Quantitative analyses on portfolios simulations : how complex should the quality stocks definition be ?

Auteur : Delhez, Rémy-Baptiste
Promoteur(s) : Antonelli, Cédric
Faculté : HEC-Ecole de gestion de l'ULg
Diplôme : Master en sciences de gestion, à finalité spécialisée en Banking and Asset Management
Année académique : 2016-2017
URI/URL : http://hdl.handle.net/2268.2/3652

QUANTITATIVE ANALYSES ON PORTFOLIOS SIMULATIONS: HOW COMPLEX SHOULD THE QUALITY STOCKS DEFINITION BE?

Jury :
Promoter :
Cédric ANTONELLI
Reader(s) :
Julien GANTER
Boris FAYS

Dissertation by
Rémy-Baptiste DELHEZ
For a Master Degree in Banking and Asset Management
Academic year 2016/2017

Acknowledgements

First and foremost, I would like to thank my promoter and readers; Cédric Antonelli, Julien Ganter and Boris Fays, to take up their time to assess this final master thesis and to share their expertise.

I would like to especially express my gratitude to Boris Fays, reader of this work, who suggested me to deepen this topic. He provided me with his wise advice and showed lots of patience.

Many thanks as well to my brother, Martin-Charles, who used his professional knowledge in the banking industry to share his advice and showed his interest in this topic.

Abstract

This thesis aims at investigating the market anomaly quality as defined by Asness, Frazzini and Pedersen (2017) in their "Quality Minus Junk" factor. The undertaken study refines the quality stocks definition and its complexity. The concept of the quality anomaly has been for years arduous to portray, as its meaning is highly subjective and differs from one academician to another. Quality is occasionally not seen as a "pure anomaly" since it consists of an aggregation of numerous factors and ratios. This memoir is willing to enlighten this interpretation puzzle.

The basic concepts of market theories and portfolio management are introduced and discussed, just like the evolution of pricing models. The most distinguished anomalies, other than quality, are acquainted as a preface for the quality concept debate. Hence, the QMJ factor (Asness, Frazzini, & Pedersen, 2017) is analyzed in its three components; profitability, growth and safety. A replica of its ratios is built using SAS software with the goal to simulate Fama/French styled long-short portfolios based on a CRSP/Compustat dataset. The computed portfolios are regressed on QMJ and analyzed using SAS Miner software, along with descriptive statistics, correlations, cumulated returns and Sharpe ratios.

The results show that the growth component may be entirely dismissed without damaging the model. The safety factors greatly matter in the regressions and strengthen their roles into quality. Return on equity, return on assets and cash flows are profitability ratios that are significant in the definition as well. While the signals of gross profits are remarkably persistent and drove the quality performance in all empirical analyses. Hence, the source of quality is identified by these late six final ratios, cutting the complexity of the definition by more than two.

Keywords: quality, factor investing, portfolio simulation, QMJ, gross profit, market anomaly.

Table of content (short version)

1. Introduction

Portfolio management is a science that has become an art. The number of investment strategies have risen in the last decades and every practitioner tries to play his cards right. Both active and passive strategies are busy finding the excess return, thus alpha has grown into an obsession. Amongst the range of already observed market anomalies, exploited in each possible way, lies one market anomaly that has not been labelled with a proper and accepted definition yet: "Quality". The concept of quality is still nebulous for most academicians, yet is already in use within many funds to produce abnormal returns. Specifically, the ultimate aim of this work is to find the source of quality stocks and refine its definition. In other words, *How complex should the quality stocks definition be?* To address this puzzle, I based this study on an existing quality model:

The "Quality Minus Junk" factor (Asness, Frazzini, & Pedersen, 2017) seeks alpha by taking long positions on quality stocks and short positions on bad quality securities, named junks. This Fama/French styled spread (Fama & French, 1993) already demonstrates great performances, both on historical datasets and in its current use (AQR, 2017). The selection of this specific model releases an exhaustive definition of quality to begin our researches. Indeed, Asness et al. (2017) define the quality concept with three major components and 16 ratios that are going to be discussed and analysed further. This broad definition serves as a basis for the empirical and quantitative analyses that will try to find main drivers into quality stocks. Extra authors definitions will be encountered to challenge the quality vision. Some authors have simpler approaches and consider that gross profit leads the way to quality investing (Novy-Marx R. , 2013).

In section 2, we will recall the concepts needed to understand portfolios management and its theories. I will dedicate a section to the Efficient Market Hypothesis (Fama E. F., 1965) as this theory is challenged everyday by investment strategies. A brief disruption between active portfolios management and passive strategies will be introduced. Afterwards, we will complete this overview of market theories with two well-known pricing models: The Capital Asset Pricing Model (Sharpe W. F., 1964) and the Arbitrage Pricing Theory (Ross, 1976). Both of these models will enlighten our understanding of market movements, assumptions and expected returns. We will then focus on the area of study of this work: market anomalies. The most influential anomalies will be discussed and analysed: such as the size effect (Banz, 1981), the value effect (Graham & Dodd, 1940), the momentum anomaly (Carhart, 1997) and some calendar effects. This will allow to deepen pricing models with the Fama and French three factor (Fama & French, 1993), the Carhart four factor model (Carhart, 1997) and the Fama and French five factor pricing model (Fama & French, 2015). To conclude the market anomaly section, I will shortly discuss some behavioural finance statements. I will assign one abbreviated section on hybrid strategies, lying between active and passive portfolio management: smart betas. Finally, a deep focus on the market anomaly that is at the heart of this thesis is performed: the quality anomaly.

In section 3, each factor and ratio of the QMJ model are going to be decomposed and explained to make the understanding of the results and interpretations clear and straightforward. With the same idea in mind, section 4 describes the methodology used to extract the sample, to compute the ratios and the portfolios simulations.

Section 5 includes every empirical and quantitative analyses. I first run some descriptive statistics to get an overview of the variables behaviour. Then, we will have a look at the cumulated spreads returns to investigate the portfolios performances. The correlations between the variables and the QMJ model are evaluated and multiple linear regressions are run on the dataset sample. To extend the researches, I will compute long-only portfolios instead of spreads and interpret the obtained measures.

The findings and outcomes of this thesis are referenced in the results, conclusions and discussion sections (5, 6 and 7).

2. Literature review

The hypotheses and postulates from different authors will bring insights about the topic and will help to build this thesis.

2.1. Efficient market hypothesis

First, we need to define a concept developed by Eugene Fama without whom there would be no purposes to speak of market anomalies. We live in a world made of financial markets and theoretically, we have been taught that these ones are efficient (Fama E. F., 1965). Meaning that, decisions on these markets are made rationally and provide a correct, fair price for stocks (Yalcin, 2010). Markets should reflect, at any time, all available information. Thus, no one should be able to buy and sell, respectively undervalued and overvalued stocks, as prices are continuously adjusted based on information or rates changes. In other words, outperform what the market is already giving as returns, should be impossible. Eugene Fama and Kenneth French also showed that returns distributions in mutual funds on the US market are very analogous to what it would be if portfolios managers did not manage anything at all (Fama & French, 2012), which greatly confirms their hypothesis.

The EMH has been used for decades with the purpose to state the basis of many financial theories. In the literature, you may find three different kinds of EMH; the weak form, the semi-strong form and the strong form. The weak form of the EMH states that assets prices already take into account all available past information. Indeed, technical researches based on historical data are then irrelevant for analysts (Burton & Shah, 2017), as these data are already imputed to the stocks prices by definition. The semi-strong form suggests that prices already reflect past and present data. Finally, the strong form stipulates that prices also include non-publicly available information known as private information or "hidden data" (Reilly & Brown, 2011). Market efficiency could be considered as a simplification of the world, which stays true for investments purposes for most individuals.

However, this is still a so-called hypothesis because empirically, it has been proven wrong many times. In the empirical world, many investment strategies are done irrationally (emotionally) which causes unexpected movements on financial markets and opens the door to some "arbitrage" opportunities (the concept of arbitrage is discussed in section 2.7). There are even well-known movements that tend to repeat themselves over the time and seem to be risk-related or rational. For instance, several studies have shown that companies with a relatively small capitalization (small caps) tend to achieve abnormally high returns in comparison to what could be explained by the market (Hackel, Livnat, & Rai, 1994). These types of movements will be discussed in section 2.5. Generally, we name these movements “anomalies” when speaking of financial markets. While the validity of the EMH is still discussed nowadays, some have already tried to find comprehension in this debate. Paul Samuelson, economics Nobel Prize of 1970, already asserted that the EMH is much suited for individual stocks than for the aggregate market (Jung & Shiller, 2005). Other specialists keep feeding the debate stating the 2007 financial crisis was led by the belief in the EMH. One example: Jeremy Grantham, analyst for GMO investment firm, blamed the EMH as main responsible for the late financial crisis stating that this theory causes dangerous underestimation of breaking bubbles (Nocera, 2009). This statement was sided by Paul Volcker, president of the Federal Reserve until 1987, who claimed that there was an unjustified faith in market efficiencies (Volcker, 2012). The effect of the cited 2008 financial crisis will be encountered as the study sample of this work includes this period of time. Moreover, the sample is extracted from the U.S. market and the Lehman shock has been powerful and surprising (see section 7).

There are two main aspects while speaking of market anomalies. As mentioned earlier, some anomalies are dedicated to the behavioral aspect of investors. The rest of the time, anomalies are dedicated to the risk aspect of investors. Both types will be discussed in this thesis.

2.2. Active versus passive portfolio management

Before entering deeply into the definition of market anomalies, the concept of active and passive portfolio management should be discussed. The return an investor tries to reproduce can described as follow:

$$\begin{aligned}\text{Total Actual Return} &= [\text{Expected Return}] + [\text{“Alpha”}] \\ &= \underbrace{\underbrace{[\text{Risk-Free Rate + Risk Premium}]}_{\textit{Passive}} + [\text{“Alpha”}]}_{\textit{Active}}\end{aligned}$$

Equation 1

Where:

i. The Risk-Free Rate (RFR) is the theoretical rate of return of an investment with zero risk. An investor would not accept additional risk except if the new rate of return exceeds the RFR. However, even the safest investment will still carry a small amount of risk. Most of the time, the rate of return of a stable government bond is used as a proxy for the RFR.
ii. The Risk Premium is the excess of return compared to the RFR that the investor should obtain when taking extra risk. It represents the additional part of returns an individual wants while investing in riskier assets than RFR products.
iii. Alpha represents the amount of value the investors add to the placement: the active return. In other words, it is the difference between the expected return and the actual one, named excess return.

Most passive portfolios managers seek the expected return, thus the RFR and the risk premium in accordance with the level of risk they are willing to accept. Contrariwise, active investors attempt to "beat the market" by pursuing actual returns greater than the expected returns in accordance with the level of risk taken, often named risk-adjusted expected returns.

In practice, passive managers usually track a benchmark index with a traditional buy and hold strategy, known as indexing strategy. There is absolutely no willing into generating some alpha in this kind of strategies. Managers are judged on their performance by how well they tracked the returns of the specific index they were chasing, meaning the aim is to minimize the difference between markets actual returns and their investments or portfolios strategies. However, some passive strategies actually seek excess returns by implementing rules and selecting specific stocks, yet they do not imply day-to-day management.

Conversely, active portfolio managers try to outperform the market. The aim is to exceed the return given by the benchmark index on a risk-adjusted basis. There are many investment strategies that tend to create the alpha introduced above. However, two main types are standing out: tactical adjustments (for instance, market and sector timing) and stock-picking (Reilly & Brown, 2011). There are even hybrid investment strategies as well, lying between the active and passive management categories, like enhanced indexing or the nowadays far-famed smart betas discussed later on. There are extreme active portfolio management as well, seeking to isolate the alpha component of the return often named "Pure Alpha" strategies.

There is obviously a significant trade-off while choosing between active and passive portfolio management: the cost. The more actively the investments are managed, the more cost linked to it. Costs and limits are discussed later on as well (section 2.7)

2.3. Capital Asset Pricing Model (CAPM)

The Capital Asset Pricing Model or CAPM, developed more than 50 years ago by (Sharpe W. F., 1964), (Lintner, 1965) and (Mossin, 1966) while continuing Markowitz theories, is largely used in the finance sector to price securities in order to compute expected returns for assets given their risk. The main idea with the CAPM is that investors must be compensated in two manners; time value of money and risk. Here is its equation:

$$E(R_i) = RFR + \beta_i[E(R_M) - RFR]$$

Equation 2

Where:

i. $E(R_i)$ is the expected return.
ii. RFR the risk-free rate.
iii. β_I is the risk measure. It is the non-diversifiable portion of a stock risk relative to the market as a whole, the volatility named systematic risk.
iv. $E(R_M)$ is the expected rate of return of the market.

This equation induces the fact that in order to seek higher returns, taking more risk is the only way. β is indeed a very convenient measure; an asset with a beta of 1,20 simply has a volatility that is 20% higher than the average on the market. The CAPM is not very different from the formula seen in section 2.2. The return is still a computation of the risk-free rate and the risk premium. But the CAPM brings a simplification as it assumes that only the overall market risk premium ($E(R_M)-RFR$) matters and not a specific risk premium for each security. Yet this market risk premium may simply be scaled up or down by the value of the Beta. It dramatically reduces the amount of computations investors need to perform when evaluating their investments.

Before exposing the CAPM model assumptions, a quick differentiation between systematic and unsystematic risk will improve our understanding. Unsystematic risk is the kind of uncertainty that appears with the firm, industry or region, a professional invests in. This part of the risk can be diminished through diversification, and is most of the time assumed to be null in well thought portfolios management strategies. Systematic risk, sometimes named market risk or un-diversifiable risk, is the kind of uncertainty inherent to the whole market. It is the volatility implied by the day-to-day fluctuations in stock prices needed to generate returns (Sharpe W. , 1964).

The CAPM, alike many theoretical models, requires some assumptions. I am only quoting here the assumptions that will be ruled out by the APT model:

i. CAPM returns are normally distributed, meaning the values plot symmetrically and are situated around the probabilities mean
ii. Investors risk-return utility functions are only quadratic, meaning all investors utility functions are approached by at maximum a second degree function (concave or convex functions)
iii. There is a mean-variance efficiency, meaning all investors seek to invest in tangent points on the efficient frontier of their utility function. Hence, the selected portfolio will depend on this utility function

The CAPM has been for decades, and still now, the basis of many financial pricing models. Another interesting features from the CAPM is the ability to graphically represent the relation between risk and required rate of return using the Security Market Line or SML.

On the SML plot, the x-axis represents the beta factor, meaning the risk of the security or the entire portfolio. The y-axis represents the results of the CAPM, the expected return for this specific security or portfolio. Then, the market risk premium is graphically represented by the slope of the SML. Here is an example of a SML representation:

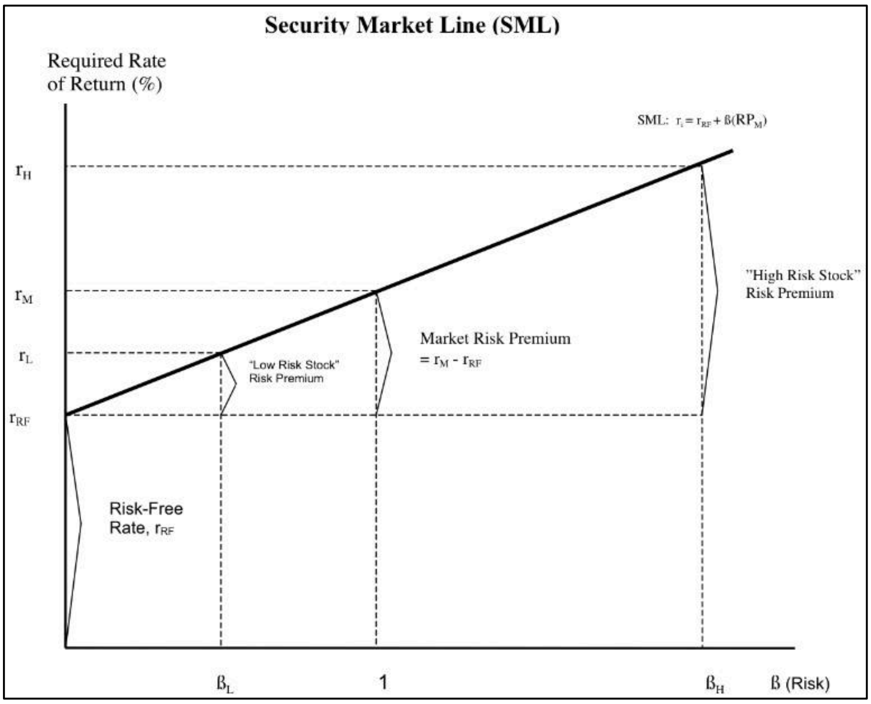

Figure 1- Security Market Line (SML)

The intercept shows the risk-free rate of return available on the market. Above the intercept, the security encounters a market risk premium represented by the slope of the line. The SML is a useful tool in determining if an asset is under or overvalued. Indeed, if a security relation between risk and expected return plots above the SML, the security is undervalued as an investor benefits from a greater returns for the inherent risk. On the contrary, overvalued securities plots under the SML since an investor would seek higher returns for the corresponding amount of risk.

Now that we have defined concepts related to the CAPM, we may have a closer look at "assets pricing" using this model. An analyst should compare the required rate of return over a specific investment horizon given by the CAPM and the estimated rate of return given by either fundamental or technical analysis. Making such a comparison, an analyst would be able to determine the appropriateness of the investment strategy. For instance, an analyst could compute this estimate with fundamental techniques by summing the forecasted capital gain and dividends gains. The difference between the estimated rate of return and the required rate of return is often named "expected alpha" or "excess return" as seen previously. Alpha may be positive, thus the security is undervalued, or negative, implying an overvalued stock. If the alpha is equal or close to zero, then the security plots directly on the SML and is properly valued in line with its systematic risk. Indeed, assuming that the CAPM provides fair results, an analyst that computes with fundamental techniques the estimated rate of return of an asset and observes that it is greater than the rate of return given by the CAPM, should expect this security to rise. Thus, buying this asset should be an appropriate investment strategy. The CAPM allows investors to implement fairly simple asset pricing strategies in order to invest properly.

2.4. Arbitrage Pricing Theory (APT)

In the last section, we highlighted the way the CAPM has enriched the investment management field. As previously mentioned, the CAPM has been one of the most useful financial models ever developed. Nonetheless, many empirical studies revealed shortcomings in this model as an explanation of the relationship between risk and return. A considerable challenge to the CAPM is the suggestion that, it is possible to use expertise on certain securities characteristics to develop profitable trading and arbitrage strategies, even after accommodating for investment risk as measured by betas (Wei, 1988).

