Quality investing : design of a quality definition from a practitioners' perspective

Auteur : Daco, Julien
Promoteur(s) : Lambert, Marie
Faculté : HEC-Ecole de gestion de l'Université de Liège
Diplôme : Master en ingénieur de gestion, à finalité spécialisée en Financial Engineering
Année académique : 2017-2018
URI/URL : http://hdl.handle.net/2268.2/4386

Avertissement à l'attention des usagers :

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

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

Jury :
Promoter :
Marie LAMBERT
Readers :
Boris FAYS
Yves FRANCIS

Dissertation by
Julien DACO
For a Master’s Degree in Business
Engineering, with a professional focus
on Financial Engineering
Academic year 2017/2018
Acknowledgements

It is with great pleasure and not without some emotion that I submit my final work at HEC Liège. As I look back on the achievements of the past 6 years, I can only thank the many people who supported me.

I would like to express my deep sense of gratitude to Marie LAMBERT, Associate Professor (Deloitte Chair of Financial Management and Corporate Valuation) at HEC Liège and Research Associate at EDHEC Risk Institute, for her guidance in the framework of this thesis as well as her invaluable support as part of the CFA Research Challenge.

I would like to extend my gratitude to Georges HÜBNER, Full Professor of Finance at HEC Liège, for having given me the opportunity to use a subsample of CRSP/Compustat in the framework of this thesis; to Boris FAYS, Doctoral Student at HEC Liège, for his precious pieces of advice; and to Yves FRANCIS, Managing Partner at Deloitte Luxembourg, for his availability and to have done me the honour of being my reader.

I would also like to once again thank Frederic ANCION, Head of Credit Risk Corporate France/Belgium at Commerzbank AG Luxembourg, for offering me the chance to integrate the Credit Risk Department as part of my internship. This experience undoubtedly had a great influence on my projects.

Last but not least, I am also grateful to all the people who support me in day-to-day life and made, and will continue to make, life so enjoyable. I first thank my relatives for the trust they have bestowed on me, and my great little sister whose turn it is to enter university. Thanks to my old brothers, Johan DZEPRAILIDIS and Xavier JUNGERS, for the rereading of my thesis, but more particularly for being such precious friends; our journey has just begun and I look forward to discover what will happen next. Thanks to Rémy-Baptiste DELHEZ, my friend and colleague, for his unfailing support in the last months.
# Table of contents

---

Acknowledgements .................................................................................................................. i  
List of figures .......................................................................................................................... ii  
List of tables .......................................................................................................................... v  
List of abbreviations .............................................................................................................. vii  

1 INTRODUCTION .................................................................................................................. 1  

2 FACTOR INVESTING ......................................................................................................... 3  

2.1 RETURN-BASED ANOMALIES AND ASSET PRICING ................................................ 3  
2.2 FACTOR RETURN DRIVERS ......................................................................................... 7  
2.3 QUALITY INVESTING ..................................................................................................... 8  

2.3.1 Quality in a nutshell..................................................................................................... 8  
2.3.2 The quality premium ................................................................................................. 8  
2.3.3 Flight to quality ........................................................................................................ 10  
2.3.4 Multi-factor investing: quality and value ................................................................. 11  

3 REVIEW OF QUALITY CHARACTERISTICS AND MODELS .................................... 15  

3.1 RETURN-BASED ANOMALIES LINKED TO QUALITY ............................................... 16  

3.1.1 Profitability ............................................................................................................... 16  
3.1.2 Growth ...................................................................................................................... 17  
3.1.3 Earnings quality ....................................................................................................... 18  
3.1.4. Safety .................................................................................................................... 20  
3.1.5. Payout .................................................................................................................... 23  
3.1.6. R&D intensity ........................................................................................................ 24  

3.2. MULTI-METRIC QUALITY SCORES ............................................................................... 25  

3.2.1 Graham ..................................................................................................................... 25  
3.2.2 Piotroski – Financial strength .................................................................................. 26  
3.2.3 Asness et al. (2013) – Quality Minus Junk ............................................................... 27  

3.3. INDUSTRY .................................................................................................................. 30  

4 EMPIRICAL PART ............................................................................................................. 33  

4.1 SURVEY OF PRACTITIONERS ..................................................................................... 33  
4.2 ANALYSIS OF SURVEY RESULTS ............................................................................. 34  

4.2.1 Identification of sub-factors valued by practitioners ................................................ 34  
4.2.2 Identification of sub-factors valued by practitioners ................................................ 35  
4.2.3 Weighting of the different dimensions ..................................................................... 39  

4.3 DETERMINATION OF SCORES ...................................................................................... 40  

4.3.1 Quality scores .......................................................................................................... 40  
4.3.2 QARP score .............................................................................................................. 42  

4.4 PORTFOLIO CONSTRUCTION AND ANALYSIS METHODOLOGY .......................... 43  

4.5 PERFORMANCE MEASURES ......................................................................................... 44  

4.5.1 Sharpe ratio .............................................................................................................. 44  
4.5.2 Sortino ratio .............................................................................................................. 44  
4.5.3 Value at risk .............................................................................................................. 45  
4.5.4 Expected shortfall ..................................................................................................... 46  
4.5.5 Maximum drawdown ............................................................................................... 46  

4.6 QMJ PORTFOLIO ANALYSIS ...................................................................................... 47  

4.6.1 Analysis and comparison of QMJ 1 vs. QMJ 2 and QMJ 3 ....................................... 47  
4.6.2 Comparison with academic definition: GPOA ......................................................... 49  
4.6.3 Analysis of individual dimension of QMJ 1 ............................................................. 50  

Julien DACO - Thesis | Page iii
4.6.4 Size premium controlling for QMJ 1 ................................................................. - 51 -
4.6.5 Recession vs expansion periods (NBER) .................................................. - 52 -
4.7 QARP LONG-SHORT PORTFOLIO ANALYSIS .................................................. - 53 -

5 CONCLUSION ........................................................................................................... - 57 -

6 BIBLIOGRAPHY ...................................................................................................... I

7 APPENDIX .............................................................................................................. IX

7.1 INTERNATIONAL PERFORMANCE OF DIFFERENT QUALITY MEASURES ......................... IX
7.2 LIST OF MEASURES DEFINITIONS ..................................................................... X
  7.2.1 Profitability measures .................................................................................. XI
  7.2.2 Earnings Quality measures ......................................................................... XI
  7.2.3 Safety measures ......................................................................................... XI
  7.1.4 Growth measures ....................................................................................... XII
7.3 CORR AND REG PROCEDURE ON QUALITY 1-2-3 .................................................. XIII
7.4 CORRELATION BETWEEN QMJ 1 AND OTHER FACTORS ....................................... XIV
7.5 CUMULATIVE RETURNS OF INDIVIDUAL MEASURES ....................................... XV
7.6 BASIC STATISTICS OF INDIVIDUAL PORTFOLIO ............................................ XVII
7.7 REGRESSION: SIZE PREMIUM CONTROLLING FOR QMJ ........................................ XVIII
7.8 PROC TTEST AND NORMALITY CHECKS ........................................................... XIX
List of figures

