U.PORTO

FEP FACULDADE DE ECONOMIA
UNIVERSIDADE DO PORTO

The Quality Elasticity of Stock Prices and the Performance of Value-Quality Portfolios

João David Ribeiro Carvalho

Dissertation
Master in Finance

Supervised by
Carlos Alves

2021

Abstract

We start this research by assessing if investors pay a premium for quality. That is, we reach for evidence indicating if high-quality stocks are relatively expensive compared to low-quality stocks. We also evaluate the price paid for quality over time and try to find any indication of "flight-to-quality" events during crises, increasing the price of quality. A third hypothesis goes over the explanatory power of quality on stock prices and its evolution. A second part of the research is dedicated to factor investing. A set of portfolios is built on quality and value simultaneously, and we carefully examine their performance. As a last research topic, we add a quality-minus-junk factor to the Three-Factor Model to assess its explanatory power on our portfolios' returns.

We work with a sample of almost 6000 companies and across 41 years. After computing a quality score for every company at every period, we compute Fama-MacBeth regressions with the PBV as explained variable to solve our first hypothesis. With this procedure, we also get a quality price and an R^2 measure for each period which provides us with the information needed for the next assessments. Furthermore, using a Fama and French (1993) methodology, we assemble our value-quality portfolios.

We find that higher-quality stocks are, in general, relatively expensive. For each additional quality score unit, a stock's PBV tends to increase from 0.24 to 0.29. We also find that quality's price doesn't increase during crises and that the explanatory power of quality on stock prices has been decreasing. Regarding our value-quality portfolios, we documented a significant outperformance for high-value-high-quality portfolios. The Three-Factor Model seems to be enough to explain the majority of their returns' variation with no significant improvement when adding the quality-minus-junk factor as an explanatory variable.

Keywords: Quality factor; factor investing; value factor; quality-minus-junk.

JEL Codes: G11; G14; G19.

Resumo

Nesta pesquisa, começamos por avaliar se os investidores pagam um prémio por qualidade, ou seja, tentamos encontrar provas que apontem para que ações de qualidade elevada sejam mais caras do que ações de baixa qualidade. Também avaliamos o preço pago por qualidade ao longo do tempo para perceber se existe procura por ações de qualidade durante momentos de crise, fazendo o preço da qualidade subir. Uma terceira hipótese põe em causa o poder explicativo do fator qualidade no preço das ações. Uma segunda parte desta pesquisa é dedicada ao investimento por fatores. Um conjunto de portfólios é construído tendo por base os fatores valor e qualidade em simultâneo de forma a que possamos avaliar a performance obtida. Por último, adicionamos um fator *quality-minus-junk* ao *Three-Factor Model* e avaliamos o seu poder explicativo relativamente aos retornos dos nossos portefólios.

Trabalhamos com uma amostra com cerca de 6000 empresas ao longo de 41 anos. Depois de calcular uma pontuação baseada na qualidade de cada ação para cada período, usamos esta pontuação para calcular regressões Fama-MacBeth com o *PBV* como variável explicada. Com este procedimento, obtemos também o preço da qualidade e o coeficiente de determinação para cada período, o que nos permite realizar as duas avaliações seguintes. Adicionalmente, usamos a metodologia de Fama e French (1993) para construir os portefólios com base nos fatores valor e qualidade.

Concluímos que ações de qualidade são geralmente mais caras. Para cada unidade adicional na pontuação da qualidade, o *PBV* tende a aumentar de 0.24 a 0.29. Também concluímos que o preço da qualidade não aumenta durante momentos de crise e que o poder explicativo da qualidade nos preços das ações tem vindo a diminuir. No que diz respeito aos portefólios, documentámos uma performance muito significativa para portefólios formados com base em ações de elevada qualidade e valor. O *Three-Factor Model* parece ser suficiente para explicar a maioria da variação dos seus retornos, não havendo uma melhoria significativa quando o fator *quality-minus-junk* é adicionado como variável explicativa.

Palavras-chave: Fator qualidade; investimento por fatores; fator valor; *quality-minus-junk*.

Códigos JEL: G11; G14; G19.

Index

Index of Tables

Index of Charts

1. Introduction

Academics and researchers in the Finance realm have tried to explain simple concepts like the price, returns, and risk of a stock and other financial assets for decades. From one perspective to the next, there seems to be always something to be added, and that's probably because these simple concepts have an inherent complex web of forces to be accounted for that push and pull in different directions which ultimately are reflections of the market's decisions, a market composed by human beings and their opinions, among other things.

One topic that has been in development for a very long time is the study of factors. Factors are characteristics that help to explain the returns of stocks, for instance. We can trace back the study of factors at least to the 1960s with the introduction of the Capital Asset Pricing Model (CAPM) and recognition of a broad market risk exposure from individual securities (Lintner, 1965; Mossin, 1966; Sharpe, 1964; Treynor, 1961). Among others, the value factor, size factor, and momentum factor are also three of the most studied and well-known characteristics of a stock that help to explain how its price behaves in the long term (Banz, 1981; Basu, 1977; Fama & French, 1992,1993; Jegadeesh & Titman, 1993). This dissertation will focus on the quality factor which started to be studied and described very recently and with little to no scientific literature covering it and its implications on prices and returns.

This research is divided into two main parts, which are also subdivided. The first one revolves around the impact of quality on stock prices and which type of relation there is between these two variables. We will assess how different levels of quality attributable to a stock explain a different price level, in other words, if higher-quality stocks demand higher prices, as some literature documents, for example, Asness et al. (2019) and Berg (2020). We believe this information is very important because it can provide interesting insights into factor investing literature. Additionally, we will study the evolution of the price given to quality over time and specifically test if there is any tendency of this quality price to appreciate during moments of crisis and market downturns. To this day, only Asness et al. (2019) present an analysis similar to this. If confirmed, this means that investors could take refuge on high-quality stocks in moments of uncertainty to protect their investments. Finally, we will do one last test to study the evolution of the explanatory power of quality on stock prices. Asness et al. (2019) and Berg (2020) conclude that this explanatory power is exceptionally low. With our analysis, we will assess if this has always been the case or if there is any relevant evolution over time.

The second part of this research goes over the performance of portfolios built on value and quality, similar to Fama and French (1993). There is evidence (Kozlov & Petajisto, 2013; Novy-Marx, 2013) suggesting that combining the value effect and the quality effect would result in good risk-adjusted performance, and based on this existent conclusion, we will search for empirical evidence within our sample. Additionally, we will evaluate if a Three-Factor Model with the addition of a quality-minus-junk (QMJ) factor, which captures the premium paid for high-quality stocks, would improve the explanatory power of a simple Three-Factor Model when regressing it against the returns obtained by the portfolios built on value and quality. This analysis is very important because if proven worthy of it, a strategy like investing in high-quality and simultaneously high-value stocks would be very easy to implement. It would fill another gap in the literature. Furthermore, we evaluate if the QMJ factor would improve the Three-Factor Model. Additionally, we build another QMJ factor to see if there is any meaningful difference at the end of the analysis.

To do all this research, we count on a large sample of almost 6000 companies across 41 years divided into quarters. Quality, as a subjective term, needs to be defined and quantified. For this, we used the definition and methodology from Asness et al. (2019), which divide quality into three main characteristics: profitability, safety, and growth. Each characteristic is measured and quantified using several proxies and culminate on a final, objective, quality score. The advantage of looking at things with this perspective comes from the way a quality score is built. It is a vague measure that grabs different characteristics that, presumably, most investors would seek and can be used to rank a set of stocks while covering several attributes and offering more information. This information is also condensed in a way that simplifies the analysis. It would be harder to pick one stock over another for their profitability, safety, and growth separately instead of their overall quality score. Having a quality score for every company at every period, we used a Fama-Macbeth (Fama & MacBeth, 1973) procedure to perform our econometric tests for the impact of the quality score on prices. Because of the nature of a Fama-MacBeth procedure, we get one regression for each period which allow us to obtain one regression coefficient associated with the quality score (which we use as quality's price) and an R^2 measure per period. These two measures allow us to test the reaction of quality prices to crises and the evolution of the explanatory power of quality on prices.

We base our initial methodology for the second part of the analysis on Fama and French (1993). Instead of building the portfolios on size and value, as these authors did, we make them on quality and value. In the end, we get nine different portfolios, each with a different combination of levels of quality and value, that are reset and rebalanced every year. We calculate the returns for each year to later present the comparisons with our benchmark and the risk-adjusted performance. Additionally, on top of testing the explanatory power of a simple Three-Factor Model on the excess returns of each portfolio and compare the results to the situation in which we add a QMJ factor to the explanatory variables, we also build a different QMJ factor to perform the same tests. The original QMJ factor from Asness et al. (2019) is built on quality and size. We build it on quality and value, adapting it in the best way possible for our analysis.

What we found initially confirms what is stated in previous literature: quality pushes up the prices of stocks, but the explanatory power of quality is overall very low. Regarding our second question, we found no positive relationship between the price of quality and the existence of a crisis. The relationship that seems to exist is negative, showing that quality follows the market more often than not. Furthermore, we found a significant decrease in the explanatory power of quality on stock prices over time.

The second part of the results reveals that portfolios built on value and quality outperform any other combination, dwarfing the returns of the S&P 500 during the period between 1986 and 2020. The risk-adjusted returns are also significantly improved. Concluding the results, we find no significant improvement to the Three-Factor Model when adding the QMJ factor to explain the returns of these portfolios. Still, the improvement is better when the version of the QMJ we created is used in place of the original one.

The remaining of this dissertation is organized as follows: Section 2 contains our literature review divided into three subsections. Section 3 introduces and synthesizes our research hypotheses. Then we present a detailed description of our methodology in Section 4. Section 5 is about our sample. The last core section, 6, fully exposes our results organized by subsections dedicated to each hypothesis, and Section 7 concludes.

2. Literature Review

2.1. Factors

Factor analysis traces back to at least the 1960s when the Capital Asset Pricing Model (CAPM) was introduced by Treynor (1962), Sharpe (1964), Lintner (1965), and Mossin (1966). This theoretical model was built on the premise that securities' returns were a linear function of the excess returns of the market, allowing an investor to predict the returns that would be appropriate for the level of risk assumed. Being the pioneer and very simple to apply, the CAPM is still one of the most well-known asset pricing models there is. However, this simplicity also brought several critics, as Fama and French (1992), that based their criticisms on the fact that CAPM left a big part of securities' returns unexplained, which ended up in low-quality empirical return observations. CAPM is formulated as follows:

$$R_i = R_f + \beta_i(R_M - R_f) \tag{2.1}$$

Being R_i the returns of security i, R_f the risk-free rate and R_M the returns of a market portfolio. The coefficient β_i represents the sensitivity of asset i to market movements and is often used as a measure of risk.

One alternative to the CAPM is the Arbitrage Pricing Theory (APT) introduced by Ross (1976). The APT asset pricing model provides a larger amount of freedom in terms of conceptualizing what determines a security's returns. In addition, it builds the framework that allows to include in the analysis any number of specific or systematic factors, not only the market risk premium, by assuming that the security's returns have some level of sensitivity to each factor.

$$R_i = R_f + \beta_{1,i}F_{1,i} + \beta_{2,i}F_{2,i} + \cdots + \beta_{k,i}F_{k,i} \tag{2.2}$$

Formula 2.2 generally describes the model behind the APT in which R_i and R_f remain the same as in formula 2.1. The F variables represent the risk premium of each factor k included. The β coefficients, as in CAPM, represent the sensitivity of the security to whatever factor they are associated with. Ross (1976) does not specify which variables should be used as factors, and that is left to be decided by whoever uses the model.

After the surge of these models and theories, the notion of factors as specific characteristics of a security that explained part of its returns started to become popular within the financial mainstream. Each factor is associated with exposure to different kinds of systematic risk. Examples of this are value, size, and momentum.

The size effect was first described by Banz (1981). Size aims to capture the excess return of a stock explained by the single fact of being a relatively low-sized company in terms of market capitalization. According to Banz (1981), firms with a relatively low market cap outperform firms with bigger market caps. Fama and French (1992, 1993) suggested that the main reason for this anomaly was the additional market risk that small stocks had compared to bigger stocks. As expected, several other authors tried to explain this market phenomenon and came up with different reasons: the spread between high- and low-quality corporate bonds (Chan, Chen & Hsieh, 1985), financial distress (Chan & Chen, 1991), liquidity (Amihud, 2002), default risk (Vassalou & Xing, 2004) and information uncertainty (Zhang, 2006).

Although being a concept firstly introduced by Basu (1977), value investing is also broadly associated with Fama and French and their Three-Factor model. The value factor brought to the table the idea that value stocks, characterized by low market-to-book ratios, usually outperform growth stocks that opposingly have high market-to-book ratios. Essentially, value stocks can be described as "cheap stocks" and growth as "expensive stocks", assuming the price is a ratio between the market value of equity and book value of equity. This can be explained by the fact that growth stocks are usually famous stocks that are overbought and commonly deteriorate the expected earnings for the investors. Value stocks represent the other side of the spectrum, being companies with low expectations of growth and oversold stocks that are somewhat forgotten by the majority of the market. This reveals to be important to explain the excess returns of stocks.

A third very well-known factor is the momentum factor. Momentum stands for something that keeps moving. In this case, the momentum effect describes the overperformance of stocks explained by past returns, which is why it is also commonly referred to as the "persistence factor". This effect was described for the first time by Jegadeesh and Titman (1993). These authors presented evidence that it was worth buying winners and selling losers for a holding period from 3 to 12 months. They tend, on average, to continue to be winners and losers, respectively. A vast amount of literature seems to confirm these findings in other scenarios like different countries and periods.