Indeed, as mentioned in section 2.1., many empirical studies on simulated portfolios have shown that stocks with low market capitalizations (i.e., “small cap”) outperformed stocks with large amount of market capitalizations (i.e., “big cap”) (Banz, 1981; Hackel, Livnat, & Rai, 1994). In the same idea, simulations have documented that securities with low price-earnings ratios similarly outran high price-earnings ones (Basu, 1977). More recently, “value” stocks (i.e., the ones with a high book-to-market value ratios) bear to produce greater risk-adjusted returns than “growth” stocks (i.e., the ones with a low book-to-market ratio) (Fama & French, 1992).

These empirical exceptions bring us back to the EMH. In an efficient market, these returns differentials should not occur. Then, there are two possibilities; either markets are not efficient and the EMH is empirically wrong, or they are efficient, yet something went wrong with the way single factor models, alike the CAPM, measure risk. Given the love economists save for their Efficient Market Hypothesis theory, the second possibility was the most considered in the early 1970s. Thus, the aim of the financial academic community was to develop an alternative asset pricing model to the CAPM that is; as intuitive, that required limited assumptions, but may allow multiple risk factors.

The Arbitrage Pricing Theory (APT), originally developed by Ross (1976) and later extended by Huberman (1982), Chamberlain & Rothschild (1983), Chen & Ingersoll (1983), Connor (1984) and many other researchers, is the reasonable alternative to the CAPM and has three major assumptions:

i. Markets are perfectly competitive
ii. More wealth is always preferable to less wealth
iii. Asset returns can be expressed as a linear function of a set of K risk factors and all the unsystematic risk is diversified away.

Moreover, the required assumptions for the CAPM theory are not required anymore (that is to say; quadratic utility functions, returns normally distributed and the mean-variance efficiency of the portfolio). Thus, the APT seems to be simpler and may explain the differentials in security prices that were empirically shown, which was the aim of the development.

Prior to discussing what is going to be named anomalies, we provide here a brief review of the basics of the APT model. This model assumes that the stochastic process (meaning that each mathematical set is uniquely associated with an element) generating asset returns may be represented as a K factor model with the following form:

$$R_i = E(R_i) + b_{i1}\delta_1 + b_{i2}\delta_2 + \cdots + b_{ik}\delta_k + \varepsilon_i \text{ for } i = 1 \text{ to } n$$

Equation 3

Where:

i. R_i is the actual return on asset i during the specified time period
ii. $E(R_i)$ is the expected return for asset i if all the risk factors have no changes
iii. b_{ij} is the reaction of asset i's returns movements to a common risk factor j
iv. δ_k is the set of common factors that influences returns of all assets
v. ε_i is the unique effect on returns, the residuals or random error
vi. n is the number of assets

The raw mathematical expression of this model requires some explanations, at least for two components. As mentioned, δ terms are different risk factors that are common to all assets. For instance, δ might include the inflation rate of the market, the growth rate measured by GDP (gross domestic product), the political context or more frequently: interest rates. Thus, we may notice a major contrast from the CAPM. Indeed, the APT implies that there are many factors that affect returns while the CAPM entails that the only relevant risk variable is the asset beta; the covariance of this asset with the market portfolio.

Then, the b_{ij} terms determine how each asset responds to its particular j^{th} common factor: δ. Meaning that, even if common factors (δ) affect all assets, the impact is not similar for each. For instance, some assets will be greatly altered by growth while some others will not at all. Hence, the greater the b_{ij} term is, the more the security is influenced by this specific common factor.

Also, similar to many mathematical models and analogous to the CAPM, the APT model assumes that the unique effect, the random error ε, will be diversified away within large portfolios following the LLN (law of large numbers) and thus may be considered null as mentioned is assumption iii. Accordingly, the mathematical expression of the APT can be simplified and finally expressed as:

$$E(R_i) = \lambda_0 + \lambda_1 b_{i1} + \lambda_2 b_{i2} + \cdots + \lambda_k b_{ik}$$

Equation 4

Where:

i. λ_0 is the expected return on an asset with zero systematic risk, thus equals to the risk-free rate
ii. λ_j is the risk premium related to the j^{th} common risk factor
iii. b_{ij} is still the elasticity between the common risk factor and the specific asset, named factor betas or factor loadings

2.5. Market anomalies

In section 2.1., dedicated to the EMH, I introduced the existence of certain types of market movements that cannot be explained by the arguments of this theoretical hypothesis. In financial theories, these specifics movements are named "anomalies". According to Tversky and Kahneman (1986, p. 252) *"an anomaly is a deviation from the presently accepted paradigms that is too widespread to be ignored, too systematic to be dismissed as random error, and too fundamental to be accommodated by relaxing the normative system"*. Thus, a financial market anomaly is a pattern in stocks returns that is not predicted by a central paradigm or theory. Most of the time, the discovery of anomalies arises from empirical tests on historical dataset. Some skeptical academicians argue that, when anomalies are discovered, their effects struggle and finally disappear as the market assimilates them (Yalcin, 2010). However, most of the anomalies that will be discussed, continue to prevail for decades. Leaving financials with two options; markets efficiencies are a craze, or our models (alike the APT) seem to be less than complete while describing financial market equilibrium (Durlauf & Blume, 2008). Both options are likely to be true while the number of documented anomalies, already large, keeps growing.

Some of the most common market anomalies will be shortly described below. The aim of this section is not to assess the consistency (their immunity to researchers' statistical tests) yet to introduce the most noticeable and researched market movements without challenging facts veracity.

In the following sections, I introduce six market anomalies; size, value, momentum, profitability, investment and some "calendar effects".

2.5.1. The size effect anomaly

The size effect refers to the positive relationship between stock returns and small market value (equity) of firms (Banz, 1981). Banz was one of the first to experience this anomaly for the U.S. stocks. This effect relies on the theory that smaller firms (companies with a small market capitalization) tend to outperform larger companies. This anomaly was quickly used as a specific risk-return factor by the Fama/French three factor model (FF3) (Fama & French, 1993). The aim was to improve the CAPM pricing model by adding some admitted anomalies and making it stick to the empirical world a bit more. The created factor is known as Small Minus Big (SMB). This theory holds on the fact that small companies have a greater amount of growth opportunities. Also, small caps tend to have a lower stock price, leaving more room to substantial prices appreciations.

2.5.2. The value effect anomaly

The value effect specifies the positive relationship between stock returns and its ratio of accounting value measures divided by the market price of the stock. This anomaly refers to "value stocks", which are securities that tend to trade at a lower price on the market relative to its fundamental accounting ratios (e.g. dividend yield, earning, book price, ...). Most of the time, this anomaly is used in its simplest form known as "book-to-market effect". An investor would compare the book value of a company with the stock price on the market ($\frac{B}{P}$) and buy if this ratio is positive. The bigger the book-to-market, the cheaper the stock. The market price of the stock tends to catch up at least the book value, which makes a long strategy successful (buy and hold). This value effect is one of the first anomaly to be recognized and can be traced to at least the 1940s with "The Security Analysis" book (Graham & Dodd, 1940). This strategy was added to the CAPM pricing model as well by the FF3 model in order to improve the fairness of the CAPM, with the additional factor known as High Minus Low (HML). As explained in the methodology section, the aim of such factors (HML and SMB) is to construct a portfolio going long on high value (respectively small caps) stocks and short on low value (respectively big caps) stocks.

2.5.3. The Fama/French three and five factor, Carhart four factor models

The introduction of the two late anomalies (size and value) allows to lay the foundations of the Fama/French three factor model, that will serve as a basis in the methodology section and quantitative analyses (Fama & French, 1993). As mentioned earlier, the FF3 expands the CPAM pricing model, by adding size and value factors to the market risk factor already accounted. The aim was indeed to take into account these two admitted anomalies to compute a pricing model that is better at measuring the market return. Here its equation:

$$r = R_f + \beta\left(R_m - R_f\right) + \beta 1 SMB + \beta 2 HML + \alpha$$

Where:

i. $R_f + \beta\left(R_m - R_f\right)$ is the traditional CAPM model.
ii. $\beta 1 SMB$ and $\beta 2 HML$ are the size and value effects with loading coefficients betas, similarly to the APT loading factors.

Many academicians add their researches to the FF3. For instance, this is the case for the four factor model that extends the pricing model by implementing a momentum effect (Carhart, 1997), introduced below:

The momentum (MOM) effect assumes that if a stock is strongly going up in the past, there are great chances that this tendency will continue in the near future. This anomaly was implemented to the FF3 by adding a new factor Up Minus Down (UMD). This fourth factor is going long on stocks that perform well, and short on stocks that are the lowest performers, lagged one month. Here is the adjusted Carhart four factor model equation:

$$r = R_f + \beta(R_m - R_f) + \beta 1SMB + \beta 2HML + \beta 3UMD + \alpha$$ *Equation 5*

There is also a Fama/French five factor model, not taking into account the UMD anomaly adjustment, yet adding two new effects. The first effects is based on the profitability anomaly as described by Novy-Marx (2013). Profitable companies tend to outperform firms with low profitability ratios. This empirical research thus generates a new factor, Robust Minus Weak (RMW). The fifth factor is willing to take into account the level of firms investment. Conservative Minus Aggressive (CMA) is the return spread between companies that invest in a conservative way minus firms that tend to heavily invest (Fama & French, 2015). Here is the five factor model equation:

$$r = R_f + \beta(R_m - R_f) + \beta 1SMB + \beta 2HML + \beta 3RMW + \beta 4CMA + \alpha$$ *Equation 6*

2.5.4. Calendar effects

Calendar anomalies refers to market movements that are not described by any ratios. There are several calendar effects representing a collection of theories suggesting that certain days, months or time period of the year are subject to above-average price changes (Patel & Sewell, 2015). The most famous calendar anomaly is the "January effect" and suggests that securities prices increase during the month of January more than in every other month. This anomaly has been researched for decades by numerous academicians and tends to repeat itself years after years. The size anomaly described earlier has a special relationship with the January effect: *"One of the biggest challenges researchers pose to any interpretation of the size effect is that it mostly resides in January (…) all of the returns to SMB and the decile size spread are concentrated in January, with no evidence of any size effect outside of January.* " (Asness, Frazzini, Israel, Moskowitz, & Perdersen, 2015, p. 21).

Another well-known calendar anomaly is the “Halloween strategy”. This strategy refers to the theory that the period from November to April has significantly greater stocks growth. The saying of this Halloween strategy is to “sell in May and go away”, meaning that an investor trusting this strategy would buy around Halloween, hold and sell in May (Dichtl & Drobetz, 2014).

There is the “Monday effect” as well, suggesting that returns on the stock market on Mondays will mostly follow the trend of the previous Friday. At the contrary, some researchers refer to the “Weekend effect”, stating that stocks returns of Mondays are lower than those of the preceding Friday. This late effect would be due to the fact that companies tend to disclose bad news on Fridays when markets close and avoid to affect their stocks prices during the weekend (Olson, Mossman, & Chou, 2015).

As we may notice, there is a difference by nature of anomalies such as size, value, momentum, profitability, investment and market movements alike calendar effects. The five firsts refer to risk based anomalies, while the last one lies mostly on the behavioral side of the finance theories. These two concepts will be discussed in the next section.

2.5.5. Risk story and behavioral story

Behavioral finance is quite taboo. Most investors convince themselves that markets are rational and that their investment choices are making it this way. Following the assumptions of the EMH, the APT or even the development of pricing models alike the FF3, the quest to compute a model that perfectly sticks to reality begins with a risk story. The only adjustments someone may implement to a model based on efficient theories is obviously risk related. These equations adaptations are simply translating the correlation between risk and return. Hence, the market is still rational and “*the assumption of rationality has a favored position in economics. It is accorded all the methodological privileges of a self-evident truth, a reasonable idealization, a tautology, and a null hypothesis. (…) The advantage of the rational model is compounded because no other theory of judgment and decision can ever match it in scope, power, and simplicity.*” (Tversky & Kahneman, 1986, p. 273). The 2002 Nobel Memorial Prize in Economic Sciences Daniel Kahneman and his mathematical psychologist collaborator Amos Tversky made several researches on behavioral finance to contest rationality on financial markets. Their aim was to mitigate the fact that any significant observed violation of the model was assumed to be quickly eliminated by learning (making a reference to the weak-form efficiency of the EMH introduced in section 2.1.), or irrelevant to economics because market forces will automatically correct it.

There is absolutely no doubt that learning implies market forces to improve efficiency. However, this effectiveness take place under conditions: accurate and immediate feedback on decision making processes. While in reality, these feedback and their accuracy (that has to be taken into account by managers, politicians, …) are commonly delayed and lacking of precision.

Investors should bear in mind that a part of the market return will never be explained by a mathematical model. A part of the observed anomalies is probably due to irrational decision-making processes on the market which are an element of the humankind behavioral story.

2.6. Smart Betas

After having discussed the concept of anomalies, it is interesting to linger over another concept that lies between actively seeking anomalies and passively tracking an index. Indeed, Smart Betas are investment strategies considered as passive portfolio management, yet that pinch some characteristics to active portfolio management styles.

Smart betas are the fastest growing categories in the investment industry right now. And it is easy to understand why. This investment strategy is trying to passively follow indices while seeking to capture some of the anomalies, the well-known returns factors like size or value, by differing on the weighting. Most of the time, benchmarks or index trackers are cap-weighted. The loadings the manager gives to each security depends on the security market capitalization over the total market capitalization. In smart betas strategies, managers set in advance and in a transparent manner, a rule that will serve as a basis for the stocks weighting. The aim of this rule is, indeed, to capture some of the alpha generated by anomalies. In other words, smart betas are striving for the best of both worlds: passively manage a portfolio while capturing the alpha generated by active portfolios management (Hsu, Kalesnik, & Li, 2012). There are three main advantages of Smart Betas strategies:

i. It has a lower cost. The portfolio is managed passively hence there are less management and administration fees
ii. Diversification is generally better. The portfolio is constructed alike a benchmark, tracking entire indices which include large amount of stocks and are thus well-diversified by nature
iii. Alpha is generated. Smart betas portfolios get a quote part of the excess return

A Smart beta, also named "strategy index", is seeking to enhance risk-adjusted returns above cap-weighted indices. There is not a single approach to these kind of strategies, each fund has its own rules that rely on managers beliefs.

2.7. Arbitrage limits, costs and mispricing

After having discussed all these anomalies, all these possibilities to beat the market. We have to mitigate these options by discussing arbitrage limits, arbitrage costs and mispricing.

As discussed in previous sections, the presented anomalies in prices are particularly hard to reconcile with standard models. Indeed, the different theoretical models may slightly differ while predicting the relationship between risk and expected returns, adding factors after factors. Yet they all assume the law of one price. An asset with identical payoffs should be traded at the same price.

Arbitrage opportunities appear when there are differentials in prices for the same asset. In the efficient market theories, such arbitrage opportunities are automatically corrected by arbitrageurs (Shleifer & Vishny, 1997). Indeed, if such a profit was possible, investors (that all have the same curved utility functions) would seek this profit. At the end of the day, prices would match again. By definition, arbitrageurs take care of mispricing and market anomalies by re-establishing market efficiency.

Yet again, this is a different story empirically speaking. Trading or simply investing is not cost-free, thus arbitrageurs face financial constraints that prevent them from eliminating mispricing. *"Limits of arbitrage are commonly viewed as one of two building blocks needed to explain anomalies. The other building block are demand shocks experienced by investors other* *than arbitrageurs"* (Vayanos & Gromb, Limits of Arbitrage: The State of the Theory, 2010, p. 2). Typically, arbitrageurs face four main limits:

i. Cost of short-selling
ii. Margin constraint
iii. Cost of equity capital
iv. Trading commissions and cost of transactions

In addition to the cost of risk, that all investors bear because of the inherent relationship between risk and return, arbitrageurs deal with short-selling cost (i). Many arbitrage opportunities imply to short-sell, to sell and asset you do not actually own, entailing the investor to borrow money. This is known as margin constraint. While short-selling, an investor actually stands the borrowing interest rate and the risk to be obliged to honour a margin call (ii), allowing more capital risk.

This is implying a new constraint, the probability to be unable to raise equity (iii). Shleifer & Vishny (1997) studied the implications of constraints on equity capital for arbitrageurs' ability to exploit mispricing. Additionally, commissions and fees (iv) are, from the beginning and for every introduced model, completely ignored. Most financials actions imply entrance, exit and administration fees that affect the way an investor should compute its expected return. An investor willing to take a position that will bring returns due to mispricing, to a market anomaly, would maybe not take it after having implemented the cost of transactions, cutting off a part of the expected return. Research has been done for decades on the invalidation of arbitrage opportunities after implementing transactions costs (Leland, 1985).

I will not go deeper into these constraints as each introduced model do not take them into account, neither will I in the quantitative analyses. However, we should keep in mind that the risk is not the only cost an investor is facing. These additional costs have two contrary implications: it is harder to get profit from market anomalies on real markets as there are more costs embedded, while these costs make it also harder to eliminate anomalies and reach market efficiency.

2.8. Focus on quality

The specific market anomaly we are interested in is quality. Investing in quality stocks has seen a growing interest over the last decades. While the majority of the previously introduced anomalies have a universally accepted definition, this is not the case for quality. Each academician and practitioner have their own interpretation with their commonalities and differences. For some, one unique ratio (gross profit over assets) could be the main driver of the quality concept (Novy-Marx R. , 2013). For Jeremy Grantham, founder of the leading industry asset management firm Grantham, Mayo and van Otterloo, quality is a combination of high returns, stable returns and low debt (GMO, 2004). For others, we should look at undervalued ratios like the return on invested capital (ROIC) (Greenblatt, 2010). According to the QMJ factor, the presently studied model, quality is a combination of 16 accounting ratios that embody three aspects; profitability, growth and safety (Asness, Frazzini, & Pedersen, 2017).

The main reason for this definitions disparity is the fact that quality is not related to one factor or one market movement that researchers can specifically implement in pricing models. Quality is a merger between several concepts professionals give to a stock. The thought behind the use of the QMJ is that this model allows to start from an exhaustive definition of what should quality be made of. The study of the 16 computed ratios will grant enough conclusions or discussions. There are yet no theoretical explanations for this anomaly to occur, thus quality is intriguing, standing out because of its complexity. While at the same time, it is somehow common sense that quality companies are worth investing in.