Figure 1: Quantile regression of Quality .............................................................................................................. 10
Figure 2: Survey results - Quality dimensions ..................................................................................................... 35
Figure 3: Survey results - Profitability measures ................................................................................................. 37
Figure 4: Survey results - Safety measures ......................................................................................................... 37
Figure 5: Survey results - Growth measures ....................................................................................................... 38
Figure 6: Survey results - Payout measures ......................................................................................................... 39
Figure 7: Survey results - Dimensions weighting ............................................................................................... 39
Figure 8: Illustration of VaR - Historical method (left) vs. Variance-Covariance method (right) .................. 46
Figure 9: Cumulative returns of QMJ 1 and GPOA ........................................................................................... 49
Figure 10: Alphas generated by ROIC, LEV, ACC and Growth 1 ...................................................................... 51
Figure 11: Cumulative return of QMJ 1 portfolio ................................................................................................ 52
Figure 12: Performance of the value (Fama French US HML). QARP and QMJ 1 factors ............................. 54
Figure 13: International performance of different Quality measures ................................................................. IX
Figure 14: Correlation between QMJ 1 and other factors .................................................................................... XIV
Figure 15: Cumulative returns of individual Profitability measures ................................................................. XV
Figure 16: Cumulative returns of Quality Earnings 1 (ACC) ........................................................................... XV
Figure 17: Cumulative returns of LEV and Altman's Z-score .......................................................................... XVI
Figure 18: Cumulative returns of individual Growth measures ....................................................................... XVI
List of tables

Table 1: Sample of quality definition from industry ............................................................. - 31 -
Table 2: List of Quality dimensions and underlying measures ................................................ - 36 -
Table 3: Basic statistical measures, normality test and tests for location on QMJ 1-2-3 ............... - 47 -
Table 4: Performance and risk measures of QMJ 1-2-3 ........................................................... - 48 -
Table 5: QMJ 1 alphas ........................................................................................................... - 48 -
Table 6: Performance of QMJ 1 vs. GPOA ............................................................................. - 50 -
Table 8: Alphas generated by GPOA ...................................................................................... - 50 -
Table 9: Size premium controlling for junk ............................................................................ - 52 -
Table 10: Performance and risk measures of QMJ 1, Value and QARP ................................. - 55 -
Table 11: Correlation between QMJ 1-2-3 ............................................................................. XIII
Table 12: Regression of QMJ 1 on QMJ 2 ............................................................................. XIII
Table 13: Regression of QMJ 1 on QMJ 3 ............................................................................. XIII
Table 14: Regression of SMB on QMJ .................................................................................... XVIII
# List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>Low accruals</td>
</tr>
<tr>
<td>AIF</td>
<td>Alternative Investment Funds</td>
</tr>
<tr>
<td>BM</td>
<td>Book-to-Market</td>
</tr>
<tr>
<td>CFOA</td>
<td>Cash flow over assets</td>
</tr>
<tr>
<td>CMA</td>
<td>Conservative Minus Aggressive</td>
</tr>
<tr>
<td>EPS</td>
<td>Earnings per Share</td>
</tr>
<tr>
<td>GARP</td>
<td>Growth at reasonable price</td>
</tr>
<tr>
<td>GMAR</td>
<td>Gross margin</td>
</tr>
<tr>
<td>GPOA</td>
<td>Gross profit over asset</td>
</tr>
<tr>
<td>HML</td>
<td>High Minus Low (Value factor)</td>
</tr>
<tr>
<td>IVOL</td>
<td>Idiosyncratic Volatility</td>
</tr>
<tr>
<td>LEV</td>
<td>Low Leverage</td>
</tr>
<tr>
<td>Mkt-rf</td>
<td>Market factor</td>
</tr>
<tr>
<td>QARP</td>
<td>Quality at reasonable price</td>
</tr>
<tr>
<td>QMJ</td>
<td>Quality Minus Junk</td>
</tr>
<tr>
<td>ROA</td>
<td>Return on assets</td>
</tr>
<tr>
<td>ROE</td>
<td>Return on equity</td>
</tr>
<tr>
<td>RMW</td>
<td>Robust Minus Weak</td>
</tr>
<tr>
<td>SMB</td>
<td>Small Minus Big (Size factor)</td>
</tr>
<tr>
<td>UMD</td>
<td>Up Minus Down (Momentum factor)</td>
</tr>
</tbody>
</table>
1 Introduction

This thesis focuses on factor investing and more particularly on the Quality factor. While the foundations for this factor have been set decades ago, it appears to become increasingly popular since the last financial crisis. At first sight two characteristics make this factor special. First of all, it is somewhat complicated to understand who is on the other side of the trade since one could reasonably assume that investors have no reason to shun Quality stocks. Then, while the Quality factor is well grounded in academic research and has solid explanations (MSCI, 2013a), no standard Quality definition exists which makes it difficult to understand what Quality really is or measures.

Hence, the thesis aims at highlighting the high complexity of quality investing and the underlying large dispersion in Quality definitions. While, historically, other factors have experienced some marginal discrepancies between measures used by practitioners and academics before converging towards the most effective one, a potential convergence appears much more challenging in the case of Quality. Indeed, Quality definitions even vary significantly among those two groups. Globally, the problem seems related to a fundamental difference in the perceptions of Quality and its purpose before being a debate about the ideal measures to capture the factor. Consequently, a generally accepted Quality definition is developed with a panel of practitioners and then compared to a purely academic definition.

In general terms, the thesis is organised as follows:

Chapter 2 – Factor Investing – operates a brief retrospective of asset pricing theory and introduces the notion of factor and its sources of returns. The Quality factor is then introduced as well as its interesting contribution in multi-factor investing. Chapter 3 – Review of Quality characteristics and models – goes over a series of return-based anomalies linked to Quality, then introduces some multi-metric Quality scores, among which the Quality Minus Junk factor which is central to the empirical part of the Thesis, and finally confronts some Quality definitions from the industry. Chapter 4 – Empirical part – aims at developing a generally accepted Quality definition from a panel of practitioners and then analyses its performance.
Finally, Chapter 5 concludes and highlights the main results of this thesis regarding Quality investing.
2 Factor Investing

2.1 Return-based anomalies and asset pricing

Before the 1970s, conventional wisdom around investing was mainly related to stock picking based on normative theories instead of using scientific methods and empirical observations. Indeed, at that time, financial data and computational power were not readily available. Investors relied primarily on the Capital Asset Pricing Model (CAPM) theory developed in the early 1960s by Sharpe (1964), Treynor (1962), Lintner (1965a, 1965b) and Mossin (1966). This single factor model, built upon the earlier research of Markowitz (1952) on modern portfolio theory, considers a world where markets are efficient and investors are rational, broadly assumes that the only risk is the market risk, and suggests the following linear relationship between risk and return:

\[ E(R_i) = R_f + \beta_a (E(R_m) - R_f) \] (2.1)

Where:

- \( E(R_i) \) denotes the expected return on a given asset;
- \( R_f \) denotes the risk-free rate of interest;
- \( E(R_m) \) denotes the expected return of the market;
- \( \beta_a \) represents the volatility (systematic risk) of the security in comparison to the market and is computed as follows:
  \[ \frac{\text{Cov}(R_i, R_m)}{\text{Var}(R_m)} \].