According to Asness, Frazzini, and Pedersen (2019) – the most influential paper for our analysis – quality is a characteristic investors are willing to pay a premium. Quality stocks are defined by having three essential characteristics, namely:

- Profitability;
- Safety;
- Growth.

Profitability is by itself a vague term that can be defined and measured by multiple different variables. Asness et al. (2019) use several profitability measures such as gross profits, margins, earnings, cash flows, etc. Safety represents the opposite of risk and is measured in two different perspectives: from a market standpoint (market beta and stock price volatility) and also from a fundamental point-of-view (credit risk, leverage, and earnings stability). Lastly, growth is simply the growth that has been seen in a company's profitability measures.

There are two additional, less studied factors we want to introduce: volatility and yield. Blitz and Vliet (2007) argue that stocks with lower volatility present higher risk-adjusted returns, which is the essential idea behind the volatility effect. Of course, volatility is a measure of risk in most cases, and with risk, investors expect higher returns, so that this idea can seem contradictory, but it should be noted that the authors here talk about risk-adjusted returns and not absolute returns.

Dividends are a heavily reported and studied in the literature, with opinions and theories shooting in every direction. For example, Litzenberger and Ramaswamy (1979), Blume (1980), and Fama and French (1988) are papers that associate dividends and stock performance by describing and analyzing the yield factor introducing another good measure of explanation of stock returns. This factor describes the outperformance of high dividend yield stocks.

After introducing the most common factors, it's time to go over a last asset pricing model, the most connected with this dissertation. The Three-Factor Model was presented by Fama and French (1992), and as the APT, it was presented as an alternative to the CAPM. As the name implies, it was based on three factors already defined above: market risk premium, value, and size. Fama and French (1993) claim and demonstrate that this model presents a strong explanatory power of stock returns.

$$R_{i,t} - R_{f,t} = \alpha_{i,t} + \beta_1(R_{M,t} - R_{f,t}) + \beta_2 SMB_t + \beta_3 HML_t + \varepsilon_{i,t} \qquad 2.3$$

Formula 2.3 shows the general expression that describes the Three-Factor Model. Regarding new notation, SMB represents the size factor, which captures the premium observed on small-sized stocks, and HML is the factor that captures the value premium. The error term is represented by ε.

Compared to the CAPM, the Three-Factor Model differs by analyzing excess return and not having the risk-free rate as the intercept. Furthermore, the addition of the size and value factors empirically improves the explanatory and predicting power of the model.

2.2. Earnings Quality

The relation between the quality factor and a concept such as earnings quality is intrinsically linked for the clear reason that quality, using the definition described above from Asness et al. (2019), includes earnings as one of its components. Furthermore, the growth of earnings and their stability can also be considered to be characteristics of their quality, so we are dealing with very similar concepts. Given this, we will find support on papers that study earnings quality as much as we find their information and findings useful for what we are trying to accomplish.

Earnings quality is addressed and studied by plenty of authors and for a very long time – at least since Graham and Dodd's book "Security Analysis" from 1934 in which the authors propose a model to value equity, based on earnings per share times a measure of quality (Dechow, Ge & Schrand, 2010). With this in mind, we will concentrate our analysis on more recent papers and other publications for two reasons: the data used is more updated to recent times, and usually, the recent works already incorporate methodology and important insights from the earlier ones, so this way, we will capture the majority of the literature efficiently.

How could we measure earnings quality? It is a very subjective question, given quality is something that ultimately depends on our interpretations and goals. Dechow et al. (2010) study several different variables that are commonly used to proxy the quality of earnings. The authors analyze over 300 papers that contain earnings quality proxied by some variable and for some reason. The proxies extracted from the papers are divided into three major categories:

1. Properties of earnings: earnings smoothness, accruals, earnings stability, and target beating, which assumes that a difference from a specific target is taken as earnings management which deteriorates earnings quality;
2. Investor responsiveness to earnings: earnings response coefficient or the R^2 of a simple regression explaining the relation between earnings and the returns of a stock;
3. External indicators of earnings misstatements: this includes indicators for errors and earnings management such as restatements and internal control deficiencies.

There is no earnings quality measure that is superior for all cases and models. Using a methodology from an earlier paper (Cronbach & Meehl, 1955) the authors reach this conclusion after classifying each paper on their analysis into two different groups – papers that work on the determinants of the earnings quality proxy, and on the other side, papers that look more into the consequences. Essentially, the first group includes research that considers the quality of earnings an independent variable and seeks what other variables explain its distribution. On the other hand, the second group includes papers with earnings quality as an explanatory variable to study their impact on a specific outcome.

Dechow et al. (2010) reach two fundamental conclusions, and none of them is completely specific and objective since this paper is not trying to reach any specific goal. Instead, it is reviewing as many earnings' quality proxies as possible, and in this case, "better" is subjective to every single paper or investigation on this matter, so every researcher must find what is most useful for themselves. In spite of it, the first conclusion is that several earnings quality proxies – the ones that involve earnings – are influenced not only by a firm's performance but also by the measurement of this performance. This happens with these specific proxies because they are based on earnings reported on an accrual basis. That is, frequently, the literature tends to ignore the differences between performance and performance measurement. The second conclusion states that these proxies are not equally influenced by performance and performance measurement, making them all different in terms of what they are capturing.

Sloan (1996) provides good insights on earnings quality and is also considered to be the first paper to dive into this specific topic, despite several other authors mentioning it before. The author divides earnings into two components – accruals and cash flows – and studies what each component brings in terms of information and how that affects stock prices. Note that accruals are the variation in current assets that does not have to do with variations in cash minus both depreciation and the variation of current liabilities, not accounting for the variation in debt and income taxes payable. The main findings are that when the accruals component is more relevant, earnings stability for the future is poor, and, contrarily, when the cash flow component stands on top, earnings tend to thrive in the future. Given this, and considering that investors tend to overweight on earnings as a whole, the author also finds that stocks with a high presence of accrual properties have negative abnormal returns in the future and vice versa for stocks with high levels of cash flow component in earnings.

Bender and Nielsen (2013) studied if earnings quality, proxied by the accruals component of earnings, was still significant to explain if stocks with high-quality earnings would generally outperform stocks with low-quality earnings. This is important because it verifies whether these observations still hold and are not an anomaly found on previous samples by chance. Another important question these authors try to answer is whether stocks with high-quality earnings are generally different in terms of risk compared to stocks with low-quality earnings. Interestingly enough, it was found that using earnings quality as a strategy stopped providing good results in the mid-2000s and bounced back at the end of 2008, suggesting the existence of a "flight to quality" after the great financial crisis. Regarding the other hypothesis stated above, the authors found no evidence suggesting that earnings quality is a good risk factor. According to them, it indicates that earnings quality may be a good alpha factor. Kozlov and Petajisto (2013) works on a similar framework but adds developed global markets instead of analyzing just the US market. The authors found that a strategy that goes long on stocks with high-quality earnings and shorts stocks with low-quality earnings generates abnormal returns and better risk-adjusted returns than the market. A significant finding is that there is a negative correlation (-0.32) between this kind of portfolio and a value portfolio, which creates opportunities for diversification and improved risk-adjusted returns – this was found in other papers that will be included later.

Novy-Marx (2013) states that the best proxy for profitability is gross profits to assets and the reason being that earnings are economic profits that are calculated by accounting rules that consider any investment as a cost and if this investment is made with the idea to bring higher returns later, then earnings begin to look like a bad way to measure profitability. The main finding of this paper is that profitability has a considerable explanatory power on returns – as high as the power for book-to-market measures (the value factor) – and that by also controlling for profitability, value strategies have their performance enhanced. It was also found that profitable companies have stocks that generate higher returns than unprofitable firms. At the same time, profitable firms tend to have lower book-to-market ratios and are bigger in market capitalization, which goes against value and size anomalies (this is similar to the finding presented by Kozlov and Petajisto (2013)). Given this, the profitability strategy presents a good opportunity for value investors to be exposed to potential additional returns coming from the profitability premium while at the same time not being necessarily exposed to additional risks and that's only possible because profitability strategies essentially go long on growth stocks.

2.3. Quality Factor

Trying to answer and study one of the conclusions he presents in his previous paper about profitability, Novy-Marx (2014) describes quality investing as being a different perspective of value investing since buying a stock below its theoretical value is virtually the same as buying a high-quality stock without paying a premium for that additional quality. His objective is to find the best quality measure that will help investors build better portfolios considering the synergy between quality and value – quality stocks at a reasonable price - and avoid investing in quality stocks with everything priced in. To find the answer to this riddle, the author combines the usual value strategies metrics with several different quality measures to extract which one gives the best overall results. Grantham's notion of quality stocks (Grantham, 2004), Greenblatt's magic formula (Greenblatt, 2010), Sloan's accrual-based measure (Sloan, 1996), Piotroski's F-score (Piotroski, 2000), and Novy-Marx's gross profitability (Novy-Marx, 2013) are the measures used to represent the quality factor, and that will be compared against each other to assess which one pairs better with a value investing strategy. The main results show that there is a slight difference when we compare investments in small-cap stocks and large-cap stocks: for small-cap stocks, the strategies that revealed to be better were not only the ones that included Piotroski's F-score but also strategies that used Novy-Marx's gross profits to assets measure to represent quality; in the large-cap world, only gross profitability showed outperformance against all others. To conclude, an additional finding from Novy-Marx (2014) that we found interesting is that adding the momentum effect to a value and quality strategy could bring even more significant improvements to the strategy in the form of reduced transaction costs and other inefficiencies of the initial value-quality strategy.

Novy-Marx (2012) does a very good job adding to previous literature on what quality investing specifically is and how we can identify a quality strategy and choose the best among several. The author describes how to measure quality with different measures (Graham's quality criteria, Grantham's high return, Greenblatt's return on invested capital, Sloan's measure of earnings quality, Piotroski's F-score, and Novy-Marx's gross profitability approach). The main conclusions are that his measure of quality seems to be the one that outperforms more, most of the time. This is especially persistent in the large-cap stock market (Novy-Marx, 2014) and for long-only investors. The argument for this last conclusion comes from the fact that long-short investors have the freedom to maximize their strategy's risk-to-reward ratio through leverage and by separating opportunity and exposure decisions. Long-only investors can't make this separation and may want to get greater exposure to opportunities with lower risk-to-reward ratios than other existent opportunities. Being so exposed to market factors, long-only investors find it difficult to joint value and quality strategies.

To finalize this literature review, we want to go into more detail on the paper that is the one with the greater presence and inspiration through this dissertation's development. Asness et al. (2019) initiate by introducing the term quality and how quality should be defined from a financial perspective: a mix between profitability, growth, and safety. With this definition already presented above, a dynamic asset pricing model was developed. This model assumes time-changing variables, hence dynamic, and a linear relationship between the price and each component of quality (profitability, growth, and safety). It aims at identifying which stock characteristics command a higher price and how important they are at determining the price. To assess this, the model is run on a series including data of 25 developed countries and for multiple decades (five for the US and almost three for every other country). The findings are that high-quality stocks are generally associated with higher prices (given by price-to-book ratio) but in the initial regression, only considering the quality score, quality and its components explain only 9% of the cross-sectional variation of prices. After controlling for other factors like size, momentum, industry-fixed, country-fixed, and firm-fixed effects the R^2 increases but only reaches values close to 50%, leaving a big chunk of the price distribution unexplained. The low explanatory power of quality on the price level of stocks puzzled the authors, and they proceeded to give three possible explanations for this:

- The prices are based on more complex quality characteristics than the ones used for the test;
- The quality factor is related with risk factors not captured by the quality measure;
- The prices fail to reflect quality due to things like behavioral finance and market constraints.

To further study the quality factor, including trying to verify any of the hypotheses above, the authors also built a quality-minus-junk factor (QMJ). The methodology used to build this factor was precisely the same used by Fama and French (1993) – buy the top 30% stocks distributed by the quality score (explained in the methodology section) and short the bottom 30%.

QMJ portfolios are divided into large-cap and small-cap stocks and present positive returns in the vast majority of countries included in the test. The results are especially good during market downturns which suggests a "flight-to-quality" event when the markets are weak which appears to be consistent with the third hypothesis suggested to explain the low explanatory power of quality on prices. This was visible because the price of quality varies over time, reaching its lowest at the peak of the internet bubble in the late 1990's/2000 and raising after the subsequent bust and in other bear markets such as the financial crisis in 2008.

3. Research Topic and Hypotheses

The literature that studies specifically the quality factor is particularly scarce compared to other factors or even other unrelated topics within finance. In this chapter, there is a detailed description of the main topics that will be empirically studied further down and what is being added to the literature available at the moment.

Logic or intuition would most likely make one think that quality stocks come with a premium in the form of a higher price. And this was the description of quality stocks given by Asness et al. (2019) – quality stocks are stocks that demand higher price premiums. The characteristics of quality are such that they attract investors and lead them to be willing to pay more for stocks with higher levels of quality. This way, following up on the findings of Sloan (1996), Dechow et al. (2010), Bender and Nielsen (2013), Novy-Marx (2012,2013,2014), and specially Asness et al. (2019) and Berg (2020), we will assess the existence of a relation between stock prices and the quality factor and the dimension of this relationship within our sample. To what extent is quality relevant to determine how cheap or expensive a stock should be, it will be questioned. The intent is to validate the results from Asness et al. (2019) and Berg (2020), compare the results, and point any new findings.

The evolution of quality's price over time will also be studied. According to Asness et al. (2019), the price of quality varies over time, being particularly expensive during market downturns and crises. This is a hypothesis that is going to be tested and verified. Furthermore, being exactly in the middle of an economic and social crisis at the writing of this dissertation, it will also be interesting to assess the impact of this particular crisis on the price of quality. This last topic will be especially interesting given that initially, in this crisis, there was a big deflationary crash on every market, particularly on the stock market, that potentially had a sudden impact on things like the price of quality. This also adds to the literature due to the recent nature of the crisis.