Quality is often view as an attractive alternative to the really common growth factor that drove the market for years (Novy-Marx R. , 2013). Both growth investing and quality investing are investment strategies that imply growth factors. Yet growth strategies focus on that area and typically concerns only growth stocks or companies whose earnings are expected to rise with an above-average rate in comparison with the market growth. While quality does take into account growth, yet not only. Quality companies are most of the time stable profit generators. By being so, these stocks are less risky with a lower volatility. This lack of volatility tends to under-price the stocks. Finance theory would have suggested that investors must over-pay for quality stocks. But it is often the inverse phenomenon that occurs given their low risk profile. Asness et al. (2017) started from this consideration to build their definition of quality: a quality stock has characteristics that, everything else equal, an investor would be willing to pay a higher price for. The next question is: on which characteristics an investor would accept to spend more? As mentioned in the beginning of this section, the answer lies in three components:

i. Profitability: everything else equal, a profitable company should command a higher stock price
ii. Growth: growing firms should be more expensive
iii. Safety: Investors attribute higher prices to safe securities

The quality anomaly is part of the risk story of the rational finance theory. However, quality stocks by definition tend to be safer than usual ones. Thus, an investor would benefit from excess return while keeping the risk quite low. This is one of the reason why quality investing is heavily investigated recently. It somehow refutes the risk return relationship theoreticians relate to so much. *"Quality stocks earn higher returns and yet appear safer, not riskier, than junk stocks."* (Asness, Frazzini, & Pedersen, 2017, p. 7).

Quality stocks have other advantages: they benefit from the "Flight to Quality" effect. Flight to quality translates the phenomenon that during volatile times, during financial crisis or periods of drawdowns, investors risk aversion increases sharply (Vayanos, 2004). Thus, there is a literal flight from risky investment to quality investment, that appear safer to investors. This phenomenon increases the productivity of quality stocks and especially during poor economic conditions.

Given these facts and the growing influence quality investing has among professional of the financial sector, different models appeared during the last five years. MSCI, which is a market indices publisher company, builds indices selecting quality stocks by computing a quality score for every single security, based on three main fundamental factors: high returns on equity, stable year-over-year earnings growth and low financial leverage. This model's structure has been applied by MSCI in order to compute an index pooling quality/growth companies which reflects the performance of *"companies with durable business models and sustainable competitive advantages"* (MSCI, 2013, p. 3). This approach is close to QMJ model, investigated in section 3.

Quality investing is also often compare to value investing. Both strategies aim to acquire excellent and productive stocks at a cheap price. A traditional value strategy achieves so by buying securities at bargain prices, with a market value lower than a book-equity value. However, *"Quality investing exploits another dimension of value"* (Novy-Marx R. , 2013, p. 27). Quality investing achieves exceptionally cheap purchases by buying uncommonly productive securities. It seems relevant to keep an eye on quality, as quality measures and especially the gross profitability that will be introduced in the next sections, help common value investors to understand the difference between undervalued stocks and value traps (cheap stocks, that are cheap for a good reason). Trading on both sides, value and quality, allows investors to increase the expected return while decreasing volatility, hence the risk. Warren Buffet is often considered as a high-quality value investor, focusing on statistically cheap stocks, but highly profitable with a durable competitive advantage and hold them for the long run (Matthews, 2014).

Nonetheless, value investing has been empirically tested and proven to work in many situations, while this is not totally the case yet for quality investing. Numerous studies have shown that holding long portfolios with only the cheapest stocks produces returns far exceeding the stock market average for the same period. There is no equivalent objective proof concerning quality investing. This is mainly due to the fact that the definition of quality is still somewhat subjective and differs greatly from one professional to another. We are going to address this problem by refining the definition of quality and its complexity, so it may be accepted more largely.

3. Modeling quality stocks

3.1. Quality Minus Junk

As mentioned several times, the quality model we are relying on is the Quality Minus Junk model developed by Asness, Frazzini and Pedersen. The version that serves as a basis for this thesis is their most current version; their working paper draft dated from June 5, 2017. Although this paper is not published yet, the results of their researches are already in use in their fund to produce excess returns. Other studies related to the QMJ model are really interesting, like the fact that monitoring quality (and so controlling junk stocks) brings back the size effect described by Fama and French (Asness, Frazzini, Israel, Moskowitz, & Perdersen, 2015). Their working paper, while not published, is already controversial. The QMJ factor is introduced as a sharpened definition of the quality anomaly, yet quality as introduced by Asness et al. cannot be considered as an anomaly itself. Indeed, their definition implies an aggregation of 16 ratios and thus is too wide to represent one specific anomaly. For instance, the size effect is an anomaly computed using the market capitalization of the stock solely. In other words, the quality concept is considered as an anomaly because it generates abnormal excess returns, yet this is a combination of many ratios rather than a single factor itself. Hence, the aim is to shrink this definition and to find the source of quality by investigating the factors and sub-factors contented in this model and evaluate their necessity.

The ultimate aim is to identify these quality stocks and develop a long/short styled portfolio with the Fama and French (1993) methodology. Trusting their theory, a Quality Minus Junk (QMJ) portfolios that goes long on high quality stocks and short on low quality stocks (junks) should earn significant risk-adjusted returns. *"Using a variety of factor models ranging from the CAPM to a 7-factor model as our risk adjustment, we show that QMJ factors earn significant abnormal returns. Looking at factor exposures and performance during distressed market conditions, quality stocks appear safer, not riskier, than junk stocks"* (Asness, Frazzini, & Pedersen, 2017, p. 26).

3.2. Factors and ratios analysis

In this section, we report, detail and discuss each variable used in the QMJ model computations, that will serve as a basis for the next quantitative sections. We define each ratio individually. Most of the variables names correspond to CRSP/Compustat data items, some are just straightforward. Time related variables refer to years space-times.

3.2.1. Profitability ratios

The profitability factor is composed of six ratios; Gross Profit Over Assets (GPOA), Return on Equity (ROE), Return on Assets (ROA), Cash Flow Over Assets (CFOA), Gross Margin (GMAR) and finally Low Accruals (ACC).

$$GPOA = \frac{(REVT - COGS)}{AT}$$

Equation 7

Where:

i. REVT is the total revenue
ii. COGS is the cost of goods sold
iii. AT is the total assets

The use of GPOA ratios in a study of return based anomalies is supplemented by literature. It has been demonstrated that securities with high profitability ratios tend to outperform (Novy-Marx, 2013).

$$ROE = \frac{IB}{SEQ}$$

Equation 8

Where:

i. IB is the income before extraordinary items, approximation of the net income
ii. SEQ are stockholders' equity or else named book-equity

Concerning the ROE, we may notice that IB is used to approximate the Net Income, even if the Net Income (NI) is available in CRSP/Compustat database. This choice is not accounted for in the working paper. In order to stay consistent with the measures, I am using IB as well in the quantitative analysis. We should keep in mind that this measure of the Net Income has not been adjusted by accounting changes, discontinued operations, extraordinary items and related taxes.

$$ROA = \frac{IB}{AT} \quad \textit{Equation 9}$$

Alike ROE, ROA uses IB instead of NI and will be used for each ratios.

$$CFOA = \frac{(IB + DP - \Delta WC - CAPX)}{AT} \quad \textit{Equation 10}$$

Where:

i. DP is the depreciation
ii. ΔWC is the change in working capital
iii. CAPX are the capital expenditures

The working capital is computed as follow:

$$WC = ACT - LCT - CHE + DLC - TXP \quad \textit{Equation 11}$$

Where:

i. ACT are the current assets
ii. LCT are the current liabilities
iii. CHE are the cash and short terms instruments
iv. DLC are the short term debts
v. TXP are the income taxes payable

The change in working capital is computed over one year. Thus, making data not available for the first year of the sample.

$$GMAR = \frac{(REVT - COGS)}{SALE} \quad \textit{Equation 12}$$

Where:

i. SALE is the total sales

$$ACC = \frac{-(\Delta WC - DP)}{AT} \quad \textit{Equation 13}$$

A minus is used before the equation, as for the accruals, the relationship between the higher the better is inversed. The accruals on the balance sheet are liabilities and non-cash-based assets, in others words, the part of earnings that are non-cash. For instance; account payables, account receivables, goodwill or deferred taxes. This part of non-cash earnings is the share that will never be spent on dividends, share repurchases, debt payment or simply reinvest. Thus, from an investor point of view, the lower the accruals, the better. The though behind the low accruals is the old saying "Cash is King". Hence, as all the quantitative analyses are based on the flipped phenomenon (e.g. the higher the GPOA the better), a minus is placed beforehand this equation. The use of low accruals in the profitability factor is advocate by the fact that firms with low accruals are less likely to suffer subsequent earnings disappointments and tend to outperform their peers with high accruals (Richardson, Sloan, Soliman, & Tuna, 2005).

3.2.2. Growth ratios

Growth ratios are computed on the same ratios as the profitability section. A five-year basis is implemented for these computations. However, it will be mentioned in the methodology section that I used in the SAS programming a three-year basis. The choice of the denominator for the growth computations will be discussed as well in the methodology section. Hereafter are introduced the components of the growth factor as developed by QMJ authors:

$$\Delta GPOA = \frac{(GP_t - GP_{t-5})}{AT_{t-5}}$$ *Equation 14*

Where:

i. GP is the Gross Profit equals to REVT-COGS as in equation 7. GP_t is the value for GP in time T and GP_{t-5} is the value five years earlier

$$\Delta ROE = \frac{(IB_t - IB_{t-5})}{SEQ_{t-5}}$$ *Equation 15*

$$\Delta ROA = \frac{(IB_t - IB_{t-5})}{AT_{t-5}}$$ *Equation 16*

$$\Delta CFOA = \frac{(CF_t - CF_{t-5})}{AT_{t-5}}$$ *Equation 17*

Where:

i. CF equals $(IB + DP - \Delta WC - CAPX)$ as in equation 10.

$$\Delta GMAR = \frac{(GP_t - GP_{t-5})}{SALE_{t-5}} \qquad \textit{Equation 18}$$

There is no growth ratio for the corresponding profitability ratio ACC (accruals). The authors do no account for the choice of not selecting low accruals as a growth ratio. The computations of these five growth ratios draw its sources and its relevance in the quality concept definition from the fact that growing firms tend to outperform firms with poor growth (Mohanram, 2005).

3.2.3. Safety ratios

The safety factor is composed of five ratios; Low Beta (BAB), Low Leverage (LEV), low credit risk score computed with Ohlson's score (O) and Altman's score (Z) and Low Earnings Volatility (EVOL).

$$BAB = -\beta \qquad \textit{Equation 19}$$

Where:

i. $-\beta$ is the market beta

A minus is placed beforehand the beta following the same reasoning as in equation 13. The lower the beta, the safer the stock. Betas are approximated with the rolling one-year standard deviation and the rolling five-year three-day correlations. As it will be mentioned in the methodology section, I do not have access to daily data samples. Hence, this measure is not part of this study.

$$LEV = \frac{-(DLTT + DLC + MIBT + PSTK)}{AT} \qquad \textit{Equation 20}$$

Where:

i. DLTT are the long-term debts
ii. DLC are the short-term debts
iii. MIBT are minority interests

iv. PSTK are preferred stocks

Once again, the low leverage (LEV) expressed the situation where the lower the leverage ratio, the safer the company, thus the expression is negative. Most of companies are willing to use debt in order to finance their operations. By doing so, a company increases its leverage ratio, the amount of investment may increase without having to raise equity. GMO has shown that since 1965, the least levered companies have an average ROE 5% higher than the most levered companies and thus confirms that safety is one of the main drivers to investment returns (Joyce & Mayer, 2012).

$$Ohlson's\ O\ score = -(-1{,}32 - 0{,}407 * \log\left(\frac{ADJASSET}{CPI}\right) + 6{,}03 * TLTA - 1{,}43 * WCTA + 0{,}076 * CLCA - 1{,}72 * OENEG - 2{,}37 * NITA - 1{,}83 * FUTL + 0{,}285 * INTWO - 0{,}521 * CHIN)$$

Equation 21

Where:

i. ADJASSET is the adjusted total asset (+10% of the difference between market equity and book equity)

ii. CPI is the price index in the country

iii. TLTA the total debt divided by the adjusted total asset

iv. WCTA is current assets minus current liabilities divided by the adjusted total asset

v. CLCA is current liabilities over current assets

vi. OENEG is a dummy variable equals to 1 if liabilities>assets, 0 if not

vii. NITA is the net income over total assets

viii. FUTL is net income before tax over total liabilities

ix. INTWO is a dummy variable equals to 1 if net income is negative, 0 if not

x. CHIN is the change in net income over a year

$$Altman's\ Z\ score = \frac{(1{,}2 * WC + 1{,}4 * RE + 3{,}3 * EBIT + 0{,}6 * ME + SALE)}{AT} \qquad \textit{Equation 22}$$

Where:

i. RE are retained earnings
ii. EBIT are the earnings before interest and taxes
iii. ME is the market equity

As it will be discussed in the methodology section, from the two late credit risk ratios and for the purpose of this study I will only be using Altman's Z score. I have access to all the necessary data, which is not the case with Ohlson's O score.

$$EVOL = -\sigma ROE \qquad \textit{Equation 23}$$

EVOL is the standard deviation of quarterly ROE over 60 quarters. Rolling quarterly data are not available and would imply to lose five years of data as well, thus this measure is set aside of the quantitative analysis.

The first safety measures used in this thesis, low leverage, relates to the literature that firms with low leverage tend to have a higher alpha (excess return) than firms with high leverage ratio (George & Hwang, 2010; Penman, Richardson, & Tuna, 2007). The second measure, the credit risk score Altman's Z, refers to the literature that firms with high credit risk tend to under-perform (Altman, 1968; Ohlson, 1980; Campbell, Hilscher, & Szilagyi, 2008).

4. Methodology

The methodology used in this thesis follows the process of portfolios simulations according to the Fama/French methodology (Fama & French, 1993). I retrieved a sample composed solely of U.S. data imported from New York Stock Exchange (NYSE). It is essential to select a specific local area to study, as risk factors behave generally differently from one market to another (Griffin, 2002). All available common stocks are captured from the Center for Research in Security Prices (CRSP), which maintains one of the largest and comprehensive historical stocks market database. CRSP data are merged with Compustat database to avoid missing values. Due to data availability, I am using a sample from 2001 to 2015. From these databases were extracted all the necessary financial ratios to compute the QMJ components as seen in section 3.

The dataset is imported into SAS Studio software, online version of SAS granted with student's credentials on SAS portal. SAS is the analytic software that will allow to manipulate CRSP/Compustat data in order to compute the different profitability, growth and safety ratios, as well as simulating the portfolios construction following the Fama/French methodology and analyse the result with statistical programming.

The first step is to reproduce each ratio that is used is the QMJ model as presented in section 3.2. For instance, concerning the profitability component of the QMJ model, I computed with SAS from CRSP/Compustat dataset sample the six ratios needed; GPOA, ROE, ROA, CFOA, GMAR and ACC. The computations of these ratios follow the equation given in section 3.2.1. The created SAS code to perform these computations can be read in Appendix A. This programming generates new columns in the dataset corresponding to each ratio. The same process is used to compute growth (Appendix B) and safety ratios (Appendix A). However, these two last need more explanations.

As seen in section 3.2.2., growth ratios are computed on the same measures as the profitability ones, but with a five-year rolling evolution formula. Since I am using a sample with security data sample from 2001 to 2015, taking a five-year basis would make my results lose the five first years of information. The first growth ratio would only appear for the beginning of the year 2006. Hence, I decided to take a three-year basis to gain two years from the dataset. The CFOA and ACC ratios both need one year of latency as well. Indeed, they are composed of the change in working capital over one year. This forced me to get only data for the change in working capital with one year latency and the growth ratio with three-year latency, losing in total only the four first years of data, instead of six. The formula for the growth ratio of GPOA then become:

$$\Delta GPOA = \frac{GP_t - GP_{t-3}}{AT_{t-3}} \qquad \textit{Equation 24}$$

Where:

i. Δ denotes the change in Gross Profit Over Asset, thus the growth
ii. GP is equal to the Gross Profit

All the others growth ratios (G_ROE, G_ROA, G_CFOA, G_GMAR) were computed following the same idea. Another important topic to mentioned, this way of computing growth refers to the computations made by Asness et al. (2017) in their Quality minus Junk working paper. We may note that the denominator is the total asset in t-3. Which is an odd way to compute a growth ratio, as the denominator will mostly differ for each ratio. Most of the time, a growth ratio is computed with the denominator equal to the numerator in t-x. The expected computation would look like this:

$$\Delta GPOA = \frac{GPOA - GPOA_{t-3}}{GPOA_{t-3}} \qquad \textit{Equation 25}$$

However, the technique used by the QMJ factor to compute growth ratios permits to understand the source of the variation on the numerator side, as the denominator stays the same in the growth and non-growth formula. The ultimate aim of this thesis is to find what ratios drive the QMJ model the most, it is important to stick to Asness et al. (2017) methodology in order to eliminate most biases and obtain meaningful results. Thus, I will use their equations to compute growth ratios.

The computations of safety ratios require some explanation as well. Two out of the five ratios, low beta (BAB) and earnings volatility (EVOL), requested daily information that I do not have access to. Moreover, the safety score is made out of two credit risk related ratios, the Ohlson-O score and the Altman-Z score. As one credit risk ratio seems enough and as the Ohlson-O score requires data that I do not have access to (prices indexes by country), I used the Altman's Z score as credit risk related ratio for the safety factor. Thus, the low leverage (LEV) ratio and the Altman's Z score (Altman_Z) were computed in SAS following the equation seen in section 3.2.3.

The code used to simulate the Fama/French portfolio construction is made out of the open source code from Wharton Research Data Services (Palacios & Vora, 2009). An example of the generated Excel file presenting the ROA ratio as a Fama/French portfolios factor can be found in Appendix C. On these Excel sheets, we may observe the 6 value-weighted portfolios simulated, denoted by HH, HL, HM, LH, LL, and LM. The portfolios are refreshed every calendar month and rebalanced to maintain value weights. The breakpoints to construct the portfolios simulations are the 30^{th} and 70^{th} percentile (Fama & French, 2015). It allows to select only the highest performing stocks minus the lowest, even if it classifies "mid-range" performing stocks. The first letter is always the size, meaning the market capitalizations of the hold securities. "H" for high market capitalizations and "L" for low ones (big caps and small caps). The second letter acts for the specific ratios targeted. In the appendix example, "H" notify high ROA value stocks and "L" low ROA value ones. We are not concerned with medium portfolios ("M") in the following computations, containing securities with medium ranked ratios.

The different ratios, parts of the QMJ factor, then become factors returns themselves by averaging the returns on the two high quality portfolios minus the average on the two low-quality (junk) portfolios. Following the ROA example, here is its equation:

$$QMJ\ with\ ROA = \frac{(Big\ Quality\ ROA + Small\ Quality\ ROA)}{2} - \frac{(Big\ Junk\ ROA + Small\ Junk\ ROA)}{2}$$

Equation 26

Where:

i. Big Quality ROA means big caps (high market capitalizations stocks) with a high value of return on assets
ii. Small Junk ROA means small caps (low market capitalizations stocks) with a low value of return on assets

In other words, the ROA factor is the average of great ROA values minus low ROA values for the two sized portfolios. When using this formula for every calendar month, the monthly returns are obtained for this specific factor. I then reproduced this procedure for each ratio and end up with QMJ and Fama/French styled portfolios with monthly returns for the 13 studied ratios.