This situation changed in the 1970s once better financial datasets and increasing computational power made the rise of empirical research possible. The story of factor investing starts at that time, when the prevailing assumptions of the CAPM began to be challenged by academic researchers. Since that moment, a lot of criticisms of CAPM have emerged and many authors have proposed alternative models to improve it. Hence, Haugen and Heins (1972) analyse the relationship between risk and realized returns and conclude, somewhat surprisingly, that this

---

1 A factor can be thought of as any characteristic related to a group of securities that is important in explaining their return and risk (MSCI, 2013a).
relation is not linear and that low risk stocks generate higher returns than the CAPM would lead to expect. This finding marks the emergence of low volatility investing and represents one of the building blocks of factor investing. Four years later, Ross (1976) came up with a different theory of what drives stock returns, the Arbitrage Pricing Theory (APT). The APT framework supports that the expected return of a security can be modelled as a function of different macroeconomic factors or theoretical market indexes (MSCI, 2013a). This multi-factor model can be credited with popularizing the original term “factor”. Unlike CAPM which contains only one fixed factor (the market factor) and one beta, APT does not precise which factors should be included in the model. Besides, the number of these factors is likely to vary over time and across markets. This requires the security’s beta in relation to each separate factor. APT is expressed as follows:

\[ E(R_i) = R_f + \beta_{i1} \cdot f_1 + \beta_{i2} \cdot f_2 + \ldots + \beta_{in} \cdot f_n + \alpha \]  \hspace{1cm} (2.2)

Where:

- \( f_j \) denotes the risk premium of factor \( j \) for \( j = 1, \ldots, n \);
- \( \beta_{ij} \) denotes the sensitivity of the \( i^{th} \) asset to factor \( j \).

Basu (1977) shows that stocks with high earnings/price ratio earned, on average, significantly higher returns than stocks with low earnings/price ratio.

Banz (1981) examines the relationship between the return and the total market value of NYSE common stocks and argues that smaller companies have had higher risk adjusted returns, on average, than larger companies.

Another contradiction is brought by DeBondt and Thaler (1985) who highlight the tendency of returns to reverse over long horizons. Indeed, they find that stocks that have had poor returns over the past three to five years have, on average, earned higher returns than winners over the next three to five years.

Chan, Hamao and Lakonishok (1991) show that book-to-market equity is significantly positively correlated with expected returns.

In the light of those findings and given the lack of consistency of CAPM, Fama and French (1992, 1993) expand on the CAPM and suggest a model that controls for the size effect as well
as the book-to-market ratio (value factor). Their Fama and French three-factor model takes the following form:

\[ E(R_i) = R_f + \beta_1 (E(R_m) - R_f) + \beta_2 \cdot SMB + \beta_3 \cdot HML \]  \hspace{1cm} (2.3)

Where:
- SMB stands for “Small (capitalization) Minus Big” and represents the historic excess returns of small companies over large ones;
- HML stands for “High (book-to-market ratio) Minus Low” and represents the historic excess returns of value stocks over growth stocks.

Just a year later, Jegadeesh and Titman (1993) came up with the momentum effect, stating that past leaders on performance are also very likely to become future winners, and mark a new milestone for factor investing. This finding marks a key turning point in asset pricing. Indeed, instead of looking inside the DNA of a security, at its risk exposures, one started to look outside, at the behaviour of investors. Academicians understood that realized returns encompass an element related to expected returns. In other words, stock prices today are influenced by investors’ expectations for tomorrow and vice versa. This statement actually lays down the foundation for behavioural finance. Consequently, Carhart (1997) extends the Fama and French three-factor model by adding momentum as a fourth factor. Hence, the Carhart four-factor model takes the following form:

\[ E(R_i) = R_f + \beta_1 (E(R_m) - R_f) + \beta_2 \cdot SMB + \beta_3 \cdot HML + \beta_4 \cdot UMD \]  \hspace{1cm} (2.4)

Where:
- UMD stands for “Up Minus Down” and represents the historic excess returns of highest momentum stocks over lowest momentum stocks for the prior year.

Recently, Fama and French (2015) have revisited their three-factor model. They suggest that the book-to-market ratio may be a noisy proxy for expected return because the market value of a stock also reflects forecast of profitability and investment (Fama and French, 2015). Therefore, they add profitability and investment factors to their three-factor model in order to better isolate the information in stock prices about expected returns. The resulting Fama and French five-factor model takes the following form:
\[ E(R_t) = R_f + \beta_1 (E(R_m) - R_f) + \beta_2 \cdot SMB + \beta_3 \cdot HML + \beta_4 \cdot RMW + \beta_5 \cdot CMA \]  

(2.5)

Where:

- RMW stands for “Robust Minus Weak” and represents the historic excess returns of the most profitable firms over the least profitable;
- CMA stands for “Conservative Minus Aggressive” and represents the historic excess returns of firms that invest conservatively over those which invest aggressively.

In fact, the factors and models described above are just a tiny snapshot of what has been documented in the literature so far. Indeed, hundreds of factors and papers attempt to explain the cross-section of expected returns. Harvey, Liu and Zhu (2015) claim that no less than 316 factors have been tested to this end and project forward 20 years assuming the rate of factor productions remains similar to last few years, which would lead to approximately 600 factors. While some of these factors are derived from economic theories, others are discovered from purely empirical exercises which leave more room for data mining and data snooping\(^2\). Hence, in the face of such a factor proliferation, the authors make a call for extreme caution since some factors would be deemed “significant” by chance and even consider that the usual statistical significance cut-offs in asset pricing tests should be raised (e.g. t-statistic that exceeds 3.0 – i.e. a p-value of 0.27% - rather than 2.0). By and large, this plethora of factors forms what Cochrane (2011) characterizes as a “zoo of factors” which are ranked among three main categories of factors: fundamental factors capturing stock characteristics, statistical factors (e.g. principal component analysis) and macroeconomic factors (e.g. GDP, yield curve, etc.).

Still, factor investing, which consists of the investment process that aims to harvest risk premia through exposure to factors (MSCI, 2013a), is increasingly in the spotlight. Among the zoo of factors, six risk premia factors are considered as well-grounded in the literature and have robust explanation regarding the way they provide persistent premia over time. Broadly speaking, these factors are:

1. **Low Size** (Small cap): This factor captures the excess returns realized by small caps over large caps.

---

\(^2\) Data snooping or data dredging consists in using data mining to discover patterns in data which can be presented as statistically significant, without first establishing any hypothesis.
2. **Value**: This factor captures the excess returns realized by stocks which have low market price relative to their fundamentals, and is commonly based on book-to-market or price-to-earnings ratios.

