., & Hwang, C.-Y. (2010). A resolution of the distress risk and leverage puzzles in the cross section of stock returns. *Journal of Financial Economics*, *96*(1), 56–79. https://doi.org/10.1016/j.jfineco.2009.11.003

Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A Test of the Efficiency of a Given Portfolio. *Econometrica*, *57*(5), 1121–1152. https://doi.org/10.2307/1913625

Griffin, J. M., & Lemmon, M. L. (2002). Book-to-Market Equity, Distress Risk, and Stock Returns. *The Journal of Finance (New York)*, *57*(5), 2317–2336. https://doi.org/10.1111/1540-6261.00497

Harris, Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., Del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature (London)*, *585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Harvey, C. R., Liu, Y., & Zhu, H. (2016). … and the Cross-Section of Expected Returns. *The Review of Financial Studies*, *29*(1), 5–68. https://doi.org/10.1093/rfs/hhv059

Hunter. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Kim, J. H. (2017). GRS.test: GRS Test for Portfolio Efficiency, Its Statistical Power Analysis, and Optimal Significance Level Calculation. R package version 1.1. https://CRAN.R-project.org/package=GRS.test

McKinney, W. (2010). Data Structures for Statistical Computing in Python. In *Proceedings of the 9th Python in Science Conference* (pp. 56-61). 10.25080/Majora-92bf1922-00a

Mirza, N., Saeed, M. S., & Rizvi, K. A. (2013). The Pricing of Size, Book to Market and Financial Leverage in Euro Stocks. *Ekonomska Istraživanja*, *26*(2), 177–190. https://doi.org/10.1080/1331677X.2013.11517615

Modigliani, F., & Miller, H. M. (1958). The Cost of Capital, Corporation Finance and the Theory of Investment. *The American Economic Review*, *48*(3), 261–297. http://www.jstor.org/stable/1809766

Obreja, I. (2013). Book-to-Market Equity, Financial Leverage, and the Cross-Section of Stock Returns. *The Review of Financial Studies*, *26*(5), 1146–1189. https://doi.org/10.1093/rfs/hhs180

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, *18*(1), 109–131. https://doi.org/10.2307/2490395

Organisation for Economic Co-operation and Development (2022). *Short-term interest rates*. Retrieved January 13, 2022, from https://doi.org/10.1787/2cc37d77-en

Ozdagli, A. K. (2012). Financial Leverage, Corporate Investment, and Stock Returns. *The Review of Financial Studies*, *25*(4), 1033–1069. https://doi.org/10.1093/rfs/hhr145

Penman, S. H., Richardson, S. A., & Tuna, İ. (2007). The Book-to-Price Effect in Stock Returns: Accounting for Leverage. *Journal of Accounting Research*, *45*(2), 427–467. https://doi.org/10.1111/j.1475-679X.2007.00240.x

Python Software Foundation (2021). Python Language Reference, version 3.10. https://www.python.org/

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Seabold, Skipper, & Perktold, J. (2010). Statsmodels: Econometric and statistical modelling with python. In *Proceedings of the 9th Python in Science Conference*.

Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance (New York)*, 19(3), 425–442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Singer, B. D. (2018). Dynamic Asset Allocation: Great Expectations. In L. Cao (Ed.), *Multi-Asset Strategies: The Future of Investment Management* (pp. 36-54). CFA Institute. https://www.cfainstitute.org/en/membership/careers/guides/multi-asset-strategies

Titman, S., & Wessels, R. (1988). The Determinants of Capital Structure Choice. *The Journal of Finance (New York)*, *43*(1), 1–19. https://doi.org/10.1111/j.1540-6261.1988.tb02585.x

Vassalou, M., & Xing, Y. (2004). Default Risk in Equity Returns. *The Journal of Finance (New York)*, *59*(2), 831–868. https://doi.org/10.1111/j.1540-6261.2004.00650.

Waskom, M. L. (2021). seaborn: statistical data visualization. *Journal of Open Source Software, 6*(60), 3021. https://doi.org/10.21105/joss.03021

EXECUTIVE SUMMARY

In this paper, we studied the influence of book value financial leverage in the pricing of euro area equities from 1989 to 2019. To do so, we used the Fama and French three-factor model (1993) as a framework, adding our leverage factor to the three original ones.

We found that a factor model including financial leverage is a better proxy for common risk factors in returns than the original Fama and French three-factor model. However, we have evidence that the value factor (HML) and the financial leverage factor (LMU) have elements in common in their stock explanatory power. Despite these similarities, neither of them should be discarded to preserve the performance of our model.

We also observed that expected common stock returns are negatively related to financial leverage, which supports George and Hwang (2010) claim that when we account for market frictions in capital structure optimisations models, firms with high distress costs select low leverage and have the greatest exposure to systematic risk. This effect dominates the strengthening effect of financial leverage on equity risk.

Moreover, we tested the robustness of our model over time. We estimated our four-factor model using rolling windows. We found that, although the performance of the model changes over time, it is always a good proxy for common risk factors in returns. Nevertheless, we did not find any statistically significant relationship between the performance of the model and any variable related to the economic situation.

We discovered something unusual in our sample: the absence of a size effect. We also noted that our model had trouble in describing small deep value portfolios returns, those giving a significant intercept. It might be of interest to further investigate these issues with the same dataset. Furthermore, we used book value of the ratio debt-to-equity to measure the financial leverage; it could be relevant to test our hypotheses by using market value instead. Finally, further research could focus on the parameters that impact the effectiveness of asset pricing models over time; and the performance of the Fama and French five-factor model (2015) plus the leverage factor.

Word count = 15,383