To conclude this part of the analysis, and taking out more information from the data already treated for previous topics, we will observe the evolution of the explanatory power of quality on prices. In other words, we will assess if there is any meaningful change in quality's ability to explain the prices of stocks over time. Asness et al. (2019) and Berg (2020) claim this explanatory power is overwhelmingly low. To add to this analysis, we will assess if this was always the case, providing new information on this phenomenon.

Now taking a different route, we extend on some findings from Kozlov and Petajisto (2013) and Novy-Marx (2013) – respectively, the negative correlation between quality portfolios and value portfolios and the enhanced performance of value portfolios when also controlled by profitability – which open up some interesting research ideas. We also base this part of the analysis on Fama and French (1993), which, as introduced in Section 2, developed a famous and worldwide accepted asset pricing model commonly referenced as the Three-Factor Model. They construct portfolios based on value and size and study the performance and exposure of these portfolios to different factors. Based on this, we will do a similar assessment, but this time with portfolios built on value and quality. The main focus here is to evaluate the overall performance and risk-adjusted performance of these portfolios. We also assess their exposure to the risk factors proposed by Fama and French (1993), as well as the quality-minus-junk factor (QMJ) created by Asness et al. (2019). Furthermore, we adapt this last factor to suit our analysis the best way possible, and this process will be further explained in the methodology section. In comparison to available literature, this part of our analysis is new in that it brings empirical verification of the hypothesis that combining value and quality strategies brings improved performance. On top of that, we add the QMJ factor to the Three-Factor Model and create a new QMJ factor built on slightly different parameters.

To summarize, we can describe our research topics with the following research hypotheses:

H1: Quality is a characteristic for which investors are willing to pay a premium.

H2: The price of quality varies over time, being more expensive in crises and times of market distress.

H3: There is a relevant evolution of the explanatory power of quality on stock prices.

H4: Portfolios built on value and quality present relevant abnormal risk-adjusted returns.

H5: The Three-Factor Model with the addition of the QMJ factor has solid explanatory power on the excess returns of value-quality portfolios.

The order in which this analysis is presented is not random. In terms of the data treatment, this is the order in which the analysis had to take place, so we believed that it is best suited to present the analysis and results in the same order so everything, even methodology, makes the most amount of sense for the reader.

4. Methodology

The topics that we aim to investigate are highly empirical. They demand high amounts of data regarding prices, earnings, revenues, assets, leverage, and the list goes on. With all this information, things like volatility and a number of ratios had to be derived later.

4.1. Quality Measure

Let's start by explaining how to measure quality and turn that characteristic into a number. Quality is a very subjective term. With the model developed by Asness et al. (2019), we can transform an otherwise subjective term into a number, being able to rank different stocks for their quality score. We calculated our quality scores using the same methodology as the aforementioned paper, with slight differences mainly due to limitations regarding our access to meaningful amounts of data.

Firstly, we divide quality into three main characteristics: profitability, growth, and safety. For each of these, we set several proxies that represent them well and are able to capture and rank every company accordingly.

For profitability, we chose four proxies, namely, return-on-equity (ROE), return-on-assets (ROA), gross income over assets (GIOA), and EBIT over Assets (EOA). The choice of these four variables was mainly based on previous literature and the availability of data for a considerable number of companies and through a fairly long timespan. ROE, ROA, and GIOA are measures also used by Asness et al. (2019), and although gross income was available for way fewer observations than the other variables, we wanted to make sure to include it since Novy-Marx (2013) points gross profits over assets as the single best measure of profitability available. For this, we didn't want to leave it out, given we still could retain a more-than-acceptable sample. To capture operational profitability, we also chose to include EOA in this analysis. The next step was to winsorize every variable at the 1st and 99th percentile for every period to tame outliers without compromising more observations at this stage. After that, we standardize the observations using z-scores in order to have an accurate way of summing and operating with values of different variables. To do this, we use the following formula:

$$z_{x,t} = \frac{r_{x,t,i} - \mu_{r_{x,t}}}{\sigma_{r_{x,t}}} \tag{4.1}$$

Being x the variable, t the time period and i the specific company. By subtracting the time period mean of a specific variable to the observed figure of that variable for a specific company i and after dividing it for the standard deviation of the same distribution, this procedure will give how many standard deviations a given observation is from the mean (calculated for each period and for every company with an available observation on that period) which is ideal for comparing very different measures. The final result is a distribution with a mean equal to 0 and a standard deviation equal to 1 for each time period.

The calculation of the profitability score uses the following formula:

$$Profitability_{i,t} = z(z_{ROE_{i,t}} + z_{ROA_{i,t}} + z_{GIOA_{i,t}} + z_{EOA_{i,t}}) \tag{4.2}$$

Essentially, we sum up every score obtained for each profitability proxy for company i at time t and with that amount for every company, we do the same process in order to standardize the profitability score and obtain this score for every company and available time period.

Regarding growth, we use the same raw data as for profitability, but we first calculate the year-over-year rate of change. This way, we obtain the growth rate for all four profitability measures, which will serve as proxies for company growth. The data treatment is the same as for profitability. The formula for growth is the following:

$$Growth_{i,t} = z(z_{\Delta ROE_{i,t}} + z_{\Delta ROA_{i,t}} + z_{\Delta GIOA_{i,t}} + z_{\Delta EOA_{i,t}}) \tag{4.3}$$

For the last component of quality – safety – the score is composed of the historical beta (HB), leverage (LEV), and stock price volatility (VAR) given by the variance of the stock price. Asness et al. (2019) also include bankruptcy risk as a measure for safety using Ohlson's O-score (Ohlson, 1980) and Altman's Z-score (Altman, 1968). We chose not to include these measures for the sake of simplification and availability of data. Every variable is standardized the same way as explained before. It is important to note that these measures proxy for the opposite of safety so, to obtain the safety score, we have to add a minus sign on the final score. We calculate the safety score using the formula below:

$$Safety_{i,t} = -z(z_{HB_{i,t}} + z_{LEV_{i,t}} + z_{VAR_{i,t}}) \qquad 4.4$$

To finally obtain the quality score, the procedure is to sum the scores of every quality component and standardize that amount for every period and every company all over again, as follows:

$$Quality_{i,t} = z(Profitability_{i,t} + Growth_{i,t} + Safety_{i,t}) \qquad 4.5$$

With a quality score for every company at every available period, we have reached the starting point of the actual analysis and exploration that seeks validation or rejection for every hypothesis we proposed to test.

4.2. Hypotheses Testing

H1

H1 is about testing the effect of quality on prices. In other words, we will study if there is a significant price difference between high-quality and low-quality stocks. To perform this test accordingly to more modern methodologies and approaches, we would use a panel data analysis methodology, given we have a large cross-section of stocks across a fairly long interval of time. Despite this, working with unbalanced panels (panels that don't have every single observation for every cross-sectional unit at every period) is particularly hard, especially to perform residuals testing. To overcome this problem and, also important, to keep faithful to the methodology from Asness et al. (2019) and Berg (2020), we used the Fama and MacBeth procedure (Fama & MacBeth, 1973) to perform our first econometric regressions. This is a simple two-step procedure that allows cross-sectional and time-series analysis to be combined. On a usual Fama-MacBeth procedure, the first step consists of regressing each cross-sectional unit against a set of explanatory variables. Usually, the dependent variable is the returns (R) of a number of stocks, and the explanatory variables are a set of risk factors (F). Consequently, for n stocks, there will be n time-series regressions and $n \times (m+1)$ regression coefficients resulting from this first step, being m the number of risk factors included. We can express this first step as follows:

$$\begin{gathered} R_{1,t} = \alpha_1 + \beta_{1,F_1} F_{1,t} + \beta_{1,F_2} F_{2,t} + \cdots + \beta_{1,F_m} F_{m,t} + \varepsilon_{1,t} \\ R_{2,t} = \alpha_2 + \beta_{2,F_1} F_{1,t} + \beta_{2,F_2} F_{2,t} + \cdots + \beta_{2,F_m} F_{m,t} + \varepsilon_{2,t} \\ \vdots \\ R_{n,t} = \alpha_n + \beta_{n,F_1} F_{1,t} + \beta_{n,F_2} F_{2,t} + \cdots + \beta_{n,F_m} F_{m,t} + \varepsilon_{n,t} \end{gathered} \tag{4.6}$$

After this, the second step consists in following a similar process but this time calculating a cross-sectional regression for every period and using the same dependent variable as before but against the betas obtained on the previous step. This time, being T the number of periods, the result will be T cross-sectional regressions and $T \times (m+1)$ new regression coefficients. The second step summarized in formulas is the following:

$$\begin{gathered} R_{i,1} = \gamma_{1,0} + \gamma_{1,1} \hat{\beta}_{i,F_1} + \gamma_{1,2} \hat{\beta}_{i,F_2} + \cdots + \gamma_{1,m} \hat{\beta}_{i,F_m} + \varepsilon_{i,1} \\ R_{i,2} = \gamma_{2,0} + \gamma_{2,1} \hat{\beta}_{i,F_1} + \gamma_{2,2} \hat{\beta}_{i,F_2} + \cdots + \gamma_{2,m} \hat{\beta}_{i,F_m} + \varepsilon_{i,2} \\ \vdots \\ R_{i,T} = \gamma_{T,0} + \gamma_{T,1} \hat{\beta}_{i,F_1} + \gamma_{T,2} \hat{\beta}_{i,F_2} + \cdots + \gamma_{T,m} \hat{\beta}_{i,F_m} + \varepsilon_{i,T} \end{gathered} \tag{4.7}$$

(IHS Eviews, 2014)

Ending this second step, we have the risk premium for each factor at time t in the form of γ.

What we just described is the usual Fama-MacBeth two-step procedure. In our analysis, following Asness et al. (2019) and Berg (2020), we should skip the first step because we already have a quality score for every cross-sectional unit across our timespan. The first beta estimations are the quality scores calculated as explained. We jumped up directly to the second step regressing each company's PBV against its quality score, developing a cross-sectional regression for each time period.

On different models aimed at estimating the same thing, we have added several controls. Firstly, we have only added a dummy variable that controls for industry effects. Different industries may inherently have different standards for the variables used as proxies in our quality score. On a third attempt, we have left the industry controls and added controls for size, momentum, earnings volatility, firm age, and a dummy variable that has a value of 1 when the firm is a dividend payer. We proxy size and momentum with market capitalization and past 12-month returns, respectively. Earnings volatility is given by the firm's ROE variance. Except for the dummy variables, every variable is standardized as previously explained. The choice of controls and respective proxies is inspired by the models from Asness et al. (2019).

Summarizing, H1 is tested with a set of Fama-MacBeth regressions through which we can assess quantitatively what is the impact of quality on the price of a stock, here represented by PBV.

H2

To obtain the price of quality over time, Asness et al. (2019) used the beta coefficient associated with the quality score for every Fama-MacBeth regression. Given that the procedure results in a regression for each time period, the quality score coefficient varies across the entire timespan. This way, we will use the coefficients that resulted from our analysis, which were also a result of our sample of data, to obtain the evolution of the price of quality across time and specifically how it performed during times of crisis.

To perform a statistical test, we run a simple regression where the variation within the series of quality prices is explained solely by a dummy variable that has assigned a value of 1 whenever the specific period is considered a period of crisis and 0 otherwise.

The regression can be expressed as follows:

$$Q_t = \alpha + \beta_t D_t + \varepsilon_t \qquad 4.8$$

Where Q represents the quality price and D the dummy variable.

For H2 to be confirmed, not only β_t has to be positive, showing that there is evidence of a higher price for quality in periods of crisis, but also it has to be statistically significant.

H3

A study about the evolution of the explanatory power of quality on prices is not something we could find in any previous literature. Given this, we found it interesting to see if there is any relevant evolution throughout the years. Here, we express the explanatory power of quality in the form of R^2. In this particular case, the R^2 of every Fama-MacBeth regression computed earlier. With a regression for each time period, it becomes easy to develop a way to assess and observe the evolution of the explanatory power of quality on prices.

On top of observing this evolution, we also added a statistical analysis that goes over this question by computing a regression with the adjusted R^2 of every Fama-MacBeth regression that was run earlier as our explained variable. In this case, the independent variable only serves the purpose of capturing the passage of time by having assigned a value of 1 for the first period, 2 for the second, and so on, until the last period, in which the variable is assigned the value of 159.

The regression is described by the following formula:

$$Adj\ R_t^2 = \alpha + \beta_t P_t + \varepsilon_t \tag{4.9}$$

Where $Adj\ R_t^2$ represents the series of adjusted R^2 measures and P_t the variable that reflects the time passing.

To confirm H3, we needed that β_t to be negative and also statically significant.

H4

This hypothesis is inspired by the work of multiple authors. The first thing to do is assemble a set of portfolios that will properly reflect the effects that we are trying to capture. In this case, the objective is to create portfolios based on quality and value and to achieve this, we based our methodology on Fama and French (1993). These authors studied the performance of portfolios constructed on value and size, so our objective was to build our portfolios using the same process and swap the size for quality.

The first step consists of ranking every available stock by their PBV and, separately, by their quality score. After that, we categorized both rankings into six divisions: high-value, medium-value, and low-value for the PBV ranking and high-quality, medium-quality, and low-quality for the quality ranking. We used the same percentile values as Fama and French (1993) for their factors: the top 30% for the high categories and the bottom 30% for the low categories. The 40% remaining represent the medium-value and medium-quality categories. This procedure is done for every single period separately.