3. **Momentum**: This factor captures the excess returns realized by past leaders on performance over poor performing stocks, and is commonly based on relative returns over the last 3 to 12 months.

4. **Low volatility**: This factor captures the excess returns realized by stocks that have lower than average volatility, and is commonly based on standard deviation over the previous years and Beta.

5. **Dividend Yield**: This factor captures the excess returns realized by stocks with higher than average dividends yields.

6. **Quality**: As further explained in section 2.3.1, there is no unified theory clearly defining the Quality factor. However, the broad idea behind this factor consists in capturing the excess returns realized by stocks with strong fundamentals.

### 2.2 Factor return drivers

The literature review allows to identify two schools of thought when it comes to explaining what drives factor returns:

- The first group envisions a world where the efficient market hypothesis holds; that is, a world where markets are efficient and investors are rational. Therefore, they consider that factor returns are linked to several sources of systematic risks. For example, in the case of the Size factor, Zhang (2006) argues that small companies earn a premium because these are less transparent, Liu (2006) mentions their lower liquidity while Dichev (1998) suggests their higher probability to be distressed. For those academicians assuming that factor premia exist as a reward for bearing systematic risk, the factors are likely to persist over time (MSCI, 2013a).

- On the other hand, the second group mainly subscribes to behavioural finance theory and considers that investors experience behavioural bias (e.g. overreaction). The factor premia would not be linked to any source of systematic risks in this case. Besides, a subgroup relates factor premia to constraints and frictions investors have to face (e.g. investment time horizon) (MSCI, 2013a). For those academicians, factors are likely to persist as long as the behavioural bias/constraints hold.
2.3 Quality investing

2.3.1 Quality in a nutshell

In contrast to well-established equity factors such as size or value, one can note that the quality factor lacks a clear and universally accepted definition. Although the intuitive notion behind this concept is that high-quality firms should have better overall performance than low-quality firms, discriminating characteristics which are taken into account have been strongly debated throughout academic literature and industry. For instance, a value stock can be defined in a number of ways (including book-to-price and earnings-to-price ratios) but each definition is logically related to the notion of cheapness and provides similar information. Similarly, when asked the hallmark characteristic of a small-cap stock, one would reply its small relative size. However, the answer is much less trivial for quality whose dispersion in definitions is much larger. One can reply using profitability, earnings quality, growth, cash generation, credit quality, payout, growth, management efficiency, or even a combination of these elements and still be correct. Generally, the term “quality” itself is an abstract concept that encompasses an element of subjectivity varying in accordance with the context. It is not surprising, therefore, that quality, once transposed in the equity universe, leads to a large number of definitions reflecting the diversity of views.

Asness, Frazzini and Pedersen (2013) broadly define quality as “characteristics that investors should be willing to pay a higher price for, everything else equal” (Asness et al., 2013, p. 2). Since investors are likely to place value on numerous fundamental characteristics spanning from profitability and growth to safety, quality appears as a multidimensional concept, which only complicates the matter. Indeed, one cannot directly compare a measure of profitability with a measure of solvency for instance. Moreover, while some academicians define quality with a single measure, others are using multi-metric definitions, which makes the comprehension of what quality really measures even harder.

2.3.2 The quality premium

While academicians seem globally convinced that high-quality companies yield higher returns, some practitioners appear in some ways more suspicious of that claim. The main reason has to do with fundamental investment beliefs and raises the question of who is on the other side of the trade. Actually, it is somewhat complex to figure out why some investors would shun high-
quality companies. Indeed, most of the characteristics attributed to quality firms seem to be characteristics that one would realistically choose to invest in and not to avoid. The second reason is more empirical and related to the large dispersions in quality definitions. Indeed, different definitions can deliver significantly different results so that practitioners implementing quality investing strategies are likely to have contradictory assessments. Nevertheless, some interesting theories attempt to explain what has driven the quality premium over time.

Asness et al. (2013) show that high-quality stocks exhibit, on average, higher prices but argue that the explanatory power of quality on price is modest. Therefore, high-quality stocks are still underpriced while bad-quality junk stocks are overpriced, which leads to high-quality stocks historically experiencing high risk-adjusted returns while junk stocks experience negative risk-adjusted returns (Asness et al., 2013). Therefore, the quality premium may be attributed to the failure of the market to price correctly quality characteristics as well as the persistence of those characteristics.

Hunstad (2014) also exhibits results which are counterintuitive for adherents to the CAPM and its “higher return equals higher risk” rule. Indeed, they show that the Sharpe ratio (see definition in section 4.5.1) for high-quality stocks is significantly larger than that of low-quality stocks. This implies that, empirically, the higher returns of those stocks is associated with significantly lower levels of risk. The explanation here provided regarding the quality premium is based on heterogeneous investing. Contrary to the CAPM which assumes universal risk aversion; that is, risk-averse investors trying to avoid risk and requiring some additional premia to undertake investments with a high degree of uncertainty, the authors assume that equity market investors are a mix of risk-averse, risk-neutral and risk-seeking profiles. This reasoning is consistent since the aggregate population exhibits risk-averse behavior (e.g. insurance products) as well as risk-seeking ones (e.g. lotteries). When considering this heterogeneous patterns of investors, the analysis of the distribution of quality returns which expands as you move from high-quality to junk stocks makes perfect sense (see Figure 1). Indeed, high-quality stocks with low levels of volatility are here the domain of risk-averse investors who require more certainty in their investments. On the other hand, low-quality stocks that are more risky/volatile appear more attractive to risk-seekers who bid the stock prices up to the point that expected returns are reduced (Hunstad, 2014).
Joyce and Mayer (2012) base their reasoning on a statement of Graham arguing that real risk was “the danger of a loss of quality and earnings power through economic changes or deterioration in management” (as cited by Montier, 2007). Therefore, they consider that a portfolio made up of firms with high and stable profits should control “real risk” and results in low and stable “price risk” (Joyce and Mayer, 2012). Profits is thus seen as ultimately driving returns. Contrary to Modigliani-Miller’s theory, they show that companies with persistently high profitability have lower leverage and vice versa. Therefore, like Asness et al. (2013) and Hunstad (2014) suggest, this creates an opportunity for investors to experience both higher returns and lower risk. Their explanation for the quality premia is based on the market’s lack of interest in the anomaly. Therefore, according to the authors, the market tends to systematically underprice quality stocks whose stability may seem unexciting while repeatedly giving capital injections to money-losing firms that survive to destroy capital in the future (Joyce and Mayer, 2012).

2.3.3 Flight to quality

Asness et al. (2013) show that the price of quality, that is, the price supplement accepted by investors for higher quality stocks, varies over time. Unsurprisingly, these price movements are quite well correlated with periods of uncertainty in the financial or international markets, which suggests that high-quality stocks are good candidates for the flight to quality described as the “knee-jerk movements towards solid fundamentals during tail events” (Joyce and Mayer, 2012,
Hence, Asness et al. (2013) observe that the price of quality was quite low leading into the 1987 crash and then increased, reaching highs in late 1990 (first Gulf War). The authors make the same analogies between the height of the internet bubble in February 2000 and in late 2002 (after the Enron and WorldCom scandals) as well as between 2007 (leading into the global financial crisis) and in early 2009 (during the height of the banking crisis).