The next step is to clean the new dataset and gather in one Excel sheet the returns for all the ratios. I add up a new column, filled with the corresponding returns of the QMJ factor. Returns from the QMJ factor are extracted from the authors' fund management website: Applied Quantitative Research management (AQR, 2017). A sample of these data can be found in Appendix D. With this clean dataset gathering every needed monthly return, I will be able to analyse their performances with descriptive statistics, statistical tests and regressions. The software I used to run these statistics is SAS Enterprise Miner. An example of the generated diagram and nodes in SASEM is given in Appendix E.

For many individuals, including myself, the complexity of financials markets and especially models have become an obstacle. We may ask ourselves if the stock selection research has not gone too wild and if a 16 ratios factor is really needed in order to define quality stocks. Moreover and as mentioned earlier, for other researchers, quality investing may have a way simpler look: *"All of the best-known notions of quality contribute, at least marginally, to investment performance. Gross profitability generally contributes the most"* (Novy-Marx, 2013, p. 28). According to Robert Novy-Marx, most of the quality anomaly performance may be attributed to the gross profit.

Following these considerations, the results of my statistical analyses on these QMJ ratios will bright to light to where the definition of quality stocks should or should no stop.

5. Quantitative analyses and results

The created dataset using SAS is imported in SAEM. The Excel dataset file can be found in Appendix D and the corresponding file in SASEM can be read in Appendix E. I am first analysing the samples some descriptive statistics to get an overview of the dataset.

5.1. Descriptive statistics

The following table shows the results after running the exploratory statistics node:

Variable	Role	Mean	Standard Deviation	Non Missing	Missing	Minimum	Median	Maximum	Skewness	Kurtosis
ACCFF	INPUT	0.045829	1.335744	126	0	-4.30818	0.094829	3.756096	-0.24565	1.343891
Altman_ZFF	INPUT	0.269671	2.110057	126	0	-4.82606	0.355764	9.075002	0.411607	2.145685
CFOAFF	INPUT	0.138718	2.027371	126	0	-4.08775	-0.10692	6.622445	0.603962	0.7671
FF3	INPUT	0.093175	3.324862	126	0	-19.72	0.17	8.75	-1.6391	9.868514
GMARFF	INPUT	-0.10932	1.408393	126	0	-4.89243	-0.1917	4.750332	0.393799	1.643631
GPOAFF	INPUT	0.400578	1.769232	126	0	-4.15509	0.188529	6.850877	0.590006	1.622797
G_CFOAFF	INPUT	0.03627	1.284895	126	0	-2.78495	-0.02204	3.144084	0.11563	-0.47462
G_GMARFF	INPUT	-0.0919	1.475658	126	0	-4.10933	-0.12551	4.156014	-0.01113	0.507778
G_GPOAFF	INPUT	0.084513	1.810258	126	0	-4.59234	0.160363	6.703466	0.411158	1.200797
G_ROAFF	INPUT	0.074198	1.832363	126	0	-3.95084	0.030331	4.850693	0.066734	-0.01281
G_ROEFF	INPUT	0.024491	1.702681	126	0	-4.64784	0.141111	4.419517	-0.22769	0.443816
HML	INPUT	-0.19101	2.134434	126	0	-8.67524	-0.20154	7.031886	-0.12278	2.388369
LowLevFF	INPUT	0.030344	1.838395	126	0	-9.78816	-0.01921	6.223201	-0.68634	7.025989
MKT	INPUT	0.598019	4.508585	126	0	-18.528	0.91068	11.61939	-0.74597	2.176201
ROAFF	INPUT	0.190012	2.333744	126	0	-5.37639	0.147182	8.459843	0.471972	0.841998
ROEFF	INPUT	0.082377	1.758717	126	0	-4.84625	0.03715	5.544344	0.183186	0.955247
SMB	INPUT	0.022883	2.045743	126	0	-5.39083	-0.08373	6.52192	0.151852	0.332318
UMD	INPUT	0.225805	4.941116	126	0	-34.5844	0.832344	11.11481	-3.04843	19.50608
rf	INPUT	0.102903	0.153223	126	0	-0.00083	0.01	0.4175	1.157916	-0.42686
QMJ	TARGET	0.310688	2.688041	126	0	-7.2658	-0.05577	9.034384	0.278532	1.559903

Table 1 - Descriptive statistics

SASEM includes 20 variables in total. The 13 ratios are displayed, nomenclated with "FF" terminology to specify they are made out of Fama/French styled portfolios. I added the QMJ factor retrieved from AQR website as a target in this dataset. The risk-free rate "rf" and the market return "MKT" are implemented as well. The risk-free rate is computed based on the US treasury bill for the corresponding period. The market return is the excess return of the market regarding this treasury bill. I also added factor from the FF4 Carhart model; SMB factor, HML, UMD and the agglomerate FF3 factor to go further in the analysis. All the measures from the dataset are displayed in percentages. Every abbreviation used are define in the lexicon at the end of this thesis.

The "Mean" column displays the different monthly returns means of each portfolio. We may definitely notice that some ratios are standing out from the others. The credit score Altman_Z (0,27), CFOA (0,14), GPOA (0,40), MKT (0,6), ROA (0,19), UMD (0,22) and the factor QMJ itself (0,31). However, some of the ratio benefiting from an interesting mean have an important standard deviation as well. MKT (4,51), UMD (4,94) have a particularly high returns volatility. The standard deviation of QMJ (2,68) is somehow quite high as well in comparison with other variables, which would mean that QMJ as a whole is more volatile that its components taken apart. We may also notice some huge drawdowns with the minimum values for MKT (-18,53) and UMD (-34,58). Thus, the market return and the momentum effect are subjected to high volatilities and exceptionally low returns period of time, which is usual considering the particularly bearish market on the studied period.

It is not surprising to observe that these two late factors have abnormally long left-tailed negative skewness, MKT (-0,74) and UMD (-3,05). Hence, the distributions of their returns display a greater chance for negative return occurrences, which correlates their standard deviations and minimum values. On the contrary, some of the ratios present noticeably right-tailed positive skewness; CFOA (0,60), GPOA (0,59), G_GPOA (Growth of GPOA, 0,41) and ROA (0,47). Thus, the chance to observe extremely negative returns occurrences is not likely. Looking at the distinct kurtosis values, LowLev (7,03) and UMD (19,51) kurtosis are unusually large. It translates the fact that high magnitude returns are more likely to occur. I have to mitigate the outstandingly poor statistics measures from the momentum factor (UMD): the study period experiences the 2007-2008 financial crisis which is distinctly rough for momentum factors.

5.2. Cumulative spreads returns

After having analysed some descriptive statistics to get a better performance understanding, I computed the cumulated returns to obtain an improved vision of these performances. Here are charted the long/short portfolios returns for profitability components:

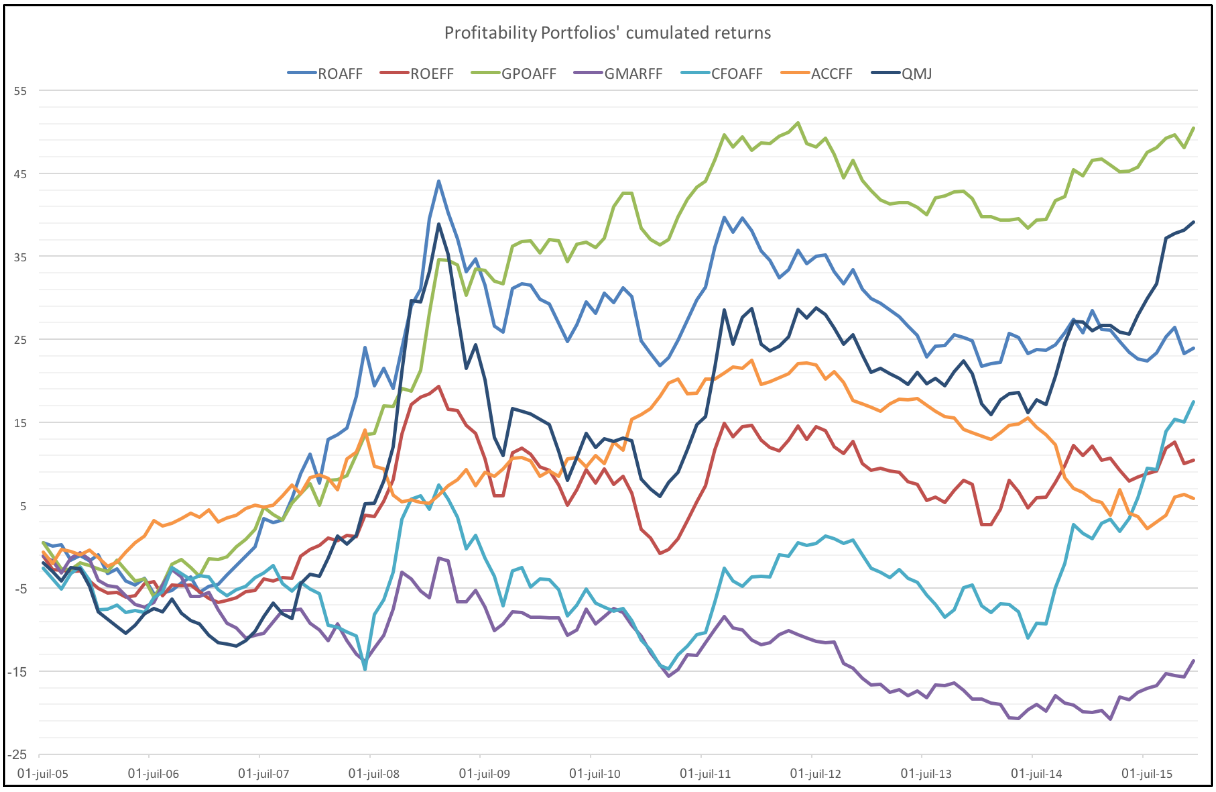

Figure 2 - Spreads cumulative returns (profitability)

We may immediately notice the extreme performance of GPOA (50,47), reaching more than 50% of returns on the studied period, giving Novy-Marx (2013) a solid support. GPOA is even recovering abnormally fast from the period 2008-20011 and its bearish market. The QMJ (39,15) factor is performing well, outrun by ROA (23,94) most of the time but finally rising off from mid-2014. CFOA (17,48) has especially great performances from 2014, almost reaching ROA at the end of the period. We may acknowledge the poor performance from GMAR (-13,77) as well, that never took over the 2008 financial crisis and produces strongly negative returns.

I computed the same cumulated returns with the growth and safety components of quality stocks:

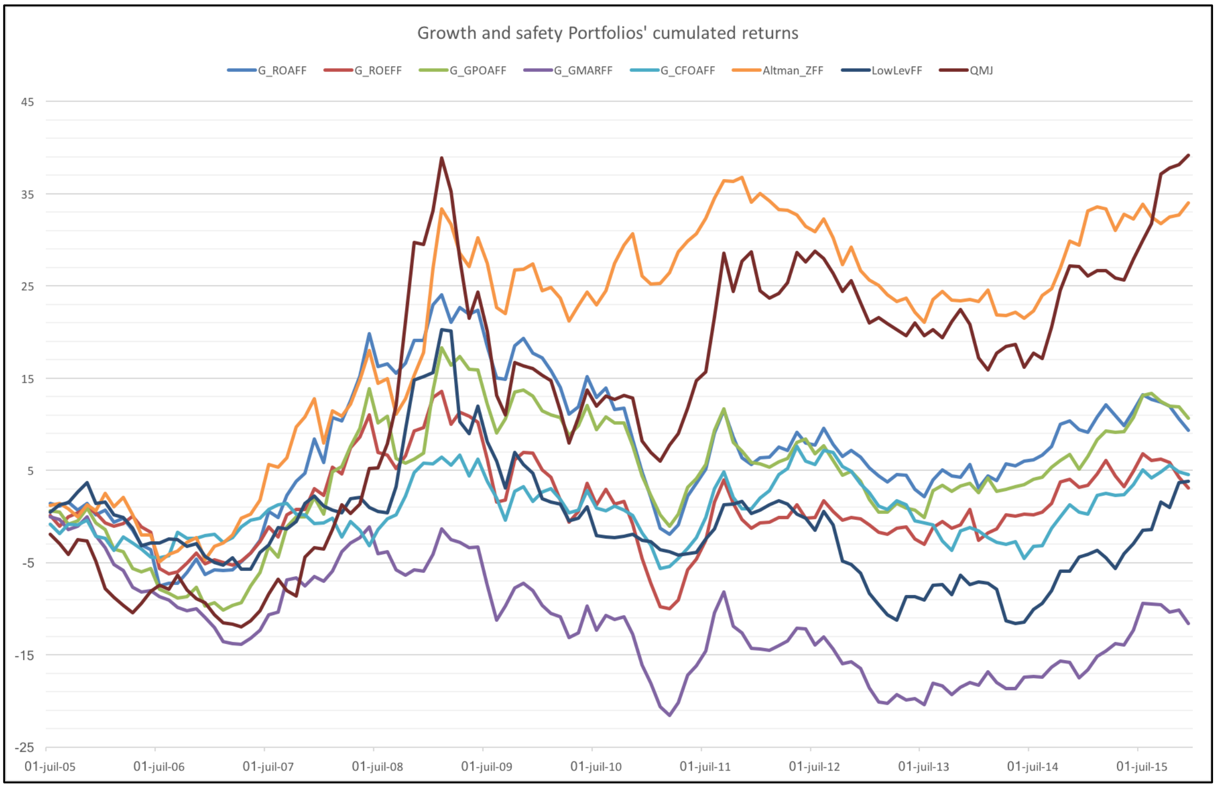

Figure 3 - Spread cumulative returns (growth & safety)

This is a different story. Most of the growth ratios portfolios are generating fairly poor performances. Ranging from G_GMAR (Gross Margin growth, -11,58) to Altman_Z (credit risk score, 33,99). It seems like the gross margin, in value and growth, are particularly producing significantly negative returns on that period. Hence, the gross margin is not likely to be correlated with the QMJ excess returns and potentially left out of the quality definition. At the contrary and surprisingly, Altman_Z produces returns close to QMJ. Thus, it seems that company with a specifically low credit risk tend to perform really well on that period. The credit risk is most of the time used as a borrowing-proof evidence for banks and loan-related firms, yet a long/short portfolio on Altman's Z score seems to select performing quality companies. G_GPOA (Gross Profit over Assets growth, 10,65) is the third performing ratio on the plot, though does not hold a candle to GPOA alone. Looking at the growth ratio as a whole, none of them seem to perform outrageously. They all suffer from the financial crisis and barely recover from it. Now, we have a better vision of the ratios and factors' performances. Yet, the matter here is to observe how they correlate to the QMJ factor and where is quality inside this model.

5.3. Pearson's correlation coefficients

The following table displays the Pearson correlation coefficients computed in SASEM with QMJ as a target variable:

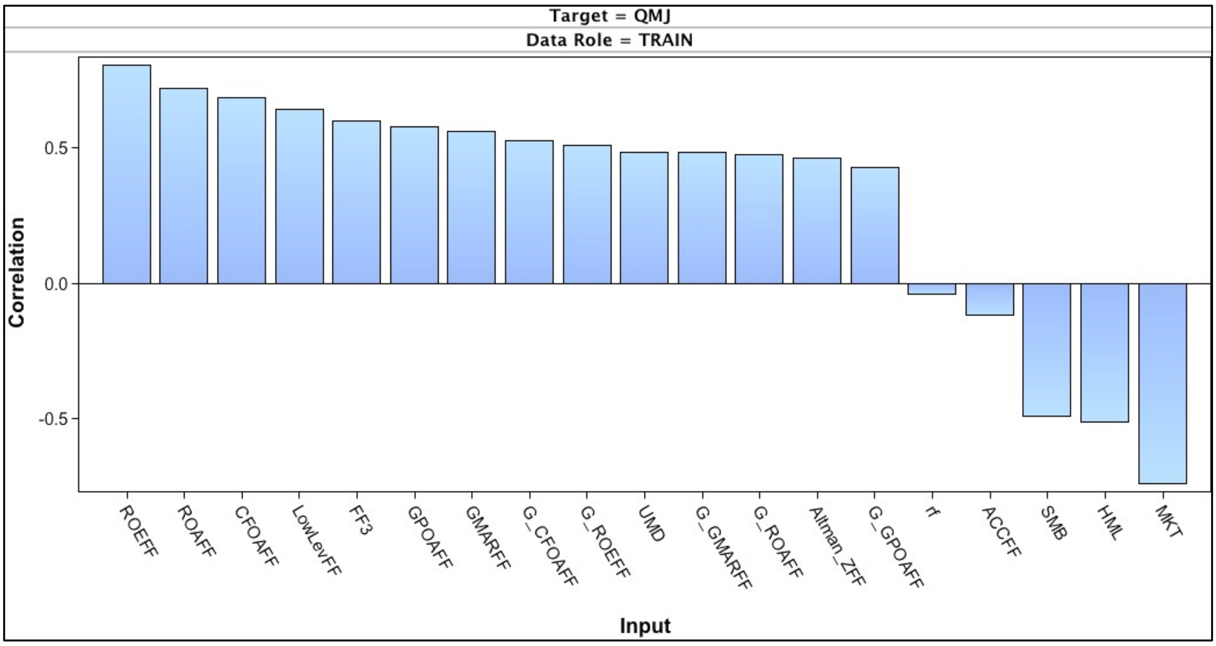

Figure 4 - Pearson's correlations

We may immediately notice that most of highly correlated ratios are profitability components ones. ROE returns seem to be highly correlated with QMJ returns and we may conclude the same for ROA and CFOA. The first non-profitability component is LowLev, which is surprising as the low leverage portfolio did not produce noticeable cumulated returns in any ways, its kurtosis value was particularly high and its mean was nothing to be concerned about. Interestingly, the FF3 (Fama/French 3 factors portfolio) seems to be somehow correlated with QMJ returns, while the components themselves of the FF3, SMB and HML are negatively correlated. This would assume that the agglomeration of the FF3 selects returns close to QMJ ones, but this does not apply to its components separately. Growth components plot on the intermediate part of the chart, not explicitly showing attractive correlation coefficients. ACC exhibits a negative correlation coefficient with QMJ returns, which is unanticipated as well. The low accruals ratio did not occur poor statistic measures or cumulated returns, however "cash is king" does not seem to relate to quality as seen by Asness et al. (2017).