The second step aims at building the portfolios. In this step, we simulate the performance of 9 different portfolios specified by the following table:

	High-Value	*Medium-Value*	*Low-Value*
High-Quality	1	2	3
Medium-Quality	4	5	6
Low-Quality	7	8	9

For each period, we sort which stocks are simultaneously classified as high-value and high-quality to put together portfolio 1 and the same is valid for the remaining eight portfolios, respectively to which value or quality ranking they are concerned with. It is important to note that this last process is just made to find which stocks should be included in which portfolio and at which period. To allocate weights to each stock, we also followed Fama and French (1993) value-weighting each portfolio by each company's market size at the respective period.

Every portfolio is recalculated and rebalanced every July of each year to keep the methodology faithful to Fama and French (1993), and with this, we obtain the returns for each portfolio across the entirety of our time sample. With the returns calculated, we will compare them with the returns of a given benchmark and obtain the abnormal return of the portfolio. In this case, the benchmark is the S&P 500 because it is a good representative of the US equity market and a value-weighted stock portfolio, characteristics also shared with our portfolios.

H5

Asness et al. (2019) developed a quality-minus-junk (QMJ) risk factor in the same way that Fama and French (1993) developed several very famous risk factors, including SMB and

HML. This new factor's main purpose is to help determine a portfolio's exposure to some quality effect. The methodology to create a QMJ factor is fairly simple: firstly, determine six value-weighted (market size-weighted, to avoid any misconception) portfolios built on size and quality, namely Small Quality, Small Neutral, Small Junk, Big Quality, Big Neutral and Big Junk; next, the authors determine the returns of a strategy that goes long the quality portfolios and short-sells the junk portfolios. They present the factor as below:

$$QMJ = \frac{1}{2}(Small\ Quality + Big\ Quality) - \frac{1}{2}(Small\ Junk + Big\ Junk) = \frac{1}{2}(Small\ Quality - Small\ Junk) + \frac{1}{2}(Big\ Quality - Big\ Junk) \quad 4.10$$

To proceed with our analysis regarding the explanatory power of the Three-Factor Model with the addition of a QMJ factor on the returns of value-quality portfolios, we used the same data series as Asness et al. (2019) for the QMJ factor. For SMB, HML, and RM-Rf we extracted the data from the work done by French (2021), which keeps track of these factors exactly how they were calculated originally.

Additionally, intending to deepen the analysis, we have developed a QMJ factor built slightly differently, to which we gave the name QMJ mark II for distinguishing purposes. This version of the factor is built considering not six size and quality portfolios but instead nine value and quality portfolios. It mimics the structure of our value-quality portfolios that are the center of this analysis. Summarizing QMJ mark II into a formula, we have the following:

$$QMJ_{II} = \frac{1}{3}(Value\ Quality + Neutral\ Quality + Growth\ Quality) - \frac{1}{3}(Value\ Junk + Neutral\ Junk + Growth\ Junk) = = \frac{1}{3}(Value\ Quality - Value\ Junk) + \frac{1}{3}(Neutral\ Quality - Neutral\ Junk) + \frac{1}{3}(Growth\ Quality - Growth\ Junk) \quad 4.11$$

5. Data and Sample

To obtain every piece of data for every raw variable, we used the Datastream database. Variables like ratios and scores were mainly calculated afterwards.

Our sample is exclusively focused on the American Stock Market. Meaning we only included companies with their shares traded in developed and regulated US exchanges (NASDAQ and New York Stock Exchange). We chose the US market because it has the most considerable amount of data available in terms of the number of companies and time length.

The final sample is composed of 5936 firms from the NASDAQ and NYSE markets, and this number includes active companies and dead companies to avoid any survivorship bias. In addition, we selected companies from every sector except the financial sector since firms from this industry have atypical capital structures and other characteristics that could bias our analysis.

Our time horizon goes from the beginning of 1980 until the end of 2020, which accounts for 41 years of data, divided into quarters. Naturally, because of the nature of our analysis and methodology, we needed this series of data for multiple variables, and the availability of data for these variables is different for each one which results in a different number of actual firm-quarter observations for every variable. Whenever we found a blank entry in the middle two existent observations, we calculated the mean of the two to determine the approximate amount that would be in that particular spot. We made this because the only reason for a blank entry in the middle of filled observations is a missed value from the database, and this way, we avoid any break in the middle of the sample and more missed observations. The only variable for which we have not done this procedure is Total Debt, as it is difficult to distinguish "0" entries and blank entries given the way the data is presented when extracted from the database. After doing this procedure, every variable is winsorized on the 1st and 99th percentiles to tame outliers while avoiding eliminating further observations.

In terms of available observations, and just as a reference, the final quality score has a total of 320498 firm-quarter observations, which for a total of 949760[1] possible entries shows how unbalanced is our panel. Although it may seem like a small relative amount, one has to consider that: 1) the quality score for company i at time t is only calculated when there is an

[1] $160 \times 5936 -$ four quartes less because by including year-over-year growth rates in the analysis, *n* years turn into *n-1* growth rates.

entry for every single variable included in the calculation for i at time t and taking into consideration the number of variables and steps to take, it is understandable that at the end, many quality score values will be missing; and 2) at the beginning of our time horizon, there are more firms with no data than the ones that have that available and only after a few years into the time horizon this is inverted.

Table 1 – Descriptive Statistics

Number of stocks	5936
Active	3078
Dead	2858
NASDAQ	3355
NYSE	2581

This table presents a small set of descriptive statistics from our entire sample. The total number of stocks included is here divided by active and dead companies and also by stocks exchanged on the NASDAQ and NYSE markets.

Table 2 – Descriptive Statistics (Cont.)

	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
# of stocks	197	516	552	583	616	640	680	737	844	909	969	1098	1230	1312	1574	1851	2056	2342	2591	2886
Av PBV	1.65	1.51	2.22	1.94	2.03	2.27	2.44	2.1	2.33	2.16	2.49	2.6	2.88	2.77	2.9	3.23	3.44	3.41	3.24	3.74
Av Size	639	620	839	762	843	1010	1139	1007	1126	1161	1286	1379	1392	1336	1533	1757	2137	2499	3087	3560

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
# of stocks	2911	2937	2881	2873	2863	2763	2710	2729	2725	2683	2674	2693	2752	2931	3077	3083	3076	3107	3083	2610
Av PBV	2.6	2.74	2.66	3.2	3.11	3.31	3.55	2.75	2.11	2.56	2.77	2.67	3.03	3.91	3.37	3.17	4.24	4.05	3.86	4.01
Av Size	2853	2465	2538	2996	3310	3576	4082	3479	2969	3656	4050	4556	5347	6130	6243	6257	7213	7824	8542	9389

This table presents another set of descriptive statistics for the same sample. We present this additional information year by year because we work with panel-type data. This way, for every year of our time horizon we show the number of stocks, the average PBV, and the average size given by the stock's market capitalization (in million USD).

6. Results

We divide this section into five. The first three go over the relation between quality and stock prices, presenting the tests and results for hypotheses H1, H2, and H3. The fourth and fifth sub-sections are about the analysis of value-quality portfolios and the QMJ factor, which covers hypotheses H4 and H5.

6.1. Quality Score and Stock Prices

The first regression we have computed represents a model based on the one from Asness et al. (2019), which aims at testing the impact of a stock's quality on its price. Initially, we calculate Fama-MacBeth regressions, simply of PBV on the quality score of each company at every period (Model 1). Secondly, we add industry-fixed effects, as part of the analysis from Asness et al. (2019), by creating a dummy variable for every sector and potentially improve the explanatory power of the model (Model 2). Finally, we also include a number of controls to improve the model further, namely, firm size, past 12-month returns, firm age, profit uncertainty, and a dummy variable to control for dividend payers (Model 3). Except for the dummy variables, every explanatory variable is standardized into z-scores, through the same process as the quality score.

These three models can be expressed by the following estimation representations:

$$PBV_{i,t} = c_{i,t} + \beta_{i,t}Q_{i,t} + \varepsilon_{i,t} \tag{6.1}$$

$$PBV_{i,t} = c_{i,t} + \beta_{1,i,t}Q_{i,t} + \beta_{2,i,t}D_{1,i,t} + \cdots + \beta_{11,i,t}D_{10,i,t} + \varepsilon_{i,t} \tag{6.2}$$

$$\begin{aligned} PBV_{i,t} = c_{i,t} + \beta_{1,i,t}Q_{i,t} + \beta_{2,i,t}Size_{i,t} + \beta_{3,i,t}Mom_{i,t} + \beta_{4,i,t}EV_{i,t} + \\ \beta_{5,i,t}Age_{i,t} + \beta_{6,i,t}DP_{i,t} + \beta_{7,i,t}D_{1,i,t} + \cdots + \beta_{16,i,t}D_{10,i,t} + \varepsilon_{i,t} \end{aligned} \tag{6.3}$$

Table 3 – Estimation Output (H1)

Explanatory Variables	Model 1 (7.1)	Model 2 (7.2)	Model 3 (7.3)
Intercept	2.786 (50.472)	2.621 (59.529)	2.673 (59.569)
$Q_{i,t}$	0.243 (9.567)	0.285 (11.647)	0.274 (11.602)
$Size_{i,t}$			0.737 (25.052)
$Mom_{i,t}$			0.539 (20.477)
$EV_{i,t}$			0.548 (10.173)
$Age_{i,t}$			-0.119 (-13.810)
$DP_{i,t}$			-0.267 (-10.416)
Average R^2	0.011	0.050	0.130
Average Adj R^2	0.007	0.041	0.118
Number of periods	159	159	159

This table presents the estimation output of the three models calculated using a Fama-MacBeth procedure. The dependent variable is the PBV for each company at the end of every quarter. The explanatory variables are, in order, the quality score, the firm size given by market capitalization, momentum expressed by the previous 12-month return, earnings volatility given by ROE's variance, the firm's age, and a dummy variable that controls for dividend-paying stocks. Model 2 and Model 3 also have dummy variables that control for industry effects. Every variable, apart from the dummies, is standardized through z-scores to have a cross-sectional mean of zero and a cross-sectional standard deviation of one. The three models are estimated using Newey and West (1987) heteroskedasticity- and autocorrelation-adjusted standard errors with 4 time-lags based on a formula[2] provided by Bali et al. (2016). Average R^2 and average adj R^2 are the simple time-series averages for the R^2 and adjusted R^2 measures for every cross-sectional Fama-Macbeth regression as in Lewellen (2015), Asness et al. (2019), and Berg (2020). In these three models, every variable is statistically significant at conventional levels, as shown by the t-statistic value indicated below every coefficient.

From Table 3 we can extract that, on average, quality, as defined in the beginning, is a characteristic for which investors are willing to pay a premium. For each additional quality score unit (which is a variation of one standard deviation), the PBV of a company is predicted by the models to increase from 0,24 on Model 1 to 0,28 on Model 2. This is consistent with

[2] $t_{lags} = 4(\frac{T}{100})^{\frac{2}{9}}$

the findings from Asness et al. (2019) and Berg (2020). Even the coefficients are particularly similar despite both authors using the natural logarithm of PBV as the dependent variable.

We also find that the explanatory power of quality, although, is overwhelmingly low. This is also one of the conclusions stated by both works cited just above. In all three models, the coefficient associated with the quality score remains relatively unchanged and is statistically significant at conventional levels every time, while the adjusted R^2 goes from less than 1% on the first model to close to 12% on the third version, when several controls are added, including industry-fixed effects which are already present on the second model. This can only mean that although quality is correlated with the PBV, it does not explain well the variation of PBV around its mean, leaving the overwhelming majority of this variation unexplained. Asness et al. (2019), in particular, have come to the same conclusion even though the models presented in the paper reveal substantially higher R^2 measures. The additional coefficients presented in Model 3 are not the focus of this analysis. What is important is that every one of them is statistically significant at conventional levels and contributes to improving the model's explanatory power.

We have found evidence that confirms the initial idea and previous findings, so we believe that quality is a characteristic for which investors are willing to pay for, although the explanatory power of this variable on price changes is very limited, therefore, we accept H1. Again, these two conclusions are very similar to those found previously, especially on Asness et al. (2019) and Berg (2020).

6.2. The Price of Quality Over Time

Something else we have proposed ourselves to assess is the evolution of quality's price over time. The price of quality is here defined as the regression coefficient associated with our quality measure present in our models. The methodology we used initially, a Fama-MacBeth procedure, provides an easy solution to observe the evolution of the price of quality: we simply extract the regression coefficient for quality for every cross-sectional Fama-MacBeth regression and compare the different coefficients we get for each period.

Chart 1 – The price of quality over time (Model 2)

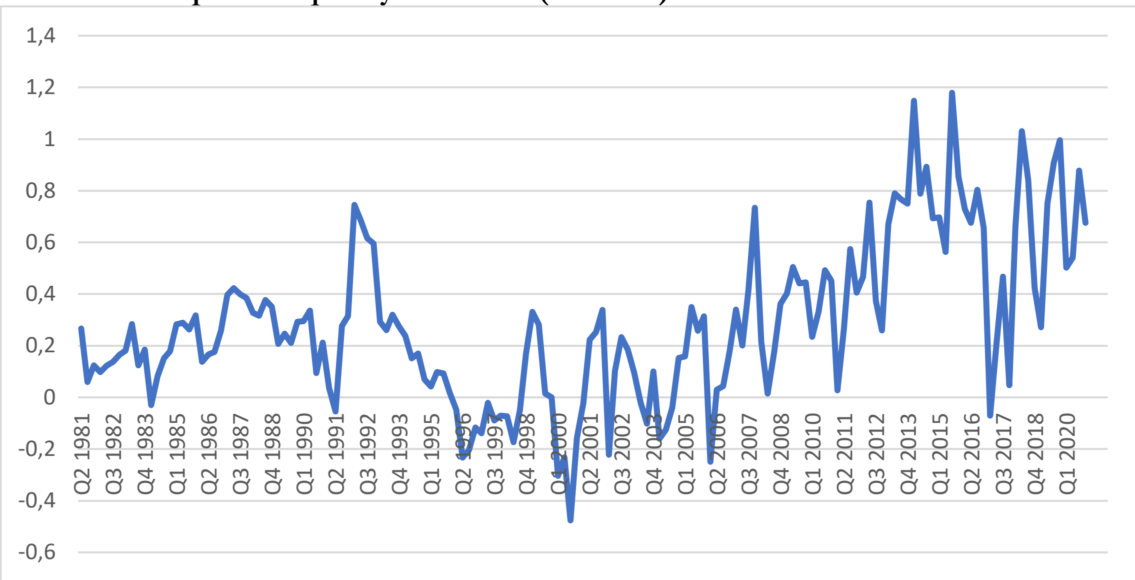


Chart 1 presents what we just stated: the evolution of the price given to quality proxied by the regression coefficients associated with the quality score in our Model 2. The idea here is to see if there is any meaningful price behavior during market downturns and extract if that behavior reveals that quality is viewed as a safe investment and if there is a "flight to quality" during these more volatile market moves, as Asness et al. (2019) found on their analysis.