Joyce and Mayer (2012) also emphasize the fact that the sole exception to the market’s inattention is found during periods of intense stress. As if investors become all of a sudden aware (consciously or subconsciously) of quality when they are worried. For his part, Hunstad (2014) makes the assumption that risk preference is not necessarily static over time. Therefore, one can reasonably assumes that there exists less risk-seekers during periods of market distress and more during times of market recovery.

Since investors are likely to change their risk appetite depending on the market conditions, quality factor performance is inevitably conditioned by those flights to quality. Indeed, the price of quality negatively predicts the future return on a high-quality stocks; that is, a higher price of quality is associated with a lower return on buying high-quality stocks (Asness et al., 2013). Therefore, a portfolio made up of high-quality stocks is expected to do very well during market drawdowns.

2.3.4 Multi-factor investing: quality and value

Literature and practice have shown that quality goes hand in hand with value investing. This strategy consists in investing in “cheap” companies displaying high valuation ratios such as book-to-market or earnings-to-price. Indeed, as argued by Basu (1977) and Fama and French (1992, 1993), those stocks tend to generate significantly higher returns than their opposite growth stocks. Explanations for this value premium are based on both systematic risk and market inefficiencies. Hence, Fama and French (1992) support the idea that high BM companies tend to be financially distressed and thus command a premium for bearing higher risk. On the other hand, Lakonishok, Sheifer and Vishny (1994) stand the market mispricing reasoning and argue that value stocks represent neglected stocks where poor prior performance has led to pessimistic expectations about future performance. However, this pessimism dissipates in the following periods, as evidenced by positive earnings surprises at earnings announcements which serve as a catalyst (La Porta et al., 1997).
While some investors are sceptical that unconditionally investing in high-quality stocks irrespective of their price is a sensible strategy, combining both quality and value factors appears as a promising investment strategy. Indeed, buying high-quality stocks makes sense but if their respective prices are over-inflated, the potential for future above average returns is seriously debatable.

This reasoning is well illustrated by the Nifty Fifty. In the 1960s and 1970s, the Nifty Fifty consisted of 50 very popular large-stocks on the NYSE which were considered by investors as highly desirable in a buy and hold strategy due to their fast-growing characteristics. Among these Nifty Fifty featured high-profile companies such as General Electric, IBM, Procter & Gamble, or Polaroid. The thing is that investors had pushed the prices of their favourite stocks to unjustified high but still continued to buy those stocks without taking their valuations into account. For instance, at the end of 1972, the Nifty Fifty was trading at a P/E of 40; that is, twice as high as the P/E of the S&P 500 at that time. While value investing was a well-rounded investment practice, the Nifty Fifty gave rise to a “growth at any price” paradigm (Kalesnik and Kose, 2014). Unfortunately for those investors, the Nifty Fifty suffered a hit during the 1970s bear market and valuations dropped to levels similar to the rest of the markets. Globally, those stocks underperformed the broad market and never caught up. Kalesnik and Kose (2014) show that, from 1973 to 2013, the S&P 500 investors would have earned 23% more than the Nifty Fifty investors.

Therefore, in the light of this example, conditioning a value investing strategy on the quality of the firms makes perfect sense. Indeed, while some firms are overpriced, others simply deserve a higher price because of their high quality. Conversely, some companies deserve a lower price because of their poor fundamentals while others are underpriced. Consequently, mixing value and quality investing allows to detect great opportunities – i.e. good companies at cheap price – while avoiding value trap – i.e. “a financial instrument (stocks or bonds) that appears cheap on historical measures or valuation grounds, such as P/E ratio, but the price never recovers to fair value” (Financial Times, 2017). Besides, Piotroski (2000) argues that the success of value strategies relies on the performance of a few firms while the poor performance of many deteriorating companies is tolerated. Consequently, given the very diverse outcomes realized within a value portfolio, discriminating between strong and weak companies appears sensible.

Whereas the benefits of combining quality and value, based on the “quality at fair price” statement, seem straightforward, it is also a good association from a diversification perspective.
Indeed, factor investing is subject to factor volatility. Whereas well-established factors have experienced risk-adjusted returns over long time periods, they also exhibit significant cyclicality over short periods of time, including periods of underperformance. Consequently, there is no free lunch captured by factor investing. In an effort to address this volatility, one can implement a multi factor strategy by selecting factors diversifying each other (MSCI, 2013a). This is the case of quality and value since high-quality firms tend to be expensive while value stocks tend to be low-quality. In other words, quality strategies are short value and vice versa (Novy-Marx, 2014). Consequently, because a quality portfolio has a negative correlation with a value portfolio, an investor willing to invest in both exposures can achieve significant diversification benefits (Kozlov and Petajisto, 2013). Similarly, Asness et al. (2013) show that the Sharpe ratio of a QARP strategy is higher than either value or quality alone.

In practice, Warren Buffett stands as a famous proponent of investment strategies mixing value and quality. The significant returns generated by his company, Berkshire Hathaway, has been a source of fascination for the investors. Frazzini, Kabiller and Pedersen (2013) show that among all U.S. stocks having been traded for more than 30 years between 1923 and 2011, Berkshire Hathaway is the one that has the highest Sharpe Ratio. Nevertheless, this performance is neither luck nor magic but rather a compensation for using leverage combined with a focus on cheap (value stocks), low-risk (with low beta and low volatility), quality (profitable, stable, growing, and with high payout ratios) stocks. Indeed, while Berkshire Hathaway generates significant alpha to traditional risk factors, this alpha becomes insignificant when controlling for the Quality-Minus-Junk factor of Asness and al. (2013) and the Betting-Against-Beta factor of Frazzini and Pedersen (2014).
3 Review of Quality characteristics and models

In order to have a comprehensive understanding of all the characteristics that have been attributed to quality through literature, this third chapter first introduces a series of return-based anomalies whose effects are linked to the outperformance of high-quality stocks. In order to do so in a structured fashion, 6 dimensions of quality have been identified and anomalies are ranked accordingly. These six quality dimensions, which will be retained and frequently mentioned in the empirical part of the thesis are the following: Profitability, Growth, Earnings quality, Safety, R&D intensity, and Payout. It is worth mentioning that one can consider that the anomalies presented are of two kinds. The first group encompasses anomalies that Kyosev, Hanauer, Huij and Landsdorp (2016) consider as very academic and which have been developed with the aim of directly capturing quality. These are the gross profitability, the accruals and the net stock issuance. The second group contains anomalies that have been linked to the quality factor and which are to some extent less specific to academics. While these are defended by academics, they also tend to be more popular among practitioners and eventually shared in their current vision of quality.