The ensuing table affects the exact correlation measures computed:

Data Role=TRAIN Type=PEARSON Target=QMJ

Input	Correlation
ROEFF	0.80771
ROAFF	0.71833
CFOAFF	0.68367
LowLevFF	0.64249
FF3	0.60056
GPOAFF	0.57829
GMARFF	0.56101
G_CFOAFF	0.52867
G_ROEFF	0.50909
UMD	0.48605
G_GMARFF	0.48540
G_ROAFF	0.47325
Altman_ZFF	0.46392
G_GPOAFF	0.42995
rf	-0.03752
ACCFF	-0.11799
SMB	-0.48973
HML	-0.51216
MKT	-0.74036

Table 2 - Pearson's correlations

With the precise measures, we may acknowledge that ROE (0,81) bear a significant correlation value. More than 80% of the ROE long/short portfolio returns are positively correlated with QMJ ones. ROE and ROA (0,72) both delivered descent statistics measures and returns, close to each other and to QMJ. CFOA (0,69) is comparable to ROA in term of correlation. GPOA (0,58) which carried outstanding performances is close to 60% correlated, which is not perfect but is far from the independence. Altman_Z (0,43) which performed adjacently to QMJ is positively correlated but the degree of correlation does not ring any alarm. We may also notice that MKT (the market return adjusted by the risk-free rate, -0,74) is highly negatively correlated with QMJ. This is totally natural as all the other variables, including QMJ, are made out of spread between long and short portfolios. It is irrelevant to compare spreads, which implies risk-adjusted return by mitigating the volatility, to non-spread returns alike MKT. Precisely, the standard deviation of MKT (4,51) is one of the highest volatility. To resolve this, an analysis of long only portfolios is carried out further in this section. A table with the corresponding p values for each correlation can be found in Appendix F. All variables are considered significant except ACC.

In the next part of this analysis, I will be using multiple linear regressions with all the variables onto QMJ as a target. This procedure is computed is SASEM with the regression modelization node as seen in Appendix E.

5.4. Multiple linear regressions

The next table displays the model fit of a standard linear regression:

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	19	827.801281	43.568488	61.26	<.0001
Error	106	75.393953	0.711264		
Corrected Total	125	903.195234			

Model Fit Statistics

R-Square	0.9165	Adj R-Sq	0.9016
AIC	-24.7079	BIC	-15.2319
SBC	32.0177	C(p)	20.0000

Table 3 - Regression fit

We may assess the fit of the regression model to the target by looking at the adjusted R-squared (0,90) which implies that the regression explains more than 90% of the variations. It is generally assumed in finance that aiming 80% fit explains relatively well the movements, while a less than 70% R-squared fit would be rejected. Hence, this regression replicates the patterns highly well with 90% R-squared. We notice the p value (Pr>F) as well, extremely close to 0 (<.0001) which obviously indicates that the model is significant. The regression model obtains a quite low Akaike's criterion (AIC, -24,71), meaning that the number of selected variables seems to be fair, but would have to be compared with other form of regressions. The Bayesian's criterion (BIC, -15,23) suggests the same conclusions. The outcome is that this regression model is a pretty good fit and can be further analysed.

In the later table, we are having a peek on the variables significance with the estimates analysis:

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	0.3072	0.0989	3.11	0.0024
ACCFF	1	0.0822	0.0722	1.14	0.2577
Altman_ZFF	1	-0.7835	0.1716	-4.57	<.0001
CFOAFF	1	0.3287	0.0969	3.39	0.0010
FF3	1	0.1235	0.0889	1.39	0.1678
GMARFF	1	-0.0771	0.1043	-0.74	0.4616
GPOAFF	1	0.2375	0.1194	1.99	0.0494
G_CFOAFF	1	-0.0367	0.1141	-0.32	0.7486
G_GMARFF	1	0.1982	0.1743	1.14	0.2580
G_GPOAFF	1	-0.0197	0.1786	-0.11	0.9123
G_ROAFF	1	-0.2612	0.1547	-1.69	0.0943
G_ROEFF	1	0.0473	0.1516	0.31	0.7555
HML	1	-0.2360	0.0870	-2.71	0.0078
LowLevFF	1	0.3359	0.0763	4.40	<.0001
MKT	1	-0.0853	0.0287	-2.97	0.0037
ROAFF	1	0.6639	0.1336	4.97	<.0001
ROEFF	1	0.2852	0.1275	2.24	0.0273
SMB	1	0.0797	0.0512	1.56	0.1227
UMD	1	0.0255	0.0393	0.65	0.5178
rf	1	-0.6971	0.5151	-1.35	0.1788

Table 4 - Regression p values

The estimates parameter would give the loading of each factor in the regression equation, QMJ being the dependent variable. The accepted level of error in this model is 5%, hence I will reject variables with a p value (Pr>t) greater than 0,05. Thus, Atman_Z (<.0001), CFOA (0,0010), GPOA (0,0494), LowLev ((<.0001), ROA (<.0001) and ROE (0,0273) are the six significant variables retained with a p value lower than 0,05. These variables are the ones that matters in the model which explains 90% of the QMJ movements. MKT (0,0037) and HML (0,0078) are significant as well in this regression. This is no surprise that none of the growth ratios are represented in this model. Their statistics and returns were not compelling, neither they are significant in this regression model. At the contrary, both of the safety features (Altman_Z and LowLev) are significant and matter in the regression. Safe companies seem to be relevant in this definition of quality securities, relating to the QMJ returns. GPOA, which was the main return producer in the previous analyses, stay a major player and is significant. ROA and ROE were the most correlated profitability ratios in the Pearson correlation procedure, both of them still matter in the regression model. Thus, the importance of ROA and ROE is confirmed in the role they have among quality. We may expect quality securities to be made out of firms with good GPOA, ROA and ROE ratios. On the following plot, we may observe how good the fit is by looking at the predicted and actual means:

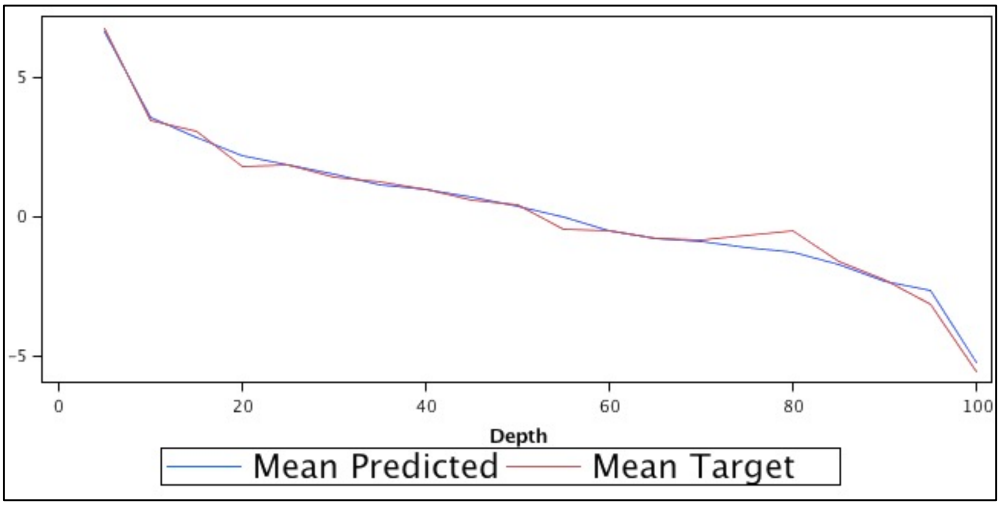

Figure 5- Predicted mean

The next table shows the results of another type of regression named multiple linear stepwise regression. The principle of this type of regression is to regress step-by-step, variables per variables, to find significant ratios that would not be considered in the first regression, or to delete some that are no longer relevant. This regression goes forward and backward. In the forward process, the regression computes one variable at a time and increase the number of variables at each step. In the backward process, the regression takes all the variables at once, then delete a non-significant variable, and re-start the process.

The selected model is the model trained in the last step (Step 7). It consists of the following effects:

Intercept Altman_ZFF CFOAFF GMARFF LowLevFF ROAFF

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	787.464954	157.492991	163.30	<.0001
Error	120	115.730280	0.964419		
Corrected Total	125	903.195234			

Model Fit Statistics

R-Square	0.8719	Adj R-Sq	0.8665
AIC	1.2875	BIC	3.4937
SBC	18.3052	C(p)	9.8656

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	0.2246	0.0897	2.50	0.0136
Altman_ZFF	1	-0.5287	0.0769	-6.87	<.0001
CFOAFF	1	0.4247	0.0567	7.49	<.0001
GMARFF	1	0.2173	0.0816	2.66	0.0088
LowLevFF	1	0.4308	0.0680	6.33	<.0001
ROAFF	1	0.9496	0.0655	14.49	<.0001

Table 5 - Stepwise

As we can see, the process ends after seven iterations. The stepwise conserves five variables; Altman_Z, CFOA, GMAR, LowLev, ROA. These ratios are all significant to the model. The model is still significant itself and as expected, the adjusted R-squared (0,87) is slightly lower than previously. This translates the fact that reducing the number of explicative variables decreases the fit of the regression. Fortunately, in this case it does not reduce the fit greatly and we still obtain a really good fit. We may also consider that this regression holds one less ratio in total, deleting ROE and GPOA as significant variables while adding GMAR. The introduction of GMAR is calling to mind, this ratio was mid-range ranked in the correlation chart (0,56) and produces abnormally negative returns (-13,77). Thus, we may or may not stick to the first regression (90% fit). However, the non-stepwise regression was taken into account variables other than the studied ratios. Indeed, the risk-free rate (rf), the market return (MKT), the size (SMB), value (HML), momentum (UMD) effects and the Fama/French 3 factor (FF3) were still accounted in the regression. The stepwise rules out these factors, which oddly implies the suppression of ROE and GPOA replaced by GMAR. Thus, this regression indicates that a model made out of the two safety ratios plus CFOA, GMAR and ROA replicates the movement of the QMJ model with almost 87% of effectiveness. We may note that in both regressions, Altman_Z seems to have a negative impact on the dependent variable QMJ, an increase in Altman's Z score would reduce the QMJ values, all else equals. Every other component has a positive estimate.

5.5. Extended analysis: Long-only cumulated returns

In the present section, I compute long-only portfolios. Returns are no longer spreads based on high quality securities minus junk ones. There is no shorting in any way, just long positions hold and rebalance every calendar month. In the following charts, we observe the returns of long only portfolios made out of the HH components: big caps high quality stocks. The aim here is to compare non-spread returns, which are more volatile but can be compare with the market return which is not a spread by definition. The long position of the QMJ factor was retrieved from AQR (2017). Here are the curves for the profitability ratios:

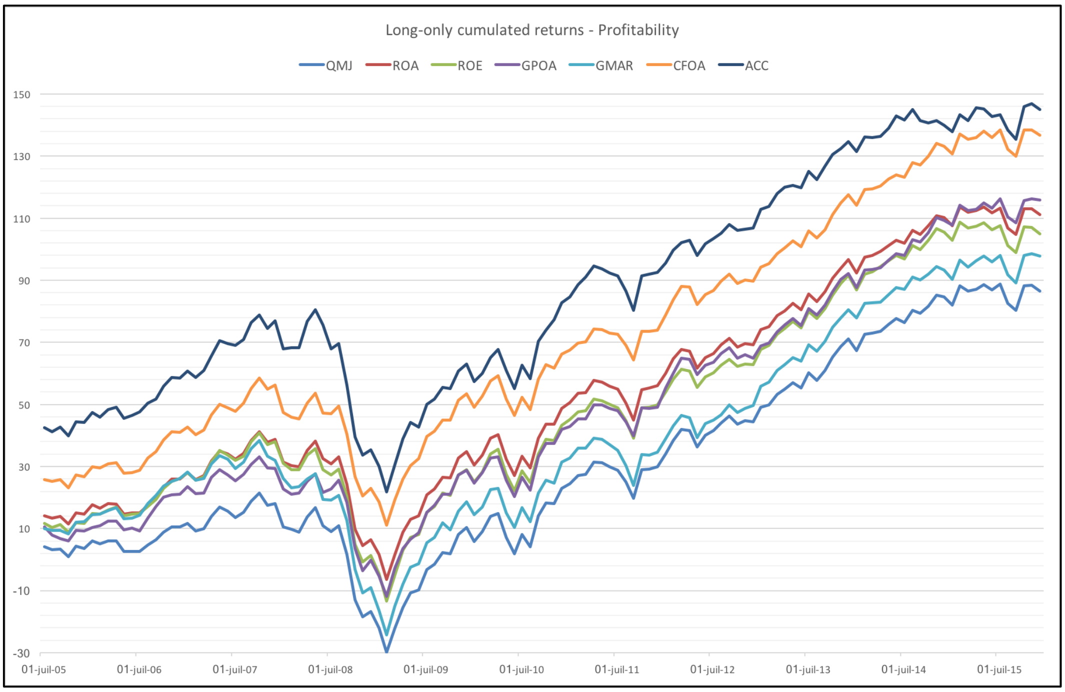

Figure 6 - Long-only cumulated returns (profitability)

We observe the absolute dominance of ACC (145) and CFOA (137). With long portfolios on quality stocks, cash seems to actually be the king. All other profitability ratios are lying between 110% to 116% of returns on 10 years. Very surprisingly, QMJ (87) is the last performer and does not seem to be made for long positions, only for spreads. The same charts with the growth components can be found in Appendix G. I computed the same chart with the safety ratios, QMJ and the market return in order to compare with what the market has done on this period:

Figure 7 - Long-only cumulated returns (safety & market)

We may notice that QMJ (87) has barely done better than the market (75) has produce. Clearly, the QMJ model is made for high minus low spread portfolios; as it produces poor excess return in its long-only form. At the contrary, both the safety ratios, Altman_Z (113) and LowLev (120) perform quite well on this 10-year period. The next sticks chart shows the all the cumulated returns by factor:

Figure 8 - Sticks chart, long-only

5.6. Extended analysis: Sharpe ratios

In this section, I compute the Sharpe ratio for each variable and portfolios' spreads. The Sharpe ratio will allow to measure the risk-adjusted returns by averaging the portfolios' returns on their standard deviation. This is useful to assess if the risk taken by the investment strategy is worth it in comparison with the return of a risk-free asset. In other words, is the excess return worth the risk or not? A negative Sharpe ratio suggests that the portfolios performed less than a risk-free asset. A positive one implies that additional risk generates excess returns. However, between 0 and 1, the excess return is too low in comparison with its risk. A Sharpe ratio equals or greater than 1 means that the excess return is not made out of too much risk taken. The following table displays the Sharpe ratios for the long/short portfolios:

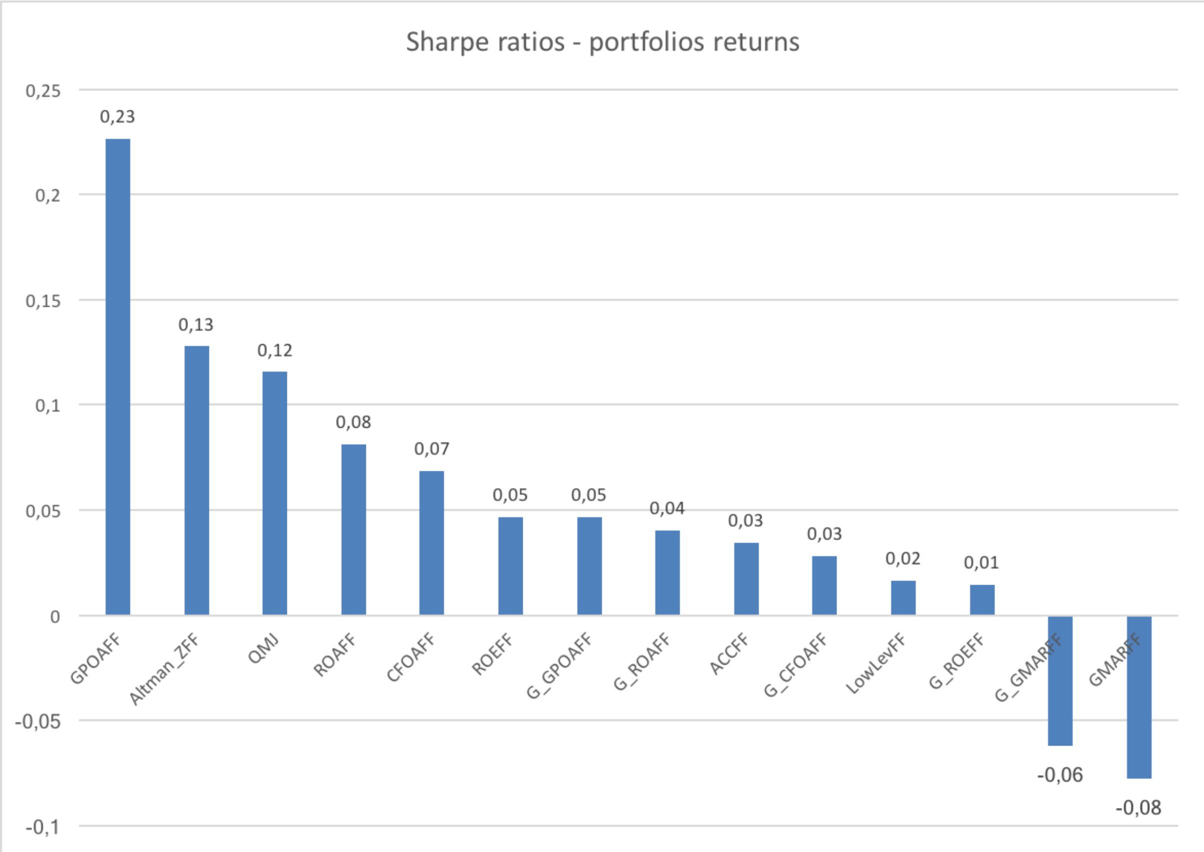

Figure 9 - Sharpe ratios

Sharpe ratio are especially low during the studied period. Even the QMJ Sharpe ratio from AQR data (from 1957 to today) is lower than 1 (approx. 0,5). We have to bear in mind that this measure has a comparative value purpose. QMJ (0,12) has a Sharpe ratio lower than one, meaning in theory that there is too much risk in comparison with the return or too low returns for a fair amount of risk. However, QMJ plots on the third place and potentially produces more risk-adjusted returns in comparison with its own components. This is the aim of Sharpe, does adding a variable to the portfolios increase the return without increasing the risk too much? Yes, this is the case here, the aggregation of variables to compute the QMJ score makes sense. However, GPOA (0,23) has the best measures once again. We cannot deny the overall performance of GPOA. Altman_Z (0,13) is comparable to QMJ and plots second. ROA (0,08) and ROE (0,05) have greater ratios than the rest of the variables: the greater the ratio, the better. Once more and well expected knowing the returns, the gross margin, both in profitability (-0,08) and growth (-0,06), are poorly performing with a negative Sharpe ratio, meaning that a risk-free asset would be a better investment strategy. Overall, every growth ratios achieve terrible effectiveness.