Within our time sample, we can include several market crashes that were violent enough to provoke any kind of reaction to the price of quality, if there is any reaction to be generated. The following examples can be included:

- The stock market crash from late 1987 – Dow Jones dropped 40% in just three months;
- The dot com bubble burst – the same index lost almost 40% from January 2000 until October 2002;
- The great financial crisis from 2007/2008 – the Dow Jones fell over 50% from October 2007 to March 2009;
- The deflationist shock caused by the COVID-19 pandemic in early 2020 – the market went down almost 40% in 6 weeks.

With these four key moments in mind, we can now look at the changes in the price of quality and how these violent shocks affect it through time.

For the first period, we see a decrease from 0.40 to 0.31 on Model 2, which indicated that the price of quality decreased during this period, being in both periods over the average of 0.28. Subsequently, from the beginning of 2000 until mid-2002 we see the price of quality go from -0.10 to 0.25 and there are two things to note here: first, the price of quality appeared to be especially low during the dot com bubble, which was something also pointed by Asness et al. (2019), the only paper found to approach this matter; second, the price of quality actually increased in a time in which the overall markets were bearish which may suggest a "flight-to-quality" event. Moving on to the great financial crisis, we see the price of quality increasing from 0.27 to 0.58 (an all-time high at that time) during the initial shock. At the end of the period, at the beginning of 2009, the price was 0.10, so despite increasing initially, the price of quality dropped with the market afterwards. Finally, during the fast market crash at the beginning of 2020, we see the price of quality again dropping from 0.61 in late 2019 to 0.46 at the end of Q1 2020 and 0.26 in mid-2020. On this last one, we also can't see any sign of a "flight-to-quality" event given the price of quality followed the market.

Chart 2 – The price of quality over time (Model 3)

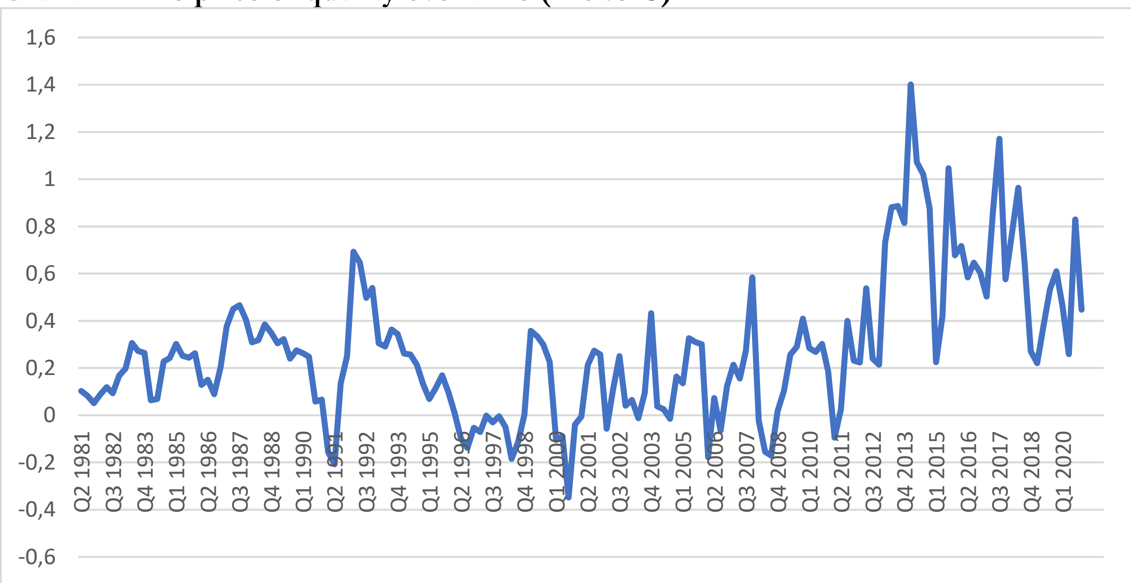


Chart 2 presents the same information as Chart 1 but this time considering the data from Model 3 instead of Model 2. As a reminder, the models are similar and only differentiate because Model 3 adds a series of control variables.

The price action of quality remains very similar, and there is no big difference to note. Apart from specific price values, the information presented before is also valid for Model 3.

Table 4 - Estimation Output (H2)

Explanatory Variables	Regression Coefficients
Intercept	0.297 (4.756)
D_t	-0.160 (-1.845)
Adj R^2	0.030
Number of periods	159

This table presents the estimation output of a simple regression with the price of quality as the dependent variable. Here, the price of quality is given by the coefficients associated with the quality score obtained on the estimation of Model 3 presented in Table 3. The only explanatory measure is a dummy variable assigned to 1 when the time period is inside of one of the crises described before, and 0 otherwise. The model is estimated using Newey and West (1987) heteroskedasticity- and autocorrelation-adjusted standard errors with 4 time-lags based on a formula provided by Bali et al. (2016). The adjusted R^2 measure and the number of periods included in the analysis are also presented. Considering only conventional levels, the variable D_t is only statistically significant at a 10% level of significance, as shown by the t-statistic value indicated below the coefficient.

Right above, in Table 4, we show the results of our econometric approach to this question. D_t, the dummy variable, is reported to have a -0.16 coefficient. Additionally, this coefficient is only significant at 10% significance, considering standard significance levels. Not surprisingly, the adjusted coefficient of determination is very low at 3%.

In conclusion, we should say that we encounter no additional demand for quality stocks during big market downturns. That effect was observed during the dot com bubble burst but in no other major bear market included. Even when it was observed, this increase in quality prices was not substantial enough to be treated as a "flight-to-quality" event because if we look at Charts 1 and 2, we see that the variation occurred on that period is not particularly big comparing with the rest of the timeframe, including moments when the markets were stable. To accept H2, not only this variation would have to be significant relative to other periods, i.e., more volatility on the price of quality, as the market crash would implicate, but also that event would have to be observable in other similar market deflationary events. Furthermore, our empirical test shows that in periods of crisis, what tends to happen is precisely the opposite: quality prices tend to follow the marker, declining, and not going against it. This way, and very much against the conclusions from Asness et al. (2019), which claim a significant increase in the price of quality on market downturns, we have to reject hypothesis H2.

6.3. The Explanatory Power of Quality on Prices

The next analysis included in this section concerns H3, which addresses the evolution of the explanatory power of quality on prices. The way to assess this evolution is very similar to how we developed the analysis on the evolution of quality's price over time. Similarly, we looked into every single Fama-MacBeth quarterly regression, but this time we extracted the R^2 for every period to observe the behavior of this measure throughout time.

Chart 3 – Regressions R2 measure over time (Model 2)

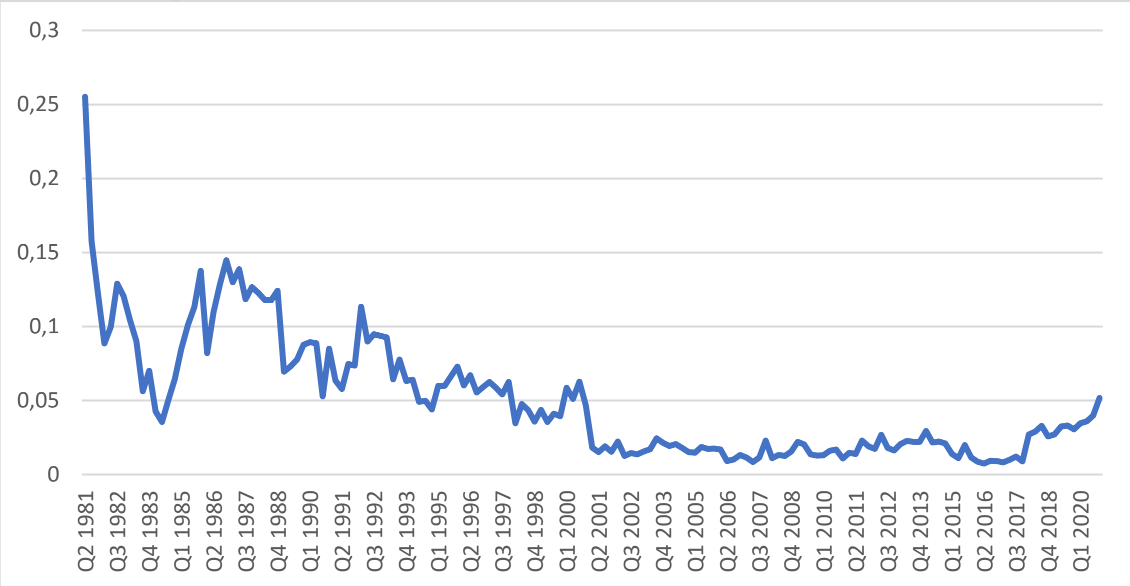


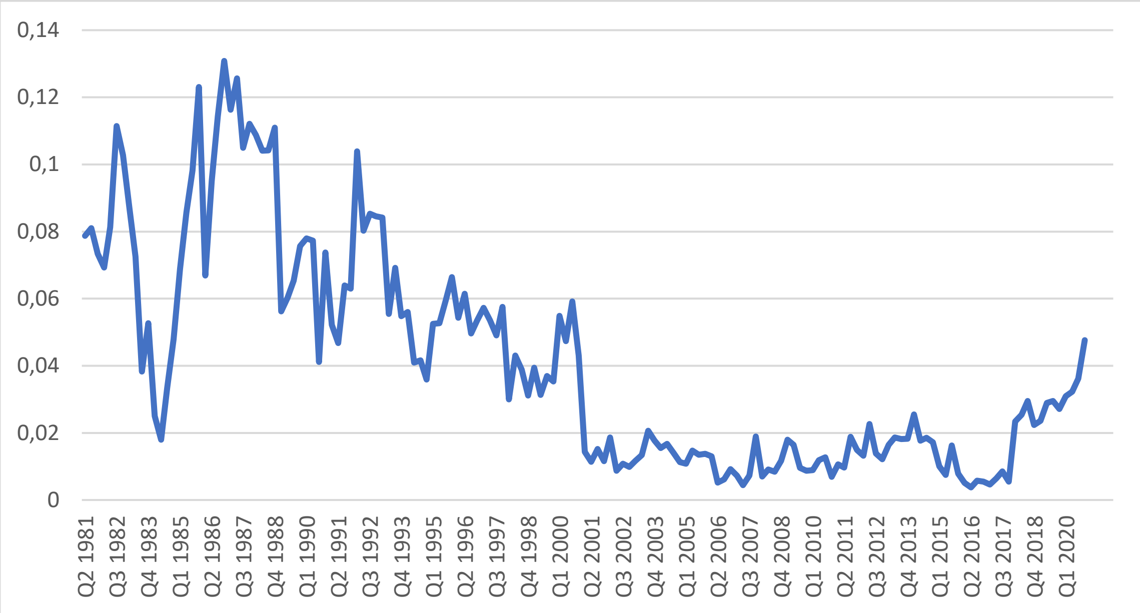
Chart 3 shows the coefficient of determination for each of the 159 cross-sectional regressions built with our Model 2. This specific model only has the quality score as an explanatory variable for prices and dummy variables that control different industries.

As it is observable, the R^2 has an undeniable downwards tendency since the beginning of our sample and pretty much until the very end, if it wasn't for the small increase since 2018.

To this day, we found no literature pointing to this phenomenon, perhaps because no other authors have encountered it on their specific samples or because they didn't notice it. Nevertheless, this relation appears to be too clear to ignore, and we currently have no answer to why this might have happened. It could be that quality has been losing the connection with prices over time, or maybe different risk factors have been arising and taking some of its place.

On Chart 4, below, we show the Adjusted R^2, which is adjusted for the number of explanatory variables in the model. The picture is slightly different, but the overall results are roughly the same: it has been decreasing with a fair level of consistency from the beginning and until 2018, when it started rising and seemed to go against the trend.

Chart 4 – Regressions Adjusted R2 measure over time (Model 2)



Below, on Charts 5 and 6, we show the same measures, respectively, but for Model 3, which includes a bigger number of model predictors. Again, the observations previously pointed remain valid for this model.