This chapter then addresses a series of quality definitions whose underlying quality scores encompass more than one metric. A particular attention is paid to the Quality Minus Junk framework developed by Asness et al. (2013) which serves as an interesting basis for the empirical research of the thesis. Finally, a sample of quality definitions developed in the industry are introduced in an attempt to understand how they differ from those supported by academicians.
3.1 Return-based anomalies linked to Quality

3.1.1 Profitability

3.1.1.1 ROE

Among the profitability dimension of quality stating that profitable firms tend to yield significantly higher average returns than unprofitable ones, ROE stands undoubtedly as one of the most popular profitability measures when it comes to describe quality stocks, especially among practitioners. Among others, Hou, Xue and Zhang (2015) show that high ROE stocks earn, on average, higher returns than low ROE stocks. Consequently, they propose an alternative to the Fama and French three-factor model that consists of a market factor, a size factor, an investment factor and a ROE factor.

3.1.1.2 ROIC

In his “Little Book That Beats the Markets”, Greenblatt (2006) highlights the importance for value investors to pay attention to quality, and more particularly to capital productivity. To this end, the author delivers a “Magic Formula” combining value and quality and aiming at buying good companies at bargain prices. His “Magic Formula” consists in investing in the firms displaying the best combined rank on their earnings yields (value metric) and ROIC (quality metric). Those measures are defined as follows:

\[
\text{Earnings yields} = \frac{EBIT}{\text{Enterprise Value}} \quad (3.1)
\]

\[
\text{ROIC} = \frac{EBIT}{\text{Net Working Capital} + \text{Net Fixed Assets}} \quad (3.2)
\]

Therefore, as an alternative to the popular ROE and ROA, Greenblatt (2006) here suggests to use the ROIC as the profitability measure. The inclusion of the EBIT instead of the reported earnings in the numerator allows to take account of the differences in the levels of debt and tax rates firms are operating with. This makes it possible to compare the operating earnings of different firms without suffering from the aforementioned distortions. The reason behind the denominator is to figure out how much capital is actually needed to conduct the business. Consequently, the net working capital and the net fixed assets are used in place of equity (for ROE calculation) and total assets (for ROA calculation). Net working capital is incorporated
into the denominator since firms need to fund their inventory and account receivables but do not have to shell out money for their accounts payables. Obviously, besides working capital requirements, a firm also has to fund the fixed assets (e.g. land, buildings, plant & equipments, etc.) required in order to conduct its business, which is materialized by the net fixed assets.

### 3.1.1.3 Gross profitability

Like a number of other academicians, Novy-Marx (2013) addresses the quality definition with the profitability theme. However, the author suggests that the profitability measures lying at the bottom of the income statement are much more polluted than those lying at the top and therefore less related to true economic profitability. As a consequence, the gross profitability scaled by assets is proposed as a game-changing top-line profitability measure which is superior in predicting future stock returns than bottom-line earnings. This ratio is very straightforward and computed as follows:

\[
\frac{\text{Gross profits} - \text{COGS}}{\text{Total assets}} = \frac{\text{Total revenues} - \text{COGS}}{\text{Total assets}}
\]  

(3.3)

The rationale behind Novy-Marx’s approach can be illustrated in the following way: it goes without saying that a company with both higher sales and lower production costs than its competitors is more profitable. Even so, this same firm may have lower earnings than competitors for a variety of reasons. For instance, the firm’s strategy may be to increase sales quickly thanks to massive advertising expenses, further reducing its net income below the level of less profitable competitors. In a similar fashion, the same reasoning can be applied to research & development expenses incurred to preserve competitive advantage or capital expenditures which increase the scale of the company’s operations while still reducing its level of free cash flows (Novy-Marx, 2013).

### 3.1.2 Growth

While value investors takes long position on value stocks and avoid low BM growth stock, also called glamour stocks, one should, however, reasonably expects that a high-quality company reports interesting growth prospects. With this in mind, Mohanram (2005) develops a score that aims at separating winners in stock performance from losers among low BM stocks. Whereas
the focus for growth companies has traditionally been on non-fundamental aspects of their operations, Mohanram’s G-score is still based on growth oriented fundamentals. He finds that future earnings performance are strongly correlated to current growth fundamentals which are effective at differentiating between future winners and losers. Indeed, companies with high G-scores generate significantly higher risk-adjusted returns than firms with low G-score (Mohanram, 2005). This suggests that the market fails to understand the aforementioned correlation between current growth oriented fundamentals and future earnings performance. This statement is in line with the tendency of markets to naïvely extrapolate current fundamentals of growth stocks (La Porta, 1996). Besides, Mohanram (2005) shows that the signals related to the stability of earnings and growth are particularly helpful in identifying stocks that are less likely to be overvalued due to naïve extrapolation of stock markets. Globally, these findings are consistent with the statement that, while some low BM stocks are value trap, others deserve their higher price given their higher quality, which in this case is materialized by good growth oriented fundamentals.

3.1.3 Earnings quality

Earnings quality is far from being a straightforward concept as reflected by the large number of proxies attached to this notion. Among others, one can cite as an example earnings persistence, magnitude of accruals, earnings smoothness, target beating or timely loss recognition. Globally, earnings quality can be defined as follows: “Higher quality earnings provide more information about the features of a firm’s financial performance that are relevant to a specific decision made by a specific decision-maker” (Dechow, Ge and Schrand, 2010, p.1). When it comes to quality investing, accruals have been quasi-unanimously adopted as the proxy that best predicts stock returns. For this reason, accruals are retained as the privileged proxy in the framework of the thesis.

3.1.3.1 Accruals

Sloan (1996) develops a well-known definition of quality focusing on earnings quality. Based on previous research from Graham, Dodd and Cottle (1962) that highlight the importance of adjusting current earnings to forecast future earnings power of a firm, the author focuses on the
different implications of cash flows and accruals for the assessment of future earnings and validates two major hypotheses:

- **H1**: The persistence of current earnings performance is decreasing in the magnitude of the accrual component of earnings and increasing in the magnitude of the cash flow component (Sloan, 1996).
- **H2**: The earnings expectations encompassed in stock prices fail to reflect fully the higher earnings persistence attributable to the cash flow component of earnings and the lower earnings persistence attributable to the accrual component of earnings (Sloan, 1996).