6. Conclusions

After having analysed the definition of quality stocks with the exhaustive QMJ model, we have a clear overview of the quality complexity. However, the quantitative analysis allows us to find the source of quality and it shows that this definition may be shrink, by a lot. QMJ in its long/short portfolio form is indeed performing well. Asness et al. (2017) model is the second performer of the analysis with almost 40% of return over the studied period of time. However, the multiple analyses performed showed that growth components have practically nothing to do with it. The five growth ratios may be entirely deleted from the quality definition and we would still obtain a pretty good fit. On the other hand, this is not the same story for safety and profitability components.

Concerning the safety ratios, the Altman's Z score (Altman_Z) and the low leverage (LowLev) ratio performed extremely well. Both are significant in the regression analysis; the credit score is close to QMJ in term of returns and the low leverage ratio is the first safety components in term of correlation. The analysis assures that safety is a major player in the definition of quality. Quality stocks are certainly made out of safe firms.

Regarding the profitability ratios, I have to emphasize the non-stop presence of the return over assets (ROA). With the second largest correlation (72%) and the third returns performance (24%), this ratio is omnipresent in the regression analyses. We may conclude the same for the return on equity (ROE), well known and used financial ratio. The return on equity is the most correlated factor and pervasive in the regressions. We absolutely cannot disregard the growth profit over assets (GPOA) role. With the highest mean (0,40), the highest performances (50%) and significant in the regression, gross profits have dominated as a return drivers in these analyses and on this period. Novy-Marx (2013) was probably right to consider profitable firms as mainly driver for excess returns. We may notice a certain cash effect. Indeed, the long-only performance of the accruals (ACC) has to be stated (145%) while the cash flow over assets (CFOA) is the fourth long/short portfolio performer and part of the regression model at each iteration. At the contrary, the gross margin (GMAR) performances are disappointing. Generating negative returns, obtaining a poor correlation measures and absent of the 90% fit regression model, this ratio may be left out of the definition.

To summarize the findings, we may state that quality stocks are mainly driven by profitable and safe firms. Growing companies do not seem to be quality ones. Here is the definition of quality securities in one table:

Profitability		Safety	
I.	Gross profit over assets (GPOA)	V.	Low leverage (LowLev)
II.	Return on equity (ROE)	VI.	Credit score (Altman's Z)
III.	Return on assets (ROA)		
IV.	Cash flow over assets (CFOA)		

Table 6 - Final results

Thus, only six ratios out of the 13 studied are detained for the definition and are driving the QMJ model. Assuming that the amount of ratio was large and comprehensive enough at the begging of the process, we may suggest a new definition of quality stocks based on Asness et al. (2017) QMJ model and portfolios: Quality consists of safe firms with low leverage, low credit risk, yet profitable with strong gross profit, return on equity, return on assets and cash flow financial statements.

7. Further discussions

In this section, the limitations and possible improvements of the carried analyses are explored. The main limitation is the time period. Indeed, I was able to extract data from a 2001-2015 sample. As explained in the methodology section, the computations of certain ratio are implying to lose three to four years of data (specifically growth ratios). The final studied period was beginning in 2005, it surely could be interesting to direct the same empirical analyses with a longer time span in order to compute more accurate answers.

The second bias could be attributed to the time context: the sample greatly suffers from the 2008 financial crisis and its stocks market road to hell. For most professionals, the late crisis was the worst since 1929 (Allayannis, 2017). The sample begins right after the internet bubble and is subject to abnormal volatility from 2007 to 2011. Once again, a longer time span sample could mitigate this extremely concentrated bearish market certified by poor Sharpe ratios in the quantitative analysis.

Equally weighted portfolios are an interesting track that could be further analysed. For the sake of consistency, I simulated portfolios with a cap-weighted technique to replicate the QMJ methodology. This technique is often used because it is assumed that big caps are safer and thus less volatile. Yet, the quality definition already suggests that only safe stocks are selected, thanks to the different safety ratios. An intriguing process would be to either to equally weight the portfolios and apply more loading on small caps, or even to "small-cap weight" the simulations. This would allow to investigate returns of smaller companies that should already be safe by definition and seek alpha somewhere else than in size.

Another track to further investigate is the construction of a quality model based on the selected ratios. Using the z-score technique that Asness et al. (2017) employ to aggregate their ratios, we could aggregate the six final ratios designated by the quantitative analysis and construct a simplified QMJ portfolios. The aim would be to compare the behaviour of the original QMJ model with the behaviour of the new definition. Would it be safer? Would it bring better Sharpe ratios and/or returns? It might produce interesting overall performances.

This thesis is based on the 2017 version of the QMJ working paper. In this late version, the authors already shrunk their definition from the previous one. In the 2013 version, they defined quality with four components instead of three nowadays. Profitability, growth and safety are unchanged, yet they were computing extra ratios from a fourth component named payout. Indeed, payout (the fraction of profits paid to shareholders in percentage, in cash or time) is often computed by practitioners to assess their investments. There are no justifications in the working paper as why they deleted this component. In the previous QMJ factor, payout was computed by equity issuance, debt issuance and net payout over profits. However, payout is often computed by the dividend yield. A new track to pursue would to be to add a payout effect, such as the dividend yield, to increase the quality selection and compare the empirical results with what has already been done.

As discussed in the conclusions section, the empirical analyses show that GPOA outruns its competitors most of the time. Gross profitability usually contributes the most to quality definition, the signal in GPOA is extremely persistent and its benefits are available to long-only investors as well as long/short investors. GPOA should be further investigated as a standalone quality factor as in big caps universe it *"largely subsumes the power of other notions of quality"* (Novy-Marx R. , 2013, p. 28). An emphasis has to be done on the safety ratios as well. Both credit risk score and low leverage ratio performed well during the analyses. It seems to me that bankers and investors have more in common than they might think. Indeed, debt is a vicious cycle, too much debt means difficulties to invest, to generate profits, to distribute dividends… A company that spends its energy to serve debts is both un-loanable to banks and not worth investing in to investors.

8. Appendices

A. SAS code, profitability and safety ratios

```
%let dir=/folders/myfolders/Master Thesis Remy/Base de données/;
libname REMY '/folders/myfolders/Master Thesis Remy/Base de données/';

Data CRSP;
Set Crsp13_remy_altmanCOPY;
ROA=IB/AT;
ROE=IB/SEQ;
GPOA=(REVT-COGS)/AT;
GMAR=(REVT-COGS)/SALE;
WC=ACT-LCT-CHE+DLC+TXP;
Run;

/**********************************************************/
proc sort data=CRSP;
by date;
run;
/*********************CFOA and ACCRUALS and LowLev**********************/

Data CRSP_WC;
Set CRSP;
date_growth=intnx('month',date,-12,'E');
format date_growth date9.;
Run;

proc sql;
create table CRSP_WC1
as select
a.*,

b.WC as WC_prior

from CRSP_WC as a
left join CRSP_WC  as b on intnx('month',a.date_growth,0,'E')=intnx('month',b.date,0,'E') and a.permno=b.permno
;
quit;

proc sort data=CRSP_WC1;
by permno date;
run;

Data CRSP_WC2;
Set CRSP_WC1;
deltaWC=WC-WC_prior;
CFOA=(IB+DP-deltaWC-CAPX)/AT;
ACC=-(deltaWC-DP)/AT;
LowLev=-(DLTT-DLC+MIBT+PSTK)/AT;
Run;

Libname out '/folders/myfolders/';
data out.CRSPratios;
Set CRSP_WC2;
Run;


/***END***/
```

Table 7 - SAS code (profitability and safety)

B. SAS code, growth ratios

```
%let dir=/folders/myfolders/Master Thesis Remy/Base de données/;
libname REMY '/folders/myfolders/Master Thesis Remy/Base de données/';

Data CRSPGROWTH;
Set remy.CRSPRATIOS;
Run;

/*******Growth ratios********/
Data CRSPgrowthrank1;
Set crspgrowth;
date_growthbis=intnx('month',date,-36,'E');
format date_growth date9.;
Run;

proc sql;
    create table CRSPgrowthrank2
    as select
    a.*,

    b.REVT as REVT_growth,
    b.COGS as COGS_growth,
    b.AT as AT_growth,
    b.IB as IB_growth,
    b.SEQ as SEQ_growth,
    b.DP as DP_growth,
    b.deltaWC as deltaWC_growth,
    b.CAPX as CAPX_growth,
    b.SALE as SALE_growth

    from CRSPgrowthrank1 as a
    left join CRSPgrowthrank1  as b on intnx('month',a.date_growthbis,0,'E')=intnx('month',b.date,0,'E') and a.perm
;
quit;

proc sort data=CRSPgrowthrank2;
by permno date;
run;

Data CRSPgrowthrank3;
Set CRSPgrowthrank2;
G_GPOA=((REVT-COGS)-(REVT_growth-COGS_growth))/AT_growth;
G_ROE=(IB-IB_growth)/SEQ_growth;
G_ROA=(IB-IB_growth)/AT_growth;
G_CFOA=((IB+DP-deltaWC-CAPX)-(IB_growth+DP_growth-deltaWC_growth-CAPX_growth))/AT_growth;
G_GMAR=((REVT-COGS)-(REVT_growth-COGS_growth))/SALE_growth;
Run;

Libname out '/folders/myfolders/';
data out.CRSPratiosandgrowth;
Set work.crspgrowthrank3;
Run;
```

Table 8 - SAS code (growth)

C. Fama/French styled six value-weighted portfolios (sample)

DATE	ROAFF	HH	HL	HM	LH	LL	LM	HH long only cumulated
31-juil-01	3,08	0,37	-2,84	-2,40	-4,34	-7,29	-2,30	0,37
31-août-01	2,39	-4,08	-6,82	-5,75	-3,18	-5,22	-2,42	-3,71
28-sept-01	0,16	-6,78	-8,46	-9,36	-15,74	-14,38	-12,91	-10,48
31-oct-01	2,71	3,47	0,32	-0,98	8,70	6,42	5,39	-7,01
30-nov-01	-0,02	7,05	7,09	6,98	8,16	8,15	5,66	0,04
31-déc-01	-1,24	-0,68	2,56	2,21	7,59	6,83	7,04	-0,65
31-janv-02	2,01	0,43	-2,62	-2,71	-0,98	-1,95	0,52	-0,21
28-févr-02	1,77	-1,95	-1,91	0,48	-1,15	-4,74	-0,78	-2,17
28-mars-02	-1,28	3,42	6,32	4,04	8,75	8,42	8,75	1,25
30-avr-02	-0,01	-6,32	-3,91	-5,71	1,80	-0,61	3,42	-5,07
31-mai-02	-0,08	-0,60	-0,48	-1,08	-4,56	-4,51	-3,27	-5,67
28-juin-02	-2,58	-8,93	-5,96	-7,30	-5,65	-3,46	-1,79	-14,60
31-juil-02	2,09	-6,71	-10,56	-6,55	-14,93	-15,27	-12,37	-21,31
30-août-02	1,55	0,94	0,77	-0,30	1,53	-1,40	0,26	-20,38
30-sept-02	4,67	-9,27	-13,48	-10,96	-5,42	-10,54	-5,52	-29,64
31-oct-02	-1,09	9,05	10,32	7,06	4,99	5,90	1,71	-20,59
29-nov-02	-8,19	3,77	11,54	5,03	7,16	15,77	5,20	-16,83
31-déc-02	4,31	-4,64	-8,69	-5,75	-4,66	-9,23	-2,57	-21,46
31-janv-03	-1,38	-3,29	-1,74	-2,65	-2,73	-1,53	-2,19	-24,75
28-févr-03	0,62	-1,12	-2,80	-1,69	-3,51	-3,06	-2,75	-25,87
31-mars-03	1,86	2,32	-1,41	0,61	1,30	1,31	1,04	-23,55
30-avr-03	-4,36	5,67	11,47	9,85	9,25	12,17	8,26	-17,88
30-mai-03	-6,45	3,82	8,84	6,74	8,75	16,62	7,65	-14,06
30-juin-03	0,42	1,63	0,40	1,27	2,62	3,00	1,97	-12,43
31-juil-03	-3,33	1,09	4,21	1,79	4,89	8,41	5,92	-11,35
29-août-03	-0,41	1,89	1,66	2,49	4,62	5,66	4,24	-9,46
30-sept-03	-1,19	-1,18	-0,84	-1,33	-2,22	-0,17	-1,72	-10,64
31-oct-03	-1,58	5,11	8,08	5,40	8,86	9,06	7,96	-5,53
28-nov-03	0,16	1,21	0,19	1,74	3,34	4,04	3,94	-4,32
31-déc-03	-1,43	3,73	5,84	5,09	1,41	2,17	3,62	-0,59
30-janv-04	-3,89	0,54	3,97	2,03	2,59	6,94	2,34	-0,04
27-févr-04	1,08	1,73	2,03	0,80	2,12	-0,33	0,98	1,69
31-mars-04	1,45	-1,47	-1,47	-1,22	2,24	-0,65	1,09	0,22
30-avr-04	4,33	0,87	-3,90	-2,85	-2,21	-6,10	-2,53	1,09
28-mai-04	-0,60	0,73	1,40	2,16	0,61	1,14	1,50	1,81
30-juin-04	0,84	1,12	1,19	3,13	4,66	2,92	6,16	2,94
30-juil-04	0,64	-4,33	-3,75	-2,15	-6,14	-8,00	-5,78	-1,39
31-août-04	-1,03	-0,13	1,55	-0,29	-1,09	-0,72	-1,13	-1,53
30-sept-04	1,34	1,14	-0,47	3,10	5,50	4,41	5,12	-0,39
29-oct-04	0,25	1,45	0,93	1,91	2,02	2,05	1,93	1,07
30-nov-04	-0,07	3,78	3,99	5,23	8,32	8,26	9,39	4,85
31-déc-04	-0,02	3,79	3,78	2,79	3,71	3,76	2,88	8,64
31-janv-05	1,93	-2,12	-3,64	-2,12	-3,02	-5,37	-2,48	6,52
28-févr-05	1,17	3,04	1,32	1,74	1,90	1,29	2,31	9,56
31-mars-05	1,73	-1,36	-3,21	-0,53	-2,31	-3,92	-1,78	8,20
29-avr-05	0,28	-2,13	-1,07	-2,06	-5,47	-7,09	-5,20	6,07
31-mai-05	0,51	3,82	3,52	3,41	6,85	6,14	6,60	9,89
30-juin-05	-1,15	-0,43	1,01	0,85	3,18	4,05	4,05	9,46
29-juil-05	0,46	4,70	2,44	3,57	5,90	7,23	6,80	14,16
31-août-05	-0,38	-0,74	-1,08	-0,61	-2,23	-1,13	-1,14	13,42
30-sept-05	0,13	0,42	0,67	1,75	0,71	0,21	1,15	13,84
31-oct-05	-1,88	-2,44	0,98	-1,95	-3,43	-3,09	-2,11	11,40
30-nov-05	0,60	3,74	3,99	3,91	5,90	4,43	4,60	15,14
30-déc-05	-0,65	-0,53	0,65	0,68	-0,17	-0,04	-0,60	14,61
31-janv-06	0,73	3,04	1,92	2,75	9,15	8,81	8,61	17,65
28-févr-06	-2,23	-1,08	1,15	1,47	-1,42	0,80	-0,53	16,57
31-mars-06	0,56	1,41	0,99	1,74	4,72	4,01	5,58	17,97
28-avr-06	-1,48	-0,13	4,00	1,29	0,85	-0,31	1,36	17,84
31-mai-06	-0,51	-3,12	-3,37	-1,97	-6,41	-5,13	-5,27	14,73
30-juin-06	0,85	0,39	-0,69	0,31	0,40	-0,21	-0,05	15,12
31-juil-06	-1,90	-0,17	1,75	-0,70	-4,87	-3,00	-3,32	14,95
31-août-06	0,08	2,61	2,02	2,75	2,40	2,83	2,81	17,56
29-sept-06	0,37	1,91	2,98	2,83	1,84	0,03	1,35	19,47
31-oct-06	0,88	3,93	2,49	2,44	5,35	5,03	5,74	23,40
30-nov-06	0,73	2,51	1,38	2,02	2,85	2,53	2,73	25,91
29-déc-06	-1,92	0,03	3,43	1,50	0,43	0,87	0,32	25,94