Chart 5 – Regressions R2 measure over time (Model 3)

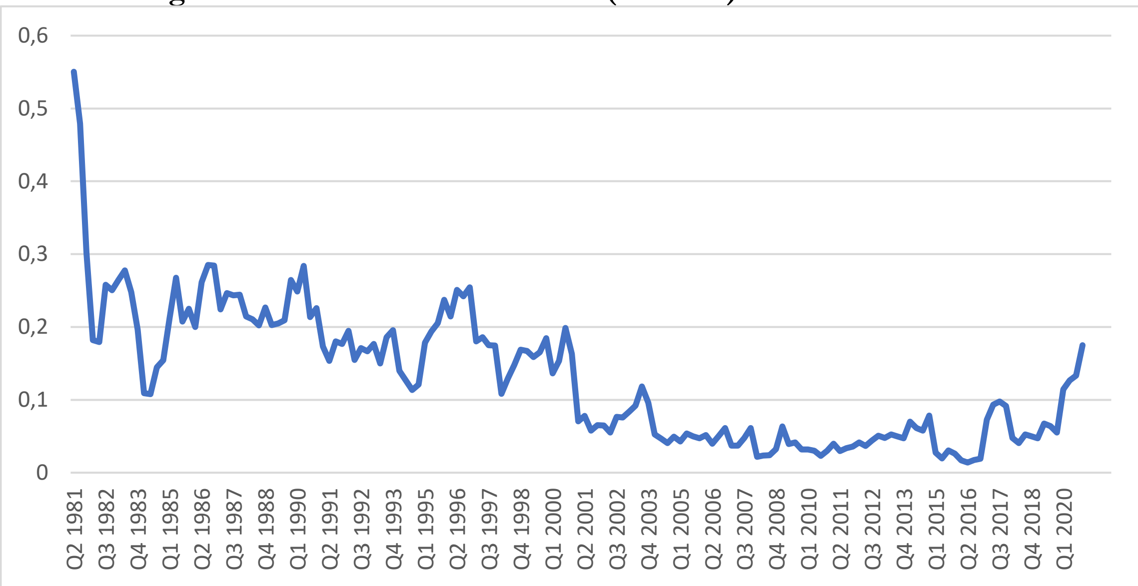


Chart 6 – Regressions Adjusted R2 measure over time (Model 3)

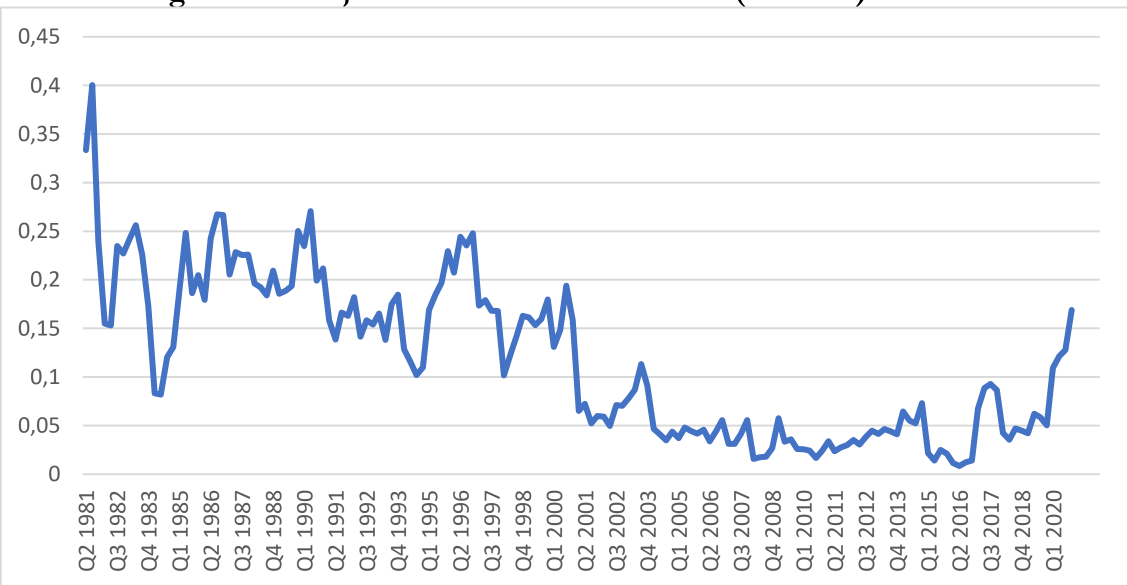


Table 5 - Estimation Output (H3)

Explanatory Variables	Regression Coefficients
Intercept	0.230 (14.594)
P_t	-0.001 (-7.105)
Adj R^2	0.623
Number of periods	159

This table presents the estimation output of a simple regression with the dependent variable being the adjusted R^2 measures obtained on the estimation of Model 3 presented in Table 3. The only explanatory measure is a variable purposefully created to reflect the passage of time by having assigned a value of t to each period, being t the number of the respective period. The model is estimated using Newey and West (1987) heteroskedasticity- and autocorrelation-adjusted standard errors with 4 time-lags based on a formula provided by Bali et al. (2016). The adjusted R^2 measure and the number of periods included in the analysis are also presented. The variable P_t is statistically significant at conventional levels, as shown by the t-statistic value indicated below the coefficient.

In Table 5 we present another econometric test, this time to further prove that, in fact, the explanatory power of quality has decreased over time. The coefficient associated with the variable that captures the passage of time appears to be negative and statistically significant even at the lowest standard significance level. On top of that, the adjusted R^2 of this regression is particularly high, revealing that 62.3% of the whole variation of this specific explained variable around its sample mean is explained solely by the time passing by.

Given all this information, it is fair to conclude that there is a relevant and negative evolution of the explanatory power of quality on prices, therefore, we accept H3. In this case, we only supply the observation, and we leave the reason for this to have happened open to future work.

6.4. Portfolios Formed on Quality

This sub-section focuses on analysing portfolios formed and ranked by quality and value, taking inspiration from Fama and French (1993). Not only that, but we also present an alternative way of building the quality-minus-junk factor firstly developed by Asness et al. (2019), which in turn was inspired by the SMB and HML factors created by Fama and French (1993), and we analyze the empirical relation between these portfolios and the QMJ factor.

Table 6 – Statistics for each portfolio formed on Quality and Value

	Value					
	Average number of firms annually			Average PBV		
Quality	Low	Medium	High	Low	Medium	High
Low	185	211	245	8.387	2.07	-0.838
Medium	202	395	257	5.766	2.067	0.489
High	253	248	139	6.503	2.136	0.456
	Average Quality Score			Average Firm Size		
Low	-1.246	-0.934	-1.124	2453	2318	1130
Medium	0.134	0.12	0.082	7069	3879	1943
High	0.992	0.883	0.952	9794	3528	1032

The nine portfolios are built on their quality and value characteristics in a similar way that Fama and French (1993) built portfolios based on value and size. Our raw sample of statistics goes from 1985 to 2020, and overall, we include 5936 different companies in the analysis. The methodology to build each portfolio can be described as follows. First, we dispose of the data (quality score and PBV) for each company and each period obtaining two panels, one for each variable. After that, for every period, we rank the stocks according to their quality score for one panel and according to their PBV for the other panel (it is important to note that the lower the PBV, the higher the value measure). For each period, we divide the stocks as "high" for the 30% highest quality stocks, "medium" for the following 40%, and "low" for the bottom 30%. The same is done for the PBV panel. With this information, we assess which stocks are at the same time high in quality and high in value so that they are included in the high-quality-high-value portfolio, and we do the same for every other portfolio respectively to their interval and for every year inside our sample. The portfolios are value-weighted, meaning that the weight given to each stock is based on their market capitalization. Finally, every portfolio is rebalanced in July, and the returns are calculated from July to June of each year. Every step of this methodology is heavily inspired by that of Fama and French (1993).

Table 6, above, shows the descriptive statistics for the nine portfolios we have built based on quality and value. The columns are about the value side, whereas the lines are allocated to the quality side of our analysis. We present four different statistics for each portfolio: the average number of stocks per portfolio for every year and the average PBV, quality score, and firm size of every stock on each portfolio and for every period.

The number of stocks is relatively balanced across the board. Naturally, portfolios on the extremes have fewer stocks than the ones in the middle were it not for the fact that the interval for the middle stocks is 40% of the available stocks and not 30% as in the extremes. Also, given that quality stocks are, mostly, growth stocks (low value), a low-quality portfolio has more stocks when combined with value stocks, less when combined with growth stocks, and vice-versa.

The averages for PBV and quality score behave as one would expect, so there is no particular observation to mention. Firm size is a variable that is not used for stock selection, and we see clearly that the closer to a growth portfolio, the higher the averages for the size of the firms included. This tells us that growth stocks tend to be bigger in market capitalization than value stocks.

Table 7 – Statistics for each portfolio formed on Quality and Value (Cont.)

		Value						
		Average Excess Returns				Average Abnormal Returns		
Quality		Low	Medium	High		Low	Medium	High
Low		6.31%	3.52%	6.73%		0.27%	-2.52%	0.69%
Medium		7.61%	5.30%	7.51%		1.57%	-0.75%	1.47%
High		10.63%	7.78%	12.58%		4.60%	1.73%	6.54%
		Sharpe Ratio				T-stat for Av. Abnormal Returns		
Low		0.275	0.172	0.328		0.113	-1.231	0.260
Medium		0.462	0.372	0.466		1.073	-0.406	0.570
High		0.656	0.552	0.617		3.214	0.865	1.900

The nine portfolios are built on their quality and value characteristics in a similar way that Fama and French (1993) built portfolios based on value and size. Our raw sample of statistics goes from 1985 to 2020, and overall, we include 5936 different companies in the analysis. The methodology to build each portfolio can be described as follows. First, we dispose of the data (quality score and PBV) for each company and each period obtaining two panels, one for each variable. After that, for every period, we rank the stocks according to their quality score for one panel and according to their PBV for the other panel (it is important to note that the lower the PBV, the higher the value measure). For each period, we divide the stocks as "high" for the 30% highest quality stocks, "medium" for the following 40%, and "low" for the bottom 30%. The same is done for the PBV panel. With this information, we assess which stocks are at the same time high in quality and high in value so that they are included in the high-quality-high-value portfolio, and we do the same for every other portfolio respectively to their interval and for every year inside our sample. The portfolios are value-weighted, meaning that the weight given to each stock is based on their market capitalization. Finally, every portfolio is rebalanced in July, and the returns are calculated from July to June of each year. Every step of this methodology is heavily inspired by that of Fama and French (1993).

Table 7 presents another set of statistics, this time related to the performance of each portfolio. Starting from the excess returns section, we present the time-series average returns for each portfolio, subtracting the risk-free rate in the respective period. It is very clear that quality is a characteristic that consistently increases the excess return of a portfolio. By controlling for the value parameter, every portfolio increases its return linearly when scaling up the quality parameter, with no exception. The two best returning portfolios are the ones that combine high-quality stock with value stocks and high-quality stocks with growth stocks. The first place is taken by combining value and quality strategies at their peak, given that portfolio HVHQ returns on average 12.58% per year above the risk-free rate.

Abnormal returns are the returns obtained by each portfolio above a certain benchmark. In this case, we have chosen the S&P 500 index as our benchmark because it represents a diversified US equity portfolio that is value-weighted, as our portfolios are. Additionally, being one of the most looked-after indices worldwide, it is easy and reliable to find data for it. Now looking at the performance reached by our portfolios, the conclusion is similar to the previous one: high-quality portfolios present an improved performance over junk portfolios, and the latter ones either present an average abnormal return close to 0 or even considerably negative for the medium-value-low-quality (MVLQ) portfolio. Again, this increase in average return rates is consistent for every step towards an increased quality level across the nine portfolios, and the two top-performers remain the same as before, as one would expect, given the only difference between excess returns and abnormal returns is the benchmark (risk-free rate on the first and S&P 500 on the latter). The first place remains to portfolio HVHQ which, on average, outperforms the S&P 500 by 6.54% in one year.

We have also included T-statistics for our time-series of average abnormal returns for each portfolio, and with these results, we can see that only two out of the nine portfolios have average abnormal returns statistically different from 0, for a 10% significance level. The two portfolios are the same as the ones mentioned before – HVHQ and LVHQ – so both of them include a high-quality characteristic in their composition. For lower conventional significance levels, only portfolio LVHQ remains with abnormal returns statistically significant, given that the volatility on the returns of HVHQ is enough to confirm the null hypothesis of a T-test.

The last section of Table 7 concerns risk-adjusted returns presented by each portfolio's Sharpe Ratio (Sharpe, 1966). As returns, by themselves, don't provide enough information to decide whether a portfolio is better than the other or not, we also calculated risk-adjusted returns, which also include the risk associated with the portfolio in the equation. Ranking the nine portfolios by returns and ranking them by Sharpe Ratio, we get slightly different results, proving thie value of this analysis. We still observe the same two portfolios on the top, as with the previous two measures, but this time they switch places on the first and second place. To put in perspective, using the same methodology[3] for the S&P 500, we obtained a Sharpe Ratio of 0.419, meaning that every single high-quality portfolio keeps outperforming the benchmark very substantially, even when considering risk-adjusted returns.