These findings give rise to a market inefficiency known as the “accrual anomaly”, stating that shares in companies that have a low (high) level of accruals tend to experience positive (negative) future abnormal returns, as if investors overemphasize accounting earnings and fail to distinguish its components and the importance of cash generation until that information impacts future earnings. Therefore, it can be implied from H2 that a simple investment strategy taking a long position in the stocks of companies reporting relatively low levels of accruals and a short position in the stocks of firms displaying relatively high levels of accruals achieves positive abnormal returns (Sloan, 1996). Consequently, accruals are isolated as a proxy for earnings quality. Broadly speaking, accruals are defined as “the difference between cash and accounting earnings, scaled by firm assets” (Novy-Marx, 2014, p.6) and the accrual component is computed as follows, using information from the balance sheet and income statement, as is common in the earnings management literature (Dechow, Sloan and Sweeney, 1995):

\[
Accruals = (\Delta CA - \Delta Cash) - (\Delta CL - \Delta STD - \Delta TP) - Dep \quad (3.4)
\]

Where:

- \( \Delta CA \) denotes change in current assets;
- \( \Delta Cash \) denotes change in cash and cash equivalents;
- \( \Delta CL \) denotes change in current liabilities;
- \( \Delta STD \) denotes change in debt included in current liabilities;
- \( \Delta TP \) denotes change in income taxes payable;
- Dep denotes depreciation and amortization.
3.1.4. Safety

3.1.4.1 Low leverage

In measuring the general level of indebtedness of a firm and providing information on its capital structure, the inclusion of the leverage in the quality definition makes sense. Indeed, this popular ratio is well-understood, widely used in practice and provides a first indication of the creditworthiness of the firm. The literature documents the leverage puzzle stating that future stock returns are negatively related to leverage. The rationale behind this puzzle is not straightforward since one may assume that investors would require a higher return from a firm with a less conservative leverage. George and Hwang (2009) provide an explanation based on differences in financial distress costs from one firm to another. Hence, firms adapt their capital structure to reflect their distress costs; that is, firms with high distress costs tend to use debt conservatively so that they have a low leverage and more generally a low probability of defaults. However, the high distress costs associated with low leverage firms makes them more exposed to systematic risk. Moreover, low leverage firms performance is affected to a higher degree when in financial distress – i.e. ROA drops and remains depressed and less foreseeable than high leverage firms (George and Hwang, 2009) – further increasing the exposure to systematic risk. Subsequently, this leverage puzzle may be attributed to the aforementioned exposure to systematic risk.

3.1.4.2 Financial distress

Since quality is often seen as the opposite of junk, the creditworthiness of a company emerges as a quasi-mandatory feature of a quality stock. In this regard, a large literature documents the use of accounting variables to estimate the probability of default that can be defined as “the probability that the borrower is unable or unwilling to fulfill terms promised under loan contract” (Cornett & Saunders, 2007, p. 309). Nevertheless, the methodologies and underlying ratios used to predict corporate defaults vary widely from one author to another. For example, after having compared a list of ratios individually, Beaver (1966) highlights the superior predictive power of the cash-flow to total-debt ratio while Tamari (1966) considers the current ratio. Besides, Hossari and Rahman (2005) isolate the ROA as the single most common ratio out of 53 corporate failure studies spanning from 1966 to 2002. Altman (1968) and Ohlson (1980) extend these findings using multivariate frameworks and propose two well-known and
widely accepted measures of financial distress; respectively, Altman’s Z-score and Ohlson’s O-score.

It is worth noting that, in line with the leverage puzzle, firms with greater distress intensity tend to deliver low average returns (Campbell, Hilscher and Szilagy, 2008; George and Hwang, 2009). This statement is supported by a whole variety of financial distress risk measures among which Altman’s Z-score (Dichev, 1998), Ohlson’s O-score (Dichev, 1998; Griffin and Lemmon, 2002), credit ratings (Avramov, Chordia, Jostova and Philipov, 2009), distance to default (Vassalou and Xing, 2004) and default risk measures from Moody’s KMV (Garlappi, Shu and Yan, 2006).

3.1.4.2.1 Altman’s Z-score

In his attempt to demonstrate the potential of financial ratio analysis, Altman (1968) develops a popular model whose purpose is to assess a firm’s financial health and subsequently to predict the probability of corporate bankruptcy. Out of a list of 22 popular variables, Altman (1968) finally identifies 5 ratios as being the best measures for bankruptcy. The resulting Z-score is computed as follows:

\[
Z = 1.2x_1 + 1.4x_2 + 3.3x_3 + 0.6x_4 + 1.0x_5
\]  

(3.5)

Where:

- \( Z \) denotes the score obtained by a given firm;
- \( x_1 \) denotes the ratio \( \frac{\text{Working capital}}{\text{Total assets}} \);
- \( x_2 \) denotes the ratio \( \frac{\text{Retained earnings}}{\text{Total assets}} \);
- \( x_3 \) denotes the ratio \( \frac{\text{EBIT}}{\text{Total assets}} \);
- \( x_4 \) denotes the ratio \( \frac{\text{Market value of equity}}{\text{Book value of total debt}} \);
- \( x_5 \) denotes the ratio \( \frac{\text{Sales}}{\text{Total assets}} \).

Stocks with a Z-score lower than 1.81 are then considered in a distress zone while those having a score higher than 2.99 are in a safe zone and are not likely to go bankrupt.
3.1.4.2.2 Ohlson’s O-score

Also in an attempt to predict corporate failure, Ohlson (1980) develops an alternative to the Altman’s Z-score. The resulting O-score, while being a more accurate predictor, is much more complex and derived from the linear combination of nine ratios. The computation of Ohlson’s O-score appears below:

\[
O = -1.32 - 0.407 \times \log \left( \frac{TA}{GNP} \right) + 6.03 \times TLTA - 1.43 \times WCTA + 0.076 \\
\times CLCA - 1.72 \times OENEG - 2.37 \times NITA - 1.83 \times FUTL \\
+ 0.285 \times INTWO - 0.521 \times CHIN
\]  

(3.6)

Where:

- TA denotes Total Assets;
- GNP denotes the GNP price-level index;
- TLTA is equal to \( \frac{Total \ Liabilities}{Total \ Assets} \);
- WCTA is equal to \( \frac{Working \ Capital}{Total \ Assets} \);
- CLCA is equal to \( \frac{Current \ Liabilities}{Current \ Assets} \);
- OENEG is equal to \( \begin{cases} 1 & \text{if Total Liabilities exceeds Total Assets} \\ 0 & \text{Otherwise} \end{cases} \);
- NITA is equal to \( \frac{Net \ Income}{Total \ Assets} \);
- FUTL is equal to \( \frac{Funds \ Provided \ by \ Operations}{Total \ Liabilities} \);
- INTWO is equal to \( \begin{cases} 1 & \text{if Net Income was negative for the last two years} \\ 0 & \text{Otherwise} \end{cases} \);
- CHIN is equal to \( \frac{NI_t - NI_{t-1}}{|NI_t| - |NI_{t-1}|} \) where NI is the net income for the most recent period.

3.1.4.3 Low beta

As explained in (Section 2.1) the Beta is a measure of the volatility (systematic risk) of a security in comparison to the market. This measure lies at the core of the CAPM stating that the expected excess return on a security equals its Beta times the expected excess returns on the market portfolio. However, after having tested a series of portfolios with different Betas, Black, Jensen and Scholes (1972) conclude that high-beta securities have significantly negative intercepts and that low-beta securities have significantly positive intercepts, which contradicts
the predictions of the model. Consequently, considering a low beta as a characteristic of quality stocks makes sense since it involves higher returns and lower volatility.