Table 9 - Fama/French portfolios

D. Clean dataset: Portfolios monthly returns and QMJ monthly returns (sample)

ATE	ROAFF	ROEFF	GPOAFF	GMARFF	CFOAFF	ACCFF	G_ROAFF	G_ROEFF	G_GPOAFF	G_GMARFF	G_CFOAFF	Altman_ZFF	LowLevFF	QMJ	MKT	rf	FF3	SMB	HML	UM
29-juil-05	0,46482877	-1,18194208	0,47535794	-1,16462005	-2,61208142	-0,65316487	1,40348152	0,11875502	0,60948147	-0,06383489	-0,83908443	1,22318288	0,49593405	-1,948494	4,09938877	0,255	-0,00500	1,99632072	-0,720494	0,
31-août-05	-0,38178505	-1,55181207	-1,63117816	-0,52172480	-1,31928179	-1,33955513	-0,24795457	-1,15769824	-0,12174392	-0,24652656	-0,98943566	0,16476363	0,73715775	-1,035938	-0,9056027	0,27833333	-0,29500	-1,0623373	1,28293891	2,
30-sept-05	0,12891042	-0,14054149	-1,55575030	-1,44877255	-1,16281088	1,68050365	0,39761701	0,96374508	-1,36434904	-1,14688523	0,99107468	-0,55251815	0,23000915	-1,124775	0,78908668	0,28666667	-0,47500	-0,9092861	0,97133108	3,
31-oct-05	-1,88009633	-0,12129359	-0,06561312	1,72191863	1,85897967	-0,27740442	-0,85282796	0,49947654	0,41767038	0,41534029	-0,11257407	-0,86723435	1,14767232	1,604228	-2,3520466	0,28916667	-0,55000	-1,1317523	-1,6772696	-
30-nov-05	0,60317652	0,08868983	0,85194554	0,64406149	0,83258065	-0,40313289	0,73134160	0,47199499	1,37559032	1,02033888	0,51434544	1,28034543	1,03592572	-0,1587684	3,79735194	0,32416667	1,15500	0,70112761	-1,4665172	0,
30-déc-05	-0,65406288	-1,22479236	-0,35609282	-0,80473720	-1,66322084	0,61205259	-1,23604406	-0,51324017	-1,58922277	-2,06556644	-1,73751654	-0,63530555	-2,25370817	-2,155016	0,01181207	0,32166667	-0,28500	0,078001	0,55165356	0,
31-janv-06	0,73491543	-0,86591344	-0,40252407	-2,47995182	-3,53081575	-0,90309188	0,50080536	-1,06921199	-0,78567414	-1,30380375	-0,21073149	1,86293289	0,18981632	-3,02704	3,64962906	0,3325	0,88500	4,57353835	0,95989631	3,
28-févr-06	-2,22533146	-0,61921392	-0,34842686	-0,61635767	0,08373007	-1,06673435	-1,35521661	-0,41211951	-2,10980059	-1,79343960	-1,31287811	-1,46150381	-1,37658481	-0,9153891	-0,513279	0,36416667	-0,20500	-0,0687694	-0,6411149	-
31-mars-06	0,56265223	0,11822281	1,41347863	-0,22395426	0,49353127	0,47449511	0,44917265	0,31522372	-0,23870669	-0,66952545	1,49216326	1,07921872	-0,34388554	-0,8977952	1,53845896	0,37583333	0,37500	3,1715409	0,09084848	1,
28-avr-06	-1,48255786	-0,58274465	-1,23625979	-1,02311045	-0,91368016	1,24379107	-1,27880381	0,81251922	-1,84150058	-1,81601715	-0,70629759	-1,94658264	-1,12515400	-0,8037307	0,91067995	0,37666667	-2,01000	-1,4152083	2,51219062	1,
31-mai-06	-0,51183781	0,16774327	-1,25832871	-1,06452887	0,26901464	1,07218464	-1,64222480	-1,16565720	-0,34534284	-0,48941606	-0,65631016	-2,16533974	-1,88109075	1,025185	-3,507097	0,3875	-2,78500	-2,1766037	2,31957575	-
30-juin-06	0,85109650	1,43406904	0,20168083	-0,27334892	-0,26340119	0,87472196	-0,46617813	-0,61327899	0,36033277	0,13168510	-0,86318240	0,02291219	0,23949949	1,318644	-0,4231451	0,395	-0,91500	-0,6085172	1,20816894	1,
31-juil-06	-1,89817125	0,30333326	-2,16389636	0,56115487	1,80039154	1,84922244	-3,95084413	-3,86376881	-2,23159505	-0,67884857	-0,07334002	-2,91827046	0,02515034	0,6848149	-0,641924	0,40583333	-3,32500	-3,2477081	2,48236488	-
31-août-06	0,07707451	-1,71181081	1,55056767	2,05435895	0,72017865	-0,68367885	0,32813285	-0,61475402	-0,45680708	-0,36142845	0,27471607	0,82074584	0,45853389	-0,4393762	2,09445346	0,41416667	0,50500	0,62523051	-1,4891581	-
29-sept-06	0,37243268	1,29567428	2,38509955	1,90547320	2,91378687	0,36695878	-0,03436789	0,21851371	-0,55496270	-0,78049399	2,53758892	0,34456998	-0,12593742	1,518	1,54603831	0,41	0,22000	-1,2275654	-0,1032699	-
31-oct-06	0,88192551	-0,10525393	0,57441623	-0,88600226	-0,75520764	0,50997232	1,17456776	0,95560755	0,20530675	-0,39409219	-0,66512389	0,93024475	-0,68305564	-1,622839	3,2808417	0,3975	-0,14000	1,76322705	0,21376743	-
30-nov-06	0,72635479	0,03715023	-0,92561243	-2,38114032	-0,68772781	0,67578322	1,44486047	1,05311259	1,01956920	0,25404327	-0,02204131	0,48681218	0,28930197	-0,9323372	1,94165403	0,4125	-0,97000	0,90861237	0,18739288	-
29-déc-06	-1,91677975	-0,79519458	-1,07269093	0,02960984	0,41546113	-0,46311562	-1,62880355	-1,09992792	-2,03626210	-0,99577411	0,28953693	-2,39877662	-1,39714826	-0,4353286	0,63907594	0,40833333	-1,65000	-0,7057608	2,56136978	0,
31-janv-07	0,74468487	-0,80110317	2,09052788	0,45108581	-0,16427103	0,84144058	0,49860699	0,39580612	0,32607581	-1,02738567	0,16429916	1,49806723	-0,68573641	-1,315349	1,52188068	0,4075	0,42500	-0,02225	0,25695846	1,
28-févr-07	0,26957254	-0,46638731	-0,05457578	-2,10031321	-1,52486468	-1,39404212	-0,10769792	-0,31995853	-0,79834341	-1,53977829	-0,85481219	0,38131716	-0,22445642	-0,9062057	-1,8448891	0,41583333	-0,69000	1,37428739	0,78239337	-
30-mars-07	1,16461773	0,20704751	0,29044141	-1,64547812	-0,70551739	0,45479368	0,07931789	-0,24007981	0,53828705	-0,26723633	0,45593125	0,84222969	0,76657797	-0,1370475	0,86201505	0,4175	0,63500	-0,5175337	0,2981628	2,
30-avr-07	1,06827517	0,38899036	1,23359506	-0,54732670	0,76187039	0,28157790	0,85011845	0,39161921	0,25359476	-0,07709396	1,16381970	1,79977640	-1,24889188	-0,2437587	3,65679393	0,40833333	0,70500	-2,0385385	-0,6962715	0,
31-mai-07	1,15268449	0,71569663	0,90310631	-1,15651501	0,35377956	0,85517446	0,93701126	0,86979848	1,82883801	0,70604208	0,74234676	0,45302519	-0,00089055	0,6429881	3,54134709	0,39916667	0,04000	-0,46975	0,45896315	-
29-juin-07	1,13916275	0,17677917	1,18840488	0,27407770	1,04874165	0,41756107	1,45082805	1,25717175	1,49496480	0,88451281	0,23977393	1,54741139	1,81602621	1,107998	-1,9064837	0,38333333	1,53000	0,76230727	-0,5895513	0,
31-juil-07	3,38147603	1,36080338	2,70192371	0,26384888	0,55338018	-0,27980901	3,06206813	1,59829112	2,82168865	1,64464875	0,96551146	3,85615322	0,84408415	1,808609	-3,5835904	0,39	4,38500	-2,7851879	-3,183859	2,
31-août-07	-0,48433191	-0,27827348	-0,88242660	1,39776315	0,95180743	0,21931618	-0,64241598	-1,05458077	-1,15564069	0,29313319	0,45975835	-0,29340569	1,85452090	1,59718	0,74471188	0,40166667	1,58500	-0,473097	-2,6691947	-
28-sept-07	0,32393834	0,40284402	-0,68399170	1,32060387	-2,19368713	1,17732734	2,45243891	2,39748882	3,21290876	3,51451066	0,19813604	1,03042882	-0,14883439	-1,265493	3,73025947	0,32583333	3,29000	-2,4849842	-1,1497685	3,
31-oct-07	2,60076735	-0,09846752	2,04809817	-0,00539105	-0,93578369	1,31936618	1,52034293	0,58826045	1,11530916	0,16803279	-1,22124951	3,30141660	0,82378895	-0,5682944	2,27374968	0,31	4,70500	-0,0837305	-2,4678149	6,
30-nov-07	3,01576243	2,73240345	1,11307461	0,18105807	1,06169268	-1,10491078	0,83519046	-0,01091609	-0,01365422	-0,87043944	-0,04391210	1,12099348	1,96679159	4,181923	-5,3098805	0,32	1,77000	-2,9701768	-1,5752036	0,
31-déc-07	2,32308966	0,80426075	1,18717676	-1,68950479	-0,80637975	1,92333948	3,69801184	2,21038712	2,13595780	1,00809579	-0,91638620	1,93465151	0,76622251	1,078446	-0,7156856	0,25666667	3,12500	-0,124028	-0,2498565	5,
31-janv-08	-3,45641807	0,46716355	-2,60603173	-0,80485768	-0,56050392	0,39725097	-2,51125140	-0,66819408	-1,84252908	-0,47403496	0,02142666	-4,81345162	-1,05068509	-0,1917929	-6,5252546	0,27416667	-8,18000	-0,3724049	4,02113588	-
29-févr-08	5,24832580	0,86504924	3,03614384	-1,26316791	-3,80529180	-0,47405867	4,85069285	3,02048820	4,60542183	1,07440930	0,54578651	3,53088049	-0,48811722	2,134694	-2,4634335	0,16	4,26500	-0,1319475	-1,0486404	7,
31-mars-08	0,54961868	-0,32369169	0,09339457	2,02601933	-0,26271846	-1,33007557	-0,38223127	-0,74105328	0,60540982	2,08507129	-2,04653120	-0,63175713	-0,28544975	2,69211	-1,2316288	0,15083333	1,11500	0,14387576	-0,6615151	2,
30-avr-08	0,82581347	0,64570373	0,49292577	-2,09722126	-0,53922672	3,74106712	2,20290283	2,83467544	2,17372818	0,99134784	1,65134879	1,37467121	1,54768133	-0,9213436	5,05094059	0,11333333	0,36000	-2,030639	-0,4635096	0,
30-mai-08	3,77128185	-0,17625960	2,47544337	-1,57339548	-0,46050959	0,75369864	2,71783978	1,14780803	1,94063934	0,55999434	-0,96756856	2,59695622	0,11927992	1,011436	2,2181654	0,1175	4,25000	2,80373848	-1,8080224	3,
30-juin-08	5,94515794	2,63647141	2,44783246	-0,90337315	-4,08774868	2,65241641	4,58521084	2,40582992	4,27680223	1,19044999	-1,65837811	3,19908606	-1,07088834	3,801915	-8,0628816	0,15416667	8,75000	-0,149464	-2,826964	1
31-juil-08	-4,65707494	-0,20342582	0,11663984	1,60254067	6,62244476	-4,30817568	-3,59655386	-4,07173079	-3,69460173	-2,94540942	1,74311301	-3,56153623	-0,45003719	0,06695195	-1,3624829	0,15583333	-6,18500	2,5148385	1,82464767	
29-août-08	2,08729727	1,86483755	3,33495322	1,52690649	1,78675442	-0,30738180	0,26314686	-0,24877982	0,74104603	0,21313552	1,18847564	0,48731454	-0,13765292	2,652394	1,07432562	0,1375	-2,27500	2,13770059	1,88239367	-
30-sept-08	-2,39919842	2,58624313	-0,06374690	3,20087122	3,31920159	-3,21523050	-0,95760159	-1,47144367	-4,59234497	-1,91829850	0,40502586	-3,81001747	2,81833840	4,169567	-9,8836882	0,14083333	-3,25500	0,1412928	3,23442911	0,
31-oct-08	4,87221260	5,54434394	2,17488873	4,39435033	6,44618975	-0,79329135	1,08188091	1,36186293	-0,54242261	-0,58336545	2,03877512	1,65553962	6,22320106	9,034384	-18,527959	0,075	7,52500	-3,4981441	-2,0461682	7,
28-nov-08	5,07958145	3,54111610	-0,29958876	-0,75818043	2,35426260	0,21850104	2,43838558	2,71176294	0,47735975	0,58908296	2,55216067	2,56027507	5,39846562	8,580549	-8,4654042	0,03666667	5,73000	-5,3908305	-2,3374809	6,
31-déc-08	2,01089019	0,84272212	2,47017158	-1,50687856	0,46918851	-0,25840104	0,03033065	0,36061400	0,67677414	-0,14945805	1,03578916	2,46744508	0,37040658	-0,1861015	2,14332528	0,00083333	0,05500	3,34960616	-1,0517524	-

'able 10 - Clean dataset

E. SAS Enterprise Miner dataset, diagram and nodes

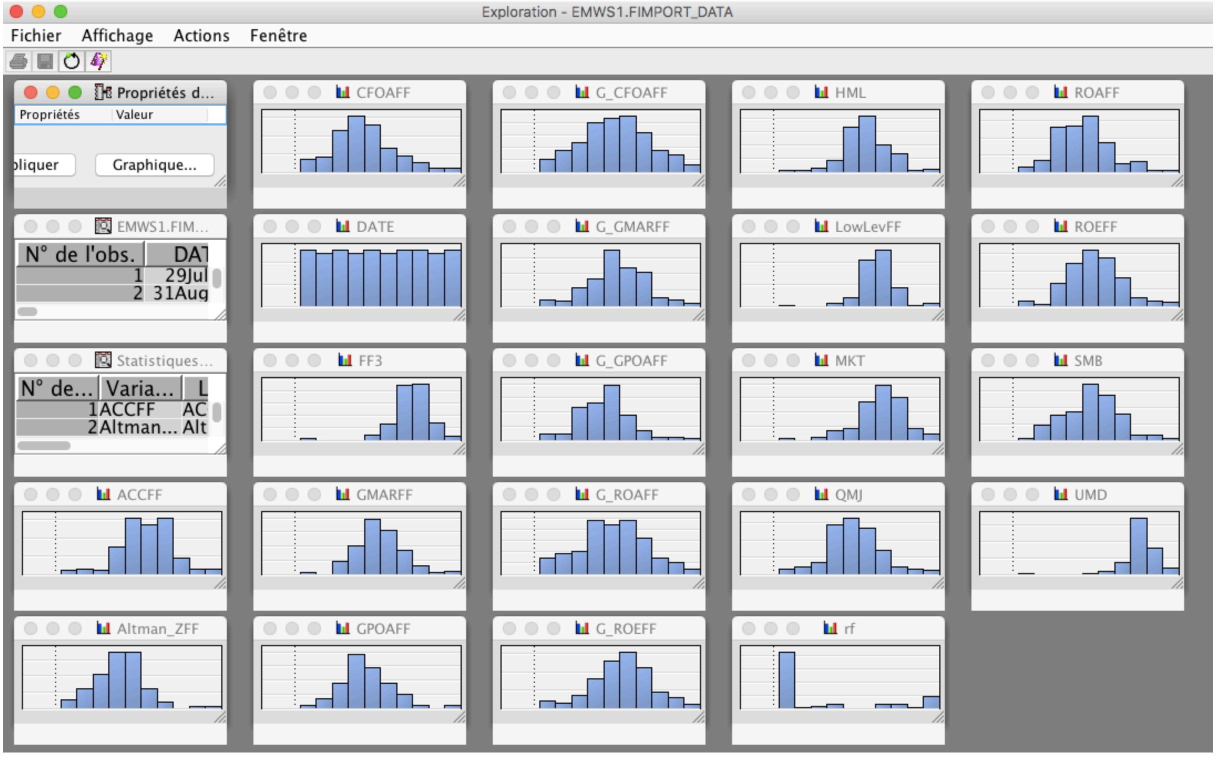

Figure 10 - SASEM variables distributions

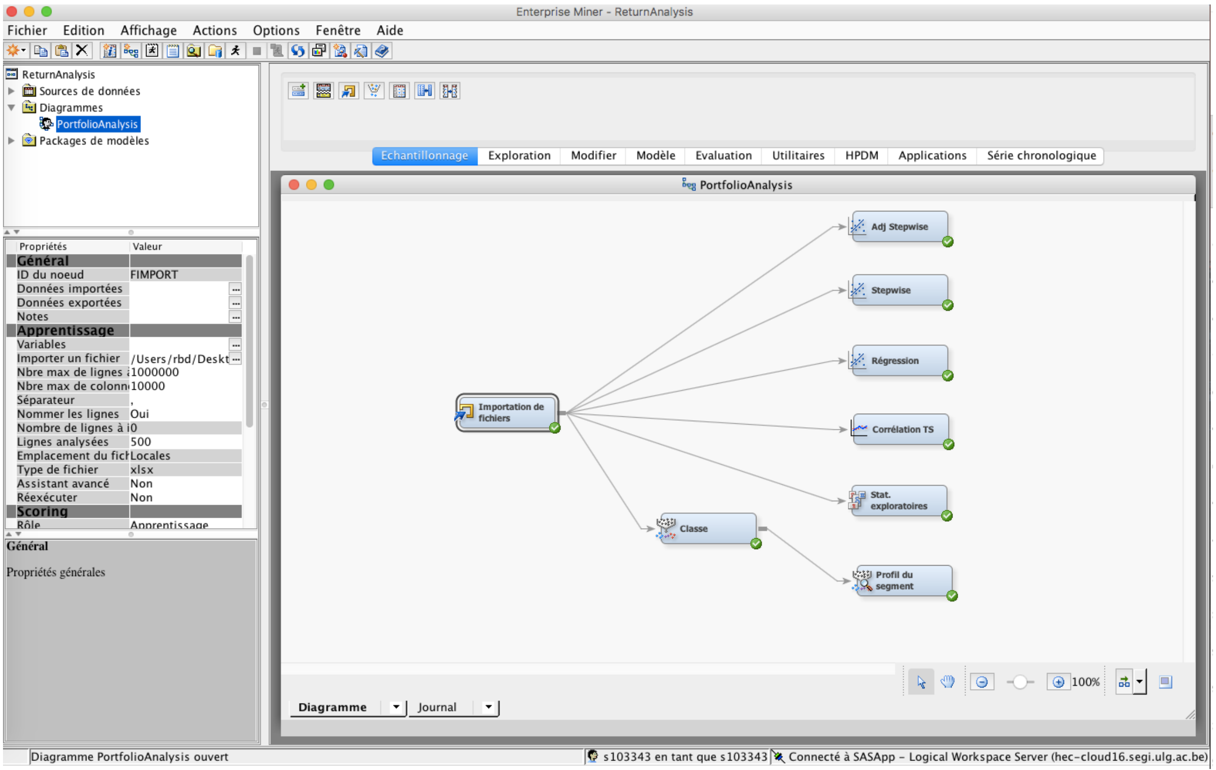

Figure 11 - SASEM diagram and nodes

F. P values correlations table

La procédure CORR

14 Variables :	ROAFF ROEFF GPOAFF GMARFF CFOAFF ACCFF G_ROAFF G_ROEFF G_GPOAFF G_GMARFF G_CFOAFF Altman_ZFF LowLevFF QMJ

Statistiques simples

Variable	N	Moyenne	Ec-type	Somme	Minimum	Maximum	Libellé
ROAFF	126	0.19001	2.33374	23.94157	-5.37639	8.45984	ROAFF
ROEFF	126	0.08238	1.75872	10.37953	-4.84625	5.54434	ROEFF
GPOAFF	126	0.40058	1.76923	50.47289	-4.15509	6.85088	GPOAFF
GMARFF	126	-0.10932	1.40839	-13.77429	-4.89243	4.75033	GMARFF
CFOAFF	126	0.13872	2.02737	17.47849	-4.08775	6.62244	CFOAFF
ACCFF	126	0.04583	1.33574	5.77444	-4.30818	3.75610	ACCFF
G_ROAFF	126	0.07420	1.83236	9.34896	-3.95084	4.85069	G_ROAFF
G_ROEFF	126	0.02449	1.70268	3.08581	-4.64784	4.41952	G_ROEFF
G_GPOAFF	126	0.08451	1.81026	10.64859	-4.59234	6.70347	G_GPOAFF
G_GMARFF	126	-0.09190	1.47566	-11.57960	-4.10933	4.15601	G_GMARFF
G_CFOAFF	126	0.03627	1.28489	4.57007	-2.78495	3.14408	G_CFOAFF
Altman_ZFF	126	0.26967	2.11006	33.97861	-4.82606	9.07500	Altman_ZFF
LowLevFF	126	0.03034	1.83839	3.82330	-9.78816	6.22320	LowLevFF
QMJ	126	0.31069	2.68804	39.14666	-7.26580	9.03438	QMJ