Chart 7 – The performance of every portfolio over time

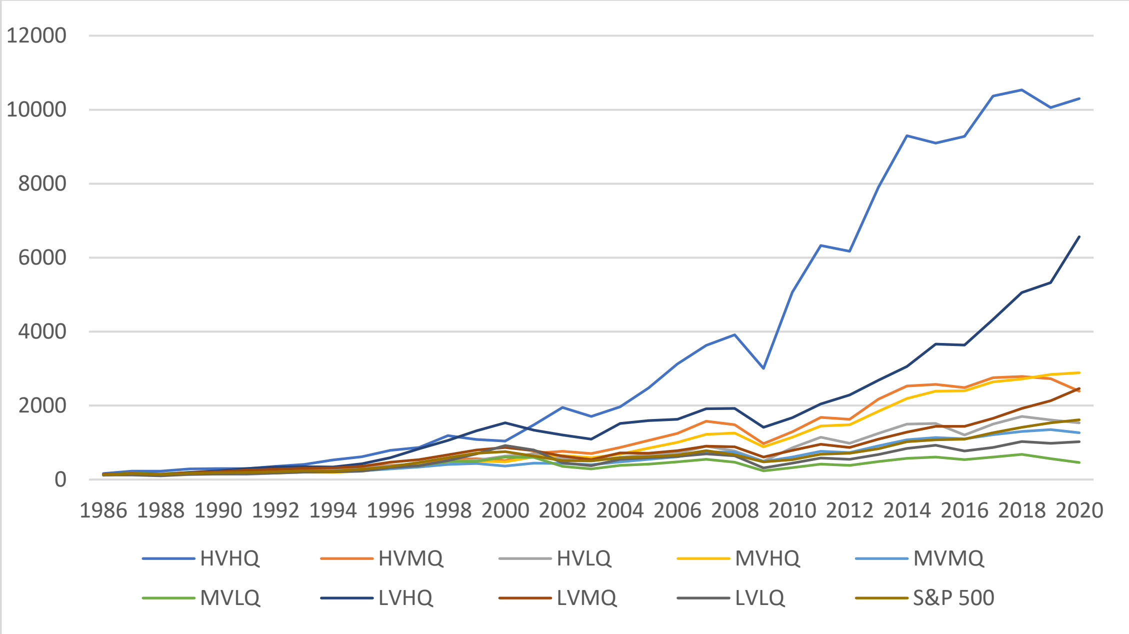


Chart 7 displays the absolute performance of 10 different portfolios: the nine we have built and studied, and our benchmark, the S&P 500. We have simulated a strategy in which $100 are invested in one portfolio at the beginning of our time horizon, July 1985, and held until the end, June 2020, to observe the differences in performance in cumulative terms across the whole period. We assume no transaction costs. From Chart 7, it is very apparent that portfolio HVHQ has taken the lead, especially since the beginning of the century when it took off from a fight with portfolio LVHQ. This latter one was distinctively the second-best at the end of the period, and the third place went for the remaining portfolio built on high quality, MVHQ.

To close this sub-section about the analysis of the performance of portfolios built on quality and value, we will conclude the assessment of H4. Kozlov and Petajisto (2013) and Novy-Marx (2013) have suggested that a high-value-high-quality portfolio would perform relatively well because of the combination of the value factor with the quality factor. According to the

[3] Having every year's return, we calculate the standard deviation of the returns. We then obtain the average excess return across the whole period and we divide this amount for the standard deviation of the returns.

authors, the risk-adjusted performance would presumably be particularly good because usually, quality stocks are growth stocks, so the strategies have a negative correlation. We were able to see that portfolios built on value and quality perform substantially better than portfolios with different characteristics and better than our benchmark, the S&P 500 index. Despite the best performing portfolio being the one with the highest degree of quality and value, we observe that the one built on high quality and low value (growth) is actually fairly close to the first one. In fact, in terms of risk-adjusted performance, this latter one outperforms the first one revealing that the overperformance may not come from the combination of the two strategies but the quality factor by itself. We also have to consider that if we look at the final results, we see no linearity regarding the value factor, meaning that the performance is not linearly enhanced as we go from a low-value portfolio to a high-value portfolio – which happens when we do the same for quality. Given this, we accept H4 because we observed relevant abnormal risk-adjusted returns on the HVHQ portfolio. Still, we want to leave the note that a LVHQ portfolio performed just as well, especially taking into account the risk inherent to each portfolio. When we look at the absolute and cumulative performance, HVHQ is a distinct winner.

6.5. Fama-French Three-Factor Model and the QMJ factor

To put an end to the analysis of value-quality portfolios and the overall *Results* section, we will now comment on the QMJ factor, our approach and development on the matter, and the results of several econometric regressions based on Fama and French (1993).

Table 8 – Statistics for each factor

	Average	St. Dev.	T-stat		Correlations				
					$R_m - R_f$	SMB	HML	QMJ	QMJ_{II}
$R_m - R_f$	8.206	13.945	3.481		1				
SMB	0.001	8.326	0.009		-0.094	1			
HML	1.357	13.977	0.574		-0.086	0.278	1		
QMJ	5.656	8.909	3.756		-0.592	-0.177	0.079	1	
QMJ_{II}	4.809	11.933	2.384		-0.554	-0.056	0.152	0.720	1

In our analysis, we include four factors: the three factors from the Three-Factor Model from Fama and French (1993), namely $R_m - R_f$, the market risk premium; SMB, the size premium; and HML, the value premium, plus the QMJ factor developed by Asness et al. (2019), which captures the premium paid for high-quality stocks. The first three factors are obtained from the work done by French (2021), which keeps track of these factors exactly how they were calculated originally. The QMJ factor was obtained from the same data series as Asness et al. (2019). We have also developed a second QMJ factor that would hopefully be more suited for our specific analysis. This QMJ_{II} factor differs from the first one in its theoretical construction. Originally, this factor was built on quality and size, but in our adaptation, we have built it on quality and value to mimic our portfolios in analysis. In this table, we present data for every factor taking into consideration our time horizon from June 1985 to June 2020. Average returns and the respective standard deviations are presented in percent (%) terms. Alongside simple averages, standard deviations, and T-statistics, we also present the correlation between each pair of factors.

Table 8 shows a set of descriptive statistics of every factor we use in the regressions presented on the following pages. It is to note that in our 36 years of data, SMB and HML average returns are not statistically significant, whereas the market risk premium, QMJ and QMJ_{II} are statistically different from 0 for any conventional significance level. Also worth noting is the fact that within our sample, we have not found a negative correlation between quality and value, at least in the form of risk factors, as we present them. The correlation between value and any of our quality factors is positive, although not very significant, being 0.079 for the factor developed by Aness et al. (2019) and 0.152 for our adapted version of it. This differs from the conclusions from Kozlov and Petajisto (2013), which found a negative correlation between value and quality, inspiring a good portion of this dissertation. Surprisingly enough, the negative correlation with quality comes from the size factor. Again, this is observable in both quality factors: -0.177 for the first and -0.056 for the second.

Table 9 – Estimation Output (H5)

		Value					
		$\beta_{R_m-R_f}$			$T-statistic_{R_m-R_f}$		
Quality		Low	Medium	High	Low	Medium	High
Low		**1.440**	**1.312**	**1.15**	14.952	18.435	8.913
Medium		**1.061**	**0.836**	**0.762**	14.588	9.071	6.840
High		**0.936**	**0.758**	**0.780**	9.746	8.190	4.562
		β_{SMB}			$T-statistic_{SMB}$		
Low		**0.268**	0.205	**0.687**	2.562	1.052	1.921
Medium		0.088	0.050	0.105	0.876	0.396	0.509
High		-0.204	0.087	**0.638**	-1.031	0.663	2.024
		β_{HML}			$T-statistic_{HML}$		
Low		-0.064	**0.266**	**0.297**	-0.958	2.098	2.561
Medium		**-0.108**	**0.410**	**0.719**	-1.824	3.015	6.979
High		**-0.178**	**0.365**	**0.685**	-2.573	2.215	5.559
		$Adjusted\ R^2$			$F-statistic$		
Low		0.780	0.804	0.665	41.202	47.452	23.471
Medium		0.869	0.783	0.742	76.087	41.881	33.541
High		0.774	0.652	0.560	39.921	22.256	15.406

This table presents the output of 9 time-series regressions with the yearly excess returns of each portfolio as the dependent variable. The explanatory variables used are the three factors from the Three-Factor Model from Fama and French (1993), namely $R_m - R_f$, the market risk premium; SMB, the size premium; and HML, the value premium. Everything is on an annual basis. The three factors are obtained from the work done by French (2021), which keeps track of these factors exactly how they were calculated originally. The data included goes from June 1985 to June 2020, which in terms of returns results in 35 time periods which is the number of observations for each of the equation regressions. From top to bottom, we present the regression coefficients for each factor and each regression, and the respective T-statistic on the right side. The last set of lines is dedicated to the $Adjusted\ R^2$ and F-statistics for each regression. All models are estimated using Newey and West (1987) heteroskedasticity- and autocorrelation-adjusted standard errors with 3 time-lags based on a formula provided by Bali et al. (2016). Coefficients that are statistically significant at a 10% level of significance are indicated in bold.

Table 9 shows a set of data demonstrating the results of nine time-series regressions inspired by Fama and French (1993) and the Three-Factor Model. The regressions are computed as follows:

$$R_{P_{i,t}} - R_{f_t} = \alpha_t + \beta_{R_m-R_f}(R_m - R_f)_t + \beta_{SMB}SMB_t + \beta_{HML}HML_t + \varepsilon_t \qquad 6.4$$

Being $R_{P_{i,t}}$ the returns of each of the nine portfolios at time t and the rest of the variables as previously described.

Now analyzing the results, we can see that every portfolio has a statistically relevant exposure to the market risk premium, which is understandable. There seems to be no clear tendency or relation between the characteristics of the portfolio and the level of exposure to this factor. The coefficients range from 0.758 and 1.44.

The second coefficient is associated with the SMB factor, which captures the exposure to the size effect. Naturally, as none of the portfolios is built with any size consideration, most portfolios show no significant exposure to the size factor.

Regarding HML, the value factor, we can see a relation between the coefficient and the characteristics of each portfolio. There is always a considerable increase in the coefficient for any given quality level as we go from a growth portfolio to a neutral and finally to a value-based portfolio. Eight out of the nine coefficients extracted from these models are statistically significant. Here the range increases, going from -0.178 to 0.719.

The adjusted R^2 obtained in the models was considerably high all around the board. Six out of the nine are above 0.70, which reflects a fairly high explanatory power associated with this combination of independent variables.

The data here presented is not to be analyzed individually but to be considered and compared with the data that will follow, which includes the factors that capture the effects that high-quality characteristics have on the prices of stocks.

Table 10 – Estimation Output (Models with QMJ)

	Value					
	$\beta_{R_m-R_f}$			$T-statistic_{R_m-R_f}$		
Quality	Low	Medium	High	Low	Medium	High
Low	**1.241**	**1.245**	**0.916**	12.547	21.027	7.943
Medium	**1.063**	**0.901**	**0.694**	8.701	6.441	4.643
High	**1.100**	**0.906**	**0.916**	10.253	7.470	3.739
	β_{SMB}			$T-statistic_{SMB}$		
Low	0.125	0.157	**0.519**	0.852	0.793	1.716
Medium	0.089	0.097	0.056	0.842	0.686	0.291
High	-0.086	0.193	**0.735**	-0.532	1.119	2.519
	β_{HML}			$T-statistic_{HML}$		
Low	-0.032	**0.277**	**0.335**	-0.357	2.286	3.286
Medium	-0.109	**0.399**	**0.731**	-1.596	3.184	6.578
High	**-0.205**	**0.341**	**0.663**	-3.700	2.384	5.293
	β_{QMJ}			$T-statistic_{QMJ}$		
Low	**-0.513**	-0.171	**-0.602**	-3.363	-0.777	-2.229
Medium	0.005	0.168	-0.174	0.034	1.108	-1.095
High	**0.422**	**0.378**	0.349	2.565	2.211	0.911
	$Adjusted\ R^2$			$F-statistic$		
Low	0.800	0.801	0.699	35.052	35.234	20.741
Medium	0.864	0.783	0.739	55.227	31.714	25.095
High	0.806	0.680	0.561	36.300	19.091	11.870

This table presents the output of 9 time-series regressions with the yearly excess returns of each portfolio as the dependent variable. The explanatory variables used are the three factors from the Three-Factor Model from Fama and French (1993), namely $R_m - R_f$, the market risk premium; SMB, the size premium; and HML, the value premium, plus the QMJ factor developed by Asness et al. (2019), which captures the premium paid for high-quality stocks. Everything is on an annual basis. The first three factors are obtained from the work done by French (2021), which keeps track of these factors exactly how they were calculated originally. The QMJ factor was obtained from the same data series as Asness et al. (2019). The data included goes from June 1985 to June 2020, which in terms of returns results in 35 time periods which is the number of observations for each of the equation regressions. From top to bottom, we present the regression coefficients for each factor and each regression, and the respective T-statistic on the right side. The last set of lines is dedicated to the $Adjusted\ R^2$ and F-statistics for each regression. All models are estimated using Newey and West (1987) heteroskedasticity- and autocorrelation-adjusted standard errors with 3 time-lags based on a formula provided by Bali et al. (2016). Coefficients that are statistically significant at a 10% level of significance are indicated in bold.

Table 10 shows a large amount of data describing the results of our nine time-series regressions that now include the QMJ factor in the analysis. Each regression goes as follows:

$$R_{P_{i,t}} - R_{f_t} = \alpha_t + \beta_{R_m - R_f}(R_m - R_f)_t + \beta_{SMB}SMB_t + \beta_{HML}HML_t + \beta_{QMJ}QMJ_t + \varepsilon_t \qquad 6.5$$

Being every variable as previously described.

Looking at the first three factors, we see no massive difference not only in terms of regression coefficient values but also in terms of statistical significance.

For the different addition to these specific models, the QMJ factor, we see that only four portfolios present statistically significant coefficients associated with this factor. And again, not surprisingly, we also see a very clear relation between the QMJ coefficient and the quality characteristics of the portfolio.

When comparing the R^2 measures, the improvement is not significant and, in some cases, there is no improvement at all. This goes in line with the fact that less than half of the QMJ coefficients across the nine models are statistically significant.

To give another try to this additional factor, we will now include our version of the QMJ factor instead of the one from Asness et al. (2019).