Coefficients de corrélation de Pearson, N = 126
Proba > |r| sous H0: Rho=0

	ROAFF	ROEFF	GPOAFF	GMARFF	CFOAFF	ACCFF	G_ROAFF	G_ROEFF	G_GPOAFF	G_GMARFF	G_CFOAFF	Altman_ZFF	LowLevFF	QMJ
ROAFF ROAFF	1.00000	0.75571 <.0001	0.75007 <.0001	0.23510 0.0081	0.25756 0.0036	0.05017 0.5769	0.74873 <.0001	0.71714 <.0001	0.72554 <.0001	0.56863 <.0001	0.37411 <.0001	0.80262 <.0001	0.39954 <.0001	0.71833 <.0001
ROEFF ROEFF	0.75571 <.0001	1.00000	0.53563 <.0001	0.36989 <.0001	0.60053 <.0001	-0.08313 0.3547	0.59950 <.0001	0.66952 <.0001	0.45989 <.0001	0.49176 <.0001	0.54973 <.0001	0.41592 <.0001	0.37204 <.0001	0.80771 <.0001
GPOAFF GPOAFF	0.75007 <.0001	0.53563 <.0001	1.00000	0.40993 <.0001	0.40506 <.0001	-0.06416 0.4754	0.56155 <.0001	0.51683 <.0001	0.64768 <.0001	0.49734 <.0001	0.47083 <.0001	0.80848 <.0001	0.38991 <.0001	0.57829 <.0001
GMARFF GMARFF	0.23510 0.0081	0.36989 <.0001	0.40993 <.0001	1.00000	0.56493 <.0001	-0.17283 0.0530	0.15466 0.0838	0.17279 0.0530	0.21795 0.0142	0.44178 <.0001	0.36382 <.0001	0.21718 0.0146	0.55151 <.0001	0.56101 <.0001
CFOAFF CFOAFF	0.25756 0.0036	0.60053 <.0001	0.40506 <.0001	0.56493 <.0001	1.00000	-0.35451 <.0001	0.19675 0.0272	0.25425 0.0041	0.10870 0.2257	0.26195 0.0030	0.71522 <.0001	0.14686 0.1008	0.50105 <.0001	0.68367 <.0001
ACCFF ACCFF	0.05017 0.5769	-0.08313 0.3547	-0.06416 0.4754	-0.17283 0.0530	-0.35451 <.0001	1.00000	-0.08886 0.3225	-0.03342 0.7103	0.02987 0.7399	-0.04663 0.6041	-0.17325 0.0524	-0.02501 0.7810	-0.18489 0.0382	-0.11799 0.1882
G_ROAFF G_ROAFF	0.74873 <.0001	0.59950 <.0001	0.56155 <.0001	0.15466 0.0838	0.19675 0.0272	-0.08886 0.3225	1.00000	0.93158 <.0001	0.85592 <.0001	0.76260 <.0001	0.50519 <.0001	0.65226 <.0001	0.28787 0.0011	0.47325 <.0001
G_ROEFF G_ROEFF	0.71714 <.0001	0.66952 <.0001	0.51683 <.0001	0.17279 0.0530	0.25425 0.0041	-0.03342 0.7103	0.93158 <.0001	1.00000	0.80556 <.0001	0.76529 <.0001	0.54000 <.0001	0.60777 <.0001	0.29174 0.0009	0.50909 <.0001
G_GPOAFF G_GPOAFF	0.72554 <.0001	0.45989 <.0001	0.64768 <.0001	0.21795 0.0142	0.10870 0.2257	0.02987 0.7399	0.85592 <.0001	0.80556 <.0001	1.00000	0.88031 <.0001	0.39967 <.0001	0.73920 <.0001	0.27138 0.0021	0.42995 <.0001
G_GMARFF G_GMARFF	0.56863 <.0001	0.49176 <.0001	0.49734 <.0001	0.44178 <.0001	0.26195 0.0030	-0.04663 0.6041	0.76260 <.0001	0.76529 <.0001	0.88031 <.0001	1.00000	0.49040 <.0001	0.56703 <.0001	0.37512 <.0001	0.48540 <.0001
G_CFOAFF G_CFOAFF	0.37411 <.0001	0.54973 <.0001	0.47083 <.0001	0.36382 <.0001	0.71522 <.0001	-0.17325 0.0524	0.50519 <.0001	0.54000 <.0001	0.39967 <.0001	0.49040 <.0001	1.00000	0.33098 0.0002	0.41307 <.0001	0.52867 <.0001
Altman_ZFF Altman_ZFF	0.80262 <.0001	0.41592 <.0001	0.80848 <.0001	0.21718 0.0146	0.14686 0.1008	-0.02501 0.7810	0.65226 <.0001	0.60777 <.0001	0.73920 <.0001	0.56703 <.0001	0.33098 0.0002	1.00000	0.49359 <.0001	0.46392 <.0001
LowLevFF LowLevFF	0.39954 <.0001	0.37204 <.0001	0.38991 <.0001	0.55151 <.0001	0.50105 <.0001	-0.18489 0.0382	0.28787 0.0011	0.29174 0.0009	0.27138 0.0021	0.37512 <.0001	0.41307 <.0001	0.49359 <.0001	1.00000	0.64249 <.0001
QMJ QMJ	0.71833 <.0001	0.80771 <.0001	0.57829 <.0001	0.56101 <.0001	0.68367 <.0001	-0.11799 0.1882	0.47325 <.0001	0.50909 <.0001	0.42995 <.0001	0.48540 <.0001	0.52867 <.0001	0.46392 <.0001	0.64249 <.0001	1.00000

Table 11 - Correlations p values

G. Long-only portfolios, growth ratios

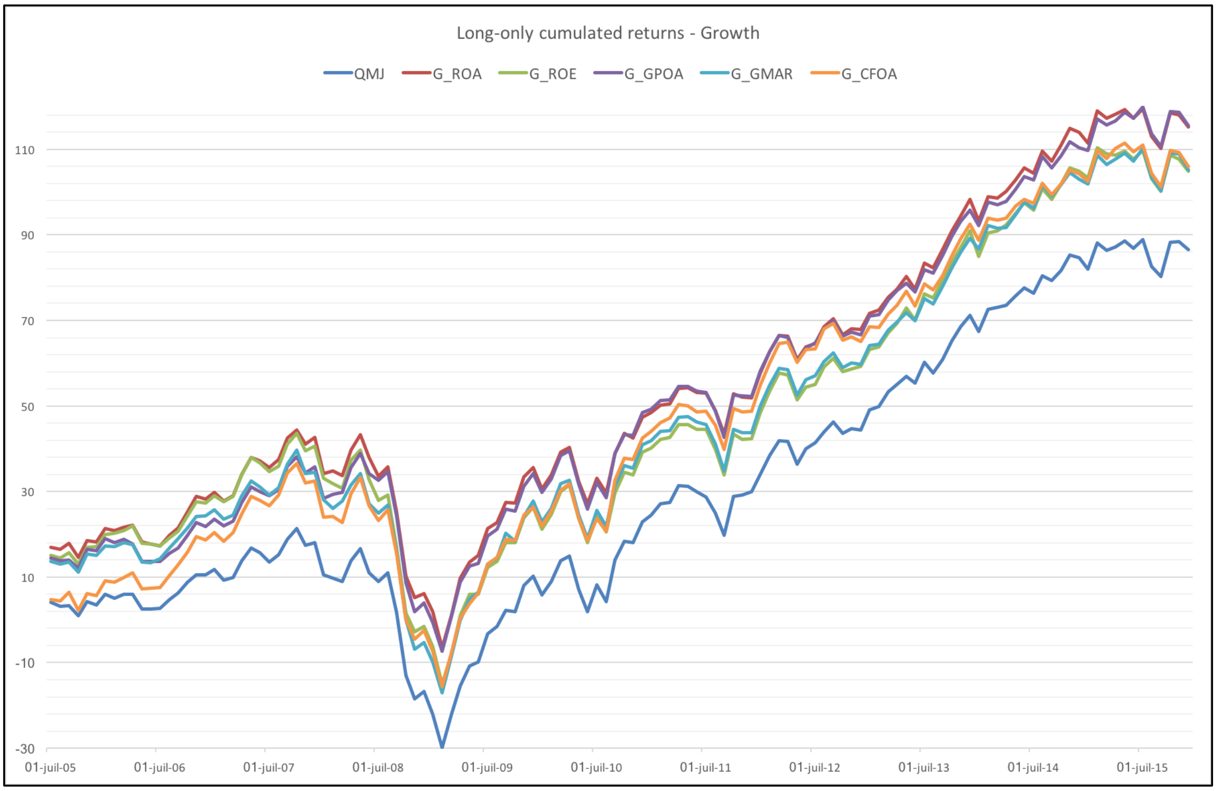

Figure 12 - Long-only cumulated returns (growth)

9. Bibliography

Allayannis, G. (2017). The Financial Crisis of 2007–2009: The Road to Systemic Risk. *Darden Business Publishing Cases*, 1-16.

Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance, 23(4)*, pp. 589-609.

AQR. (2017, July 15). *AQR - Quality Minus Junk: Factors, Monthly.* Retrieved from AQR: https://www.aqr.com/library/data-sets/quality-minus-junk-factors-monthly

Asness, C., Frazzini, A., & Pedersen, L. (2017). Quality minus junk. *Fama-Miller Working Paper*.

Asness, C., Frazzini, A., Israel, R., Moskowitz, T., & Perdersen, L. (2015). Size matters, if you conrol your junk. *Fama-Miller Working Paper*.

Banz, R. (1981). "The Relationship Between Return And Market Value Of Common Stocks". *Journal Of Financial Economics, 9*, pp. 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.

Burton, F., & Shah, S. (2017). *Efficient Market Hypothesis.* CMT Level I 2017: An Introduction to Technical Analysis.

Campbell, J., Hilscher, J., & Szilagyi, J. (2008). In search of distress risk. *Journal of Finance, 63*, pp. 2899–2939.

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

Chamberlain, G. (1983). Funds, factors, and diversification in arbitrage pricing models. *Econometrica: Journal of the Econometric Society*, 1305-1323.

Chamberlain, G., & Rothschild, M. (1983). Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets. *Econometrica, Vol. 51*(No. 5), 1281-1304.

Chen, N. F., & Ingersoll, J. E. (1983). Exact pricing in linear factor models with finitely many assets: A note. *The Journal of Finance, 38(3)*, 985-988.

Connor, G. (1984). A unified beta pricing theory. *Journal of Economic Theory, 34(1)*, 13-31.

Dichtl, H., & Drobetz, W. (2014). The case of the "Halloween Indicator", Are stock markets really so inefficient? *Finance Research Letters, 11(2)*, pp. 112-121.

Durlauf, S., & Blume, L. (2008). *The New Palgrave Dictionary of Economics* (Vol. 6). Basingstoke: Palgrave Macmillan.

Fama, E. F. (1965). The behavior of stock-market prices. *The journal of business, 38(1)*, 34-105.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Fama-Miller Working Paper, 2014. Acesso em, 17*.

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

Fama, E., & French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics, 33(1)*, pp. 3-56.

Fama, E., & French, K. (2012). Size, value, and momentum in international stock returns. *Journal of financial economics, 105(3)*, pp. 457-472.

George, T., & Hwang, C. (2010). A Resolution of the Distress Risk and Leverage Puzzles in the Cross Section of Stock Returns. *Journal of Financial Economics, 96*, pp. 56-79.

GMO. (2004). *The case for quality - The danger of junk.* Boston: GMO.

Graham, B., & Dodd, D. (1940). *Security Analysis: Principles and Technique.* New York: McGraw-Hill Book Company, Inc.

Greenblatt, J. (2010). *The little book that still beats the market.* New Jersey: John Wiley & Sons.
Griffin, J. (2002). Are the Fama and French Factors Global or Country Specific? *Rev Financ Stud, 15 (3*, 783-803.
Hackel, K., Livnat, J., & Rai, A. (1994). The free cash flow/small-cap anomaly. *Financial Analysts Journal*, pp. 33-42.
Hsu, J., Kalesnik, V., & Li, F. (2012). An Investor’s Guide to Smart Beta Strategies. *AAII Journal*, 11-16.
Huberman, G. (1982). A simple approach to arbitrage pricing theory. *Journal of Economic Theory, 28(1)*, 183-191.
Joyce, C., & Mayer, K. (2012). *Profits for the long run: Affirming the case for quality.* GMO.
Jung, J., & Shiller, R. (2005). Samuelson's dictum and the stock market. *Economic Inquiry, 43(2)*, 221-228.
Leland, H. E. (1985). Option pricing and replication with transactions costs. *The journal of finance, 40(5)*, pp. 1283-1301.
Lintner, J. (1965, December 20). Security Prices, Risk and Maximal Gains from Diversification. *Journal of Finance, 4*, pp. 587–615.
Matthews, R. B. (2014, March 3). *'Deep value' VS. 'High quality': Berkshire numbers rekindle debate*. Retrieved from The globe and mail: https://www.theglobeandmail.com/globe-investor/investment-ideas/berkshires-latest-numbers-rekindle-investment-strategy-debate/article17240125/
Mohanram, P. (2005). Separating Winners from Losers among Low Book-to-Market Stocks using Financial Statement Analysis. *Review of Accounting Studies, 10*, pp. 133–170.
Mossin, J. (1966, October). Equilibrium in a Capital Asset Market. *Econometrica, 34, no. 4*, pp. 768–783.
MSCI. (2013, May). Retrieved from MSCI: https://www.msci.com/eqb/methodology/meth_docs/MSCI_Quality_Indices_Methodology.pdf
Nocera, J. (2009). Poking holes in a theory on markets. *New York Times*, p. 5.
Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics, 108(1)*, pp. 1-28.
Novy-Marx, R. (2013). The quality dimension of value investing. *Rnm. simon. rochester. edu*, 1-54.
Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of accounting research, 18(1)*, pp. 109-131.
Olson, D., Mossman, C., & Chou, N. T. (2015). The evolution of the weekend effect in US markets. *The Quarterly Review of Economics and Finance, 58*, pp. 56-63.
Palacios, L., & Vora, P. (2009, April). *fama_french_factor_replication.sas.* Retrieved from The Wharton School: https://wrds-web.wharton.upenn.edu/wrds/research/applications/sas_files/fama_french_factors_replication.sas
Patel, N., & Sewell, M. (2015). Calendar anomalies: a survey of the literature. *International Journal of Behavioural Accounting and Finance, 5(2)*, pp. 99-121.
Penman, S., Richardson, S., & Tuna, I. (2007). The Book-to-Price Effect in Stock Returns: Accounting for Leverage,. *Journal of Accounting Research, 45(2)*, pp. 427-467.
Reilly, F., & Brown, K. (2011). *Investment analysis and portfolio management.* Cengage Learning.
Richardson, S. A., Sloan, R. G., Soliman, M. T., & Tuna, I. (2005). Accrual reliability, earnings persistence and stock prices. *Journal of accounting and economics, 39(3)*, pp. 437-485.

Ross, S. (1976). The arbitrage theory of capital asset pricing. *Journal of economic theory, 13(3)*, 341-360.
Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance, 19(3)*, pp. 425-442.
Sharpe, W. F. (1964, September 19). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance, 3*, pp. 425–442.
Shleifer, A., & Vishny, R. W. (1997). The limits of arbitrage. *The Journal of Finance, 52(1)*, pp. 35-55.
Tversky, A., & Kahneman, D. (1986). Rational Choice and the Framing of the Decision. *The Journal of Business, Vol.59*.
Vayanos, D. (2004). *Flight to quality, flight to liquidity, and the pricing of risk.* Cambridge: National Bureau of Economic Research.
Vayanos, D., & Gromb, D. (2010, March). Limits of Arbitrage: The State of the Theory. *No. dp650*. Cambridge, MA 02138: Financial Markets Group.
Volcker, P. (2012). Unfinished business in financial reform. *International Finance, 15(1)*, pp. 125-135.
Wei, K. (1988). An Asset-Pricing Theory Unifying the CAPM and APT. *The Journal of Finance, 43(4)*, 881-892.
Yalcin, K. C. (2010). Market Rationality: Efficient Market Hypothesis versus Market Anomalies. *European Journal of Economic and Political Studies*, pp. 23-38.

10. Table of content (long version)

11. Table of figures

12. Table of tables

13. Lexicon

ACC	Low accruals ratio
Altman_Z	Low credit risk ratio
APT	Arbitrage pricing theory
Caps	Refers to the market capitalization of a firm
CAPM	Capital asset pricing model
CFOA	Cash flow over assets ratio
CMA	Conservative minus aggressive
CRSP	The center for research in security prices database
FF3	Refers to the Fama/French three factor model
FF4	Idem to the four factor
G_	Refers to a ratio, in growth form
GMAR	Gross margin ratio
GPOA	Gross profit over assets ratio
HML	High minus low portfolios
LowLev	Low leverage ratio
MOM	Refers to UMD portfolios
NYSE	New York stock exchange
QMJ	Quality minus junk portfolios
RFR	Risk free rate
RMW	Robust minus weak portfolios
ROA	Return on assets
ROE	Return on equity
SAS	Statistical analysis software
SASEM	Data miner version of SAS
SMB	Small minus big portfolios
UMD	Up minus down portfolios

14. Executive summary

This thesis aims at investigating the market anomaly quality as defined by Asness, Frazzini and Pedersen (2017) in their “Quality Minus Junk” factor. The undertake study refines the quality stocks’ definition and its complexity. The concept of the quality anomaly has been for years arduous to portray, as its meaning is highly subjective and differs from one academician to another. Quality is occasionally not seen as a “pure anomaly” since it consists of an aggregation of numerous factors and ratios. This memoir is willing to enlighten this interpretation puzzle.

The basic concepts of market theories and portfolio management are introduced and discussed, just like the evolution of pricing models. The most distinguished anomalies others than quality are acquainted as a preface for the quality concept debate. Hence, the QMJ factor is analyzed in its three components; profitability, growth and safety. A replica of its ratios is built using SAS software with the goal to simulate Fama/French styled long-short portfolios based on a CRSP/Compustat dataset. The computed portfolios are regressed on QMJ and analyzed using SAS Miner software, along with descriptive statistics, correlations, cumulated returns and Sharpe ratios.

The results show that the growth component may be entirely dismissed without damaging the model. The safety factors greatly matter in the regressions and strengthen their role into quality. Return on equity, return on assets and cash flows are profitability ratios which are significant in the definition as well. While the signals of gross profits are remarkably persistent and drove the quality performance in all empirical analyses. Hence, the source of quality is identified and corresponds to six final ratios, cutting the complexity of the definition by more than two.

Keywords: quality, factor investing, portfolio simulation, QMJ, gross profit, market anomaly.