Table 11 – Estimation Output (Models with QMJ II)

	Value					
	$\beta_{R_m-R_f}$			$T-statistic_{R_m-R_f}$		
Quality	Low	Medium	High	Low	Medium	High
Low	**1.098**	**1.062**	**0.842**	10.346	9.446	9.580
Medium	**1.089**	**0.982**	**0.770**	11.587	11.004	5.971
High	**1.009**	**0.963**	**1.030**	9.490	9.876	5.733
	β_{SMB}			$T-statistic_{SMB}$		
Low	0.115	0.093	**0.550**	0.652	0.522	1.900
Medium	0.101	0.116	0.109	1.042	0.948	0.520
High	-0.171	0.178	**0.750**	-0.975	1.299	2.354
	β_{HML}			$T-statistic_{HML}$		
Low	0.025	**0.332**	**0.377**	0.267	3.796	4.593
Medium	**-0.116**	**0.372**	**0.717**	-1.818	3.644	7.181
High	**-0.197**	**0.311**	**0.620**	-3.229	2.581	5.523
	$\beta_{QMJ_{II}}$			$T-statistic_{QMJ_{II}}$		
Low	**-0.720**	**-0.525**	**-0.647**	-6.916	-3.394	-4.942
Medium	0.058	**0.307**	0.017	0.525	3.880	0.165
High	0.154	**0.429**	**0.525**	1.508	4.122	2.566
	$Adjusted\ R^2$			$F-statistic$		
Low	0.883	0.869	0.760	64.893	57.316	27.891
Medium	0.866	0.827	0.733	55.899	41.611	24.359
High	0.778	0.744	0.619	30.703	25.719	14.815

This table presents the output of 9 time-series regressions with the yearly excess returns of each portfolio as the dependent variable. The explanatory variables used are the three factors from the Three-Factor Model from Fama and French (1993), namely $R_m - R_f$, the market risk premium; SMB, the size premium; and HML, the value premium, plus the QMJ factor developed by Asness et al. (2019), which captures the premium paid for high-quality stocks. Everything is on an annual basis. The first three factors are obtained from the work done by French (2021), which keeps track of these factors exactly how they were calculated originally. In this case, the fourth factor is represented by our adaptation, more in line with our analysis. This QMJ_{II} factor differs from the first one in its theoretical construction. Originally, this factor was built on quality and size, but in our adaptation, we have built it on quality and value to mimic our portfolios. The data included goes from June 1985 to June 2020, which in terms of returns results in 35 time periods which is the number of observations for each of the equation regressions. From top to bottom, we present the regression coefficients for each factor and each regression, and the respective T-statistic on the right side. The last set of lines is dedicated to the $Adjusted\ R^2$ and F-statistics. All models are estimated using Newey and West (1987) heteroskedasticity- and autocorrelation-adjusted standard errors with 3 time-lags based on a formula provided by Bali et al. (2016). Coefficients that are statistically significant at a 10% level of significance are indicated in bold.

Lastly, we present in Table 11 the estimation output for the last set of regressions. In this case, the format of the model remains untouched, and we add to the analysis our QMJ_{II} measure instead of the standard one, which was used on the model detailed in Table 10. The main objective here is to assess if there is any observable improvement to the model due to the adapted quality factor measure.

For the first three factors, the coefficient estimates change slightly. Every statistically significant coefficient for these three factors remained like that, and we actually observe an additional coefficient as significant related to the HML factor.

Regarding the QMJ factor, we see very clear and substantial changes in the coefficients, and we go from four out of nine statistically significant coefficients on the first set of models to six out of nine on this set.

Finally, looking at the R^2 of every model, we can see that for six models, there is an increase in the coefficient of determination which indicates that, overall, the quality of the model seems to improve slightly from the use of our version of the QMJ factor. Nonetheless, compared with the first set of regressions presented in this section, in Table 9, the increases are not impressive, being relatively significant in some models and completely insignificant in others. With all this in mind, we have to reject H5 because there is no clear sign of an improvement of our models when adding the QMJ factor. The explanatory power of the model is relatively strong all across the board but does not get much better when we add the new factor. It is important to underline one more time that despite everything, the version of QMJ we presented here, QMJ_{II}, revealed to be considerably more helpful in most regressions.

To put everything in good perspective and to finally end this section, we want to point out an inherent problem of this kind of analysis. As explained by Fama and French (1993) and also by Merton (1980), stock returns naturally tend to be volatile, and because of that, returns are many times not statistically different from 0, even when they are considerable, due to high volatility. This does not mean, however, that the models lack explanatory power, as already seen.

7. Conclusions

7.1 Findings

In this dissertation, we aimed to verify previous findings regarding the quality factor and increase and improve the amount of information disposable in the literature. Again, since it was developed and described relatively recently, it very well opens the opportunity to come up with new conclusions and discoveries, and we gladly took advantage of that.

In the first stage of our analysis, we detected that, as expected and already documented by Asness et al. (2019) and Berg (2020), high-quality stocks are relatively more expensive than low-quality, or junk, stocks. On average, and across all the models we presented, we found that for every additional quality score unit, the PBV increases from 0.243 to 0.285. It is very important to note that despite being statistically significant, the explanatory power of quality on prices is overwhelmingly low. With this finding, we confirm what has been stated by the authors cited at the beginning of the paragraph, which represent the only literature available directly related to this question.

We have also included two more topics in this first stage that studies the relation of quality and prices. One goes over the evolution of the price given to quality across our time horizon. It was suggested and shown by Asness et al. (2019) that quality was particularly cheap when stocks were expensive, as in the dot com bubble around 1998, and particularly expensive in moments of market distress and crisis, like in 2008. This suggests that investors see quality as a safer investment and that in times of fear and uncertainty, there are "flight-to-quality" events. However, within our sample, we could not encounter any meaningful result that would lead us to conclude such a finding. In fact, we concluded the opposite and we go very much against the only study that tests this hypothesis leaving a different conclusion in the literature.

Furthermore, we study the explanatory power of quality over time to understand better why it has been so low. We were able to discover a very clear down-trending relation between the R^2 measure of our regressions and the passage of time. We have no clear explanation for this observation. This assessment provides insight never before documented and builds upon the study of the explanatory power of quality on stock prices, previously mentioned by Asness et al. (2019) and Berg (2020).

The second stage of our analysis works similarly to Fama and French (1993). Kozlov and Petajisto (2013) and Novy-Marx (2013) suggest that there is a negative correlation between quality and value. In conjunction with the work from the authors of the Three-Factor Model, we have used this idea and developed a set of portfolios built on quality and value. Our first hypothesis concerning these portfolios regards their performance. Additionally, we assess the ability of a Three-Factor Model with an additional quality-minus-junk (QMJ) factor (Asness et al., 2019) to explain their returns.

We have concluded that portfolios built on value and quality present relevant excess, abnormal, and risk-adjusted returns. Out of nine, the overall best-performing portfolio is the one with the highest level of quality and the highest level of value, confirming the idea taken from Kozlov and Petajisto (2013) and Novy-Marx (2013). Even though the combination of high-quality and high-value was revealed to be the best one, the three best portfolios in every performance measure were always the ones with the highest quality measure for any given level of value. The second-best portfolio in terms of absolute performance was the one with high-quality and low-value (growth) stocks, showing that high-quality stocks were the edge of the strategy. Another interesting finding was that this latter portfolio (HQLV) was revealed to be less risky than HQHV, being the leader in terms of risk-adjusted performance. All these findings are new and contributive for the literature, given no other article or paper has tested this strategy. We would say that an assessment of this hypothesis, especially with results like this, brings significant value to the already existing literature on factor investing.

Concerning the addition of a QMJ factor to the Three-Factor Model, the conclusions were not very appealing to our hypothesis. By adding this factor to our models, the R^2 measure did not improve considerably and consistently. We also developed another version of the factor built slightly differently, and the panorama improved, but not by a large portion. The main contribution we have to offer with this part of the study is the new approach for building a QMJ factor. Additionally, we verify that the classic Three-Factor Model remains an exceptional asset pricing model with no significant need of adding any other explanatory variables, maintaining its simplicity and ease of application.

7.2 Limitations and Suggestions

In hindsight, there were a couple of limitations present during the development of this dissertation and when performing all these tests. Firstly, although we managed to work with a generous amount of data (almost 6000 companies across 41 years), our sample seems limited when comparing with Asness et al. (2019) and Berg (2020), which include more years and more countries in the analysis. This is particularly expressive during the first years of our sample, which begins in 1980.

The second limitation we can recognize regards the methodology used to perform the tests of our hypothesis H1. The Fama-MacBeth procedure is very time-consuming, simplified, and outdated to modern econometrics, as Berg (2020) admitted. An analysis with pooled or panel data treatment would probably be of simpler application and better results. We opted for the first method to stay truthful to the original papers by Asness et al. (2019) and Berg (2020). Additionally, our data would culminate in a very unbalanced panel, understandably, which makes it very difficult to work within a panel framework.

Lastly, not as a limitation of the work itself but more a limitation for future application is the fact that computing a quality score is an outstandingly long and time-consuming process. The amount of data needed and the number of steps until the final score make this a tedious process.

To finalize, we will leave a couple of suggestions for future research on this matter. The first one is about our discovery of a decreasing R^2 over time on cross-sectional regressions of a quality score on the PBV of stocks. We feel like we filled a gap in the literature, giving another step on trying to explain the low explanatory power of quality on prices, but at the same time, we encountered a new gap to be explored. An explanation for the decrease of this measure over time can be another important step in the right direction. The second suggestion concerns the negative correlation between quality and size we found and shown in Table 8. In the same way a negative correlation between quality and value previously found inspired this dissertation (even though that correlation appeared to be positive in our sample), this particular finding can also be the root of future work.

8. Bibliography

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance, 23*(4), 589-609.

Amihud, Y. (2002). Illiquidity and Stock Returns: Cross-section and Time-series Effects. *Journal of Financial Markets, 5*, 31-56.

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

Bali, T. G., Engle, R. F., & Murray, S. (2016). *Empirical Asset Pricing: The cross-section of stock returns*. John Wiley & Sons.

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

Basu, S. (1977). Investment Performance of Common Stocks in Relation to their Price-earnings Ratios: A Test of the Efficient Market Hypothesis. *The Journal of Finance, 32*(3), 663-682.

Bender, J., Briand, R., Melas, D., & Subramanian, R. A. (2013). Foundations of Factor Investing. Research Insights, MSCI.

Bender, J., & Nielsen, F. (2013). Earnings Quality Revisited. *The Journal of Portfolio Management, 39*(4), 69-79.

Berg, T. (2020). Quality Minus Junk – A Study Across 44 Countries (Master's Thesis, University of Stavanger, Stavanger, Norway). Retrieved from https://uis.brage.unit.no/uis-xmlui/handle/11250/2682140.

Blitz, D. C., & Vliet, P. (2007). The Volatility Effect. *The Journal of Portfolio Management, 24*(1), 102-113.

Blume, M. E. (1980). Stock Returns and Dividend Yields: Some More Evidence. *The Review of Economics and Statistics, 62*(4), 567-577.

Chan, K. C., & Chen NF. (1991). Structural and Return Characteristics of Small and Large Firms. *The Journal of Finance, 46*(4), 1467-1484.

Chan, K. C., Chen NF., & Hsieh, D. (1985). An Exploratory Investigation of the Firm Size Effect. *Journal of Financial Economics, 14*, 451-471.

Dechow, P., Ge, W., & Schrand, C. (2010). Understanding Earnings Quality: A Review of the Proxies, their Determinants and their Consequences. *Journal of Accounting and Economics, 50*, 344-401.

Fama, E., & French, K. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics, 22*(1), 3-25.

Fama, E., & French, K. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance, 47*(2), 427-466.

Fama, E., & French, K. (1993). Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics, 33*, 3-56.

Fama, E., & MacBeth J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. The *Journal of Political Economy, 81*(3), 607-636.

French, K. R. (2021). *Fama/French 3 Factors*. MBA - Tuck School of Business. Retrieved May 18, 2021, from https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Graham, B., & Dodd D. L. (1934). *Security Analysis*. McGraw-Hill Book Co.

Grantham, J. (2004). *The Case for Quality – The Danger of Junk*. [White Paper]

Greenblatt, J. (2010). *The Little Book That Still Beats the Market*. John Wiley & Sons.

HIS Eviews (2014). Fama-MacBeth Two-Step Regression.

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

Lewellen, J. (2015). The Cross-section of Expected Stock Returns. *Critical Finance Review, 4*(1), 1-44.

Lintner, J. (1965). The valuation of risk assets on the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics, 47*(1), 13-37.

Litzenberger, R. H., & Ramaswamy, K. (1979). The effect of personal taxes and dividends on capital asset prices: Theory and empirical evidence. *Journal of Financial Economics, 7*(2), 163-195.

Kozlov, M., & Petajisto, A. (2013). Global Return Premiums on Earnings Quality, Value, and Size.

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica, 34*(4), 768-783.

Merton, R. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics, 8*(4), 323-361.

Newey, W. K., & West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica, 55*(3), 703-708.

Novy-Marx, R. (2012). Quality Investing. *University of Rochester*, Working Paper.

Novy-Marx, R. (2013). The Other Side of Value: The Gross Profitability Premium. *Journal of Financial Economics, 108*(1), 1-28.

Novy-Marx, R. (2014). The Quality Dimension of Value Investing. *University of Rochester*, Working Paper.

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research, 18*(1), 109-131.

Piotroski, J. (2000). Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers. *Journal of Accounting Research, 38*, 1-41.

Ross, S. A. (1976). The Arbitrage Theory of Capital Asset Pricing. *Journal of Economic Theory, 13*, 341-360.

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

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

Sloan, R. G. (1996). Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings? *The Accounting Review, 71*(3), 289-315.

Treynor, J. L. (1962). Market value, time, and risk.

Vassalou, M., & Xing, Y. (2004). Default Risk in Equity Returns. *The Journal of Finance, 59*(2), 831-868.

Zhang, X. F. (2006). Information Uncertainty and Stock Returns. *The Journal of Finance, 61*(1), 105-136.