### 3.1.4.4 Low idiosyncratic volatility

While the Beta measures the systematic risk of a security, the idiosyncratic volatility refers to the unsystematic risk; that is, the risk that is inherent to a particular security and which can be mitigated through diversification in an investment portfolio. Like Beta, it has been shown that, around the world, stocks displaying recent past high idiosyncratic volatility tend to have much lower returns than stocks with recent past low idiosyncratic volatility (Ang, Hodrick, Xing and Zhang, 2009).

### 3.1.5. Payout

#### 3.1.5.1 Low external financing

Several studies highlight long-run abnormal stock returns following seasoned equity offerings (Loughran and Ritter, 1995), share repurchase announcements (Ikenberry, Lakonishok and Vermaelen, 1995) and stock mergers (Loughran and Vijh, 1997). From a behavioural perspective, this literature can be interpreted in the following way: in order to exploit mispricing of their securities in capital markets, firms issue equity when the stock price is overvalued and repurchase equity when it is undervalued. Pontiff and Woodgate (2008) broaden these studies and show that measures of share issuance exhibits a strong negative cross-sectional relation with the stock returns of U.S. firms for holding periods ranging from 1 month to 3 years. McLean, Pontiff and Watanabe (2009) expand this finding to a sample of 41 non-U.S. countries. Interestingly, the statistical significance attributed to the annual share issuance is greater than previously documented factors such as book-to-market, size, and momentum (Pontiff and Woodgate, 2008 ; McLean et al., 2009).

Just as in the case of equity issuance, future stock returns tend to be unusually low following bank borrowings (Billett, Flannery and Garfinkel, 2001) and debt offerings (Spiess and Affleck-Graves, 1999). Aforementioned research was dedicated to individual categories of financing transactions. Consequently, these studies do not take account of refinancing transactions leading to little or no net change in total capital. Indeed, some transactions only involve a shift
from one category to another. For instance, a firm may issue debt to repurchase equity. In that respect, Richardson and Sloan (2003) analyse the relation between external financing as a whole and the future stock returns and show that controlling for refinancing transactions results in superior predictive power. As a consequence, it is not so much the type of transactions that matter, but more specifically the extent to which a financing transaction impacts the change in net external financing.

3.1.5.2 Dividend yield

The notion that dividend yield, expressed as dividend-to-price ratio, forecasts future returns has a long tradition among both academicians and practitioners and dates back to Dow (1920). Many studies, among which Miller and Scholes (1982), then investigated the relation between dividend yield and future returns and demonstrate a significantly positive relation between dividend yield and stock returns. The rationale behind this statement is that stock prices are low relative to dividends when expected returns and discount rates are high, so that dividends yields should reflect changes in expected returns (Kotecha and Yadav, 1995).

In practice, many investors have implemented investment policies based on the aforementioned observation. For instance, the Dow dividend strategy consists in investing in the highest-yielding stocks from the 30 Dow Industrials. Historically, these stocks have been qualified as “Dogs of the Dow” since they often encompass some of the previous year’s worst performers. The behaviour of these stocks may, therefore, be explained by the market overreaction hypothesis (Domian, Louton, & Mossman, 1998).

3.1.6. R&D intensity

In a similar way as an investor that assigns good growth oriented fundamentals to high-quality stocks, one may focus more specifically on the firm’s R&D intensity. Joyce and Mayer (2012) argue that true competitive equilibrium is rarely seen in the economy and that persistent winners and persistent losers can be observed. To that extent, R&D is an activity likely to create a corporate moat that protects profitability prospects from competitive pressures and that keeps the firm among winners who tend to reports persistent and above-market profitability (Joyce and Mayer, 2012). Interestingly, Ciftci and Cready (2011) highlight the scale effect of R&D stating that the positive relation between R&D intensity and subsequent earnings increases with
firm size, and that the positive relation between R&D intensity and future earnings volatility decreases with firm size.

From a stock return perspective, R&D activities may depress current earnings and book value but is likely to boost future growth through sales and earnings growth (Mohanram, 2005). Lev and Sougiannis (1996) find a significant relation between firms’ R&D capital and future stock returns. They, therefore, suggest that this effect is either due to a systematic mispricing of R&D intensive company stocks or a risk premium for extra-market risk associated with R&D. Chan, Lakonishok and Sougiannis (2001) confirm the excess returns attributed to R&D intensive firms and show that R&D intensity is positively related to return volatility. Finally, Penman and Zhang (2002) state that the stock market tends to ignore the influences of conservative accounting on future earnings. Indeed, conservative accounting makes firms expense outlays such as R&D even if these items create intangible assets. Consequently, these unrecorded intangible assets reduce book values, making it more likely that a company reports a low BM for accounting reasons rather than over-valuation (Penman and Zhang, 2002).

An interesting measure to capture the R&D Intensity is the Price-to-Research (PRR) ratio which is used to compare the price of a company’s stock with its ability to potentially generate future profits from innovation. PRR is computed as follows:

\[
PRR = \frac{\text{Market capitalization}}{\text{R&D expense}}
\]  

(3.7)

3.2. Multi-metric Quality scores

3.2.1 Graham

Although considered as a pioneer of value investing, Graham (1973) had already realised that value and quality went hand in hand. His strategy was to buy undervalued stocks that still meet a series of quality characteristics. It is, therefore, not surprising that, among the 7 investment criteria from his famous “Intelligent Investor”, 5 are directly related to quality – i.e. adequate enterprise size, strong financial condition, earnings stability, uninterrupted dividend payments and EPS growth – and 2 relates to value – i.e. moderate price-to-earning and market-to-book ratios.
In his attempt to turn Graham’s five quality criteria into a quality investing strategy, Novy-Marx (2013) suggests a quality score ranging from 1 (lowest quality) to 5 (highest quality) and ascribable to each stock. This Graham score, also called G-score (different than Mohanram’s G-score), assigns:

1. One point if the current ratio is larger than 2;
2. One point if the firm’s net current assets exceed long term debt;
3. One point if the firm has a ten year history of positive earnings;
4. One point if the firm has a ten year history of returning cash to shareholders;
5. One point if the firm’s EPS is at least a third higher than it was 10 years earlier.

3.2.2 Piotroski – Financial strength

In contrast with the score developed by Mohanram (2005) focusing on low BM, the quality definition proposed by Piotroski (2000) is built in an attempt to separate winners from losers in the framework of a value investing strategy. Since high BM firms tend to be financially distressed (Fama and French, 1995), the quality score is based on financial strength indicators, such as profitability, operating efficiency, liquidity and leverage. Indeed, the rationale behind it is that examining fundamental signals reflecting changes in these parameters is likely to provide better prediction of future firm performance. Overall, the implied F-score relies on nine fundamental signals and is derived by summing nine corresponding binary variables taking either value 0 (indicating weakness) or 1 (indicating strength). More precisely, Piotroski’s F-score is computed as follows:

- **Profitability:**
  - One point if ROA is positive;
  - One point if CFO\(^3\) (cash flow from operations) is positive;
  - One point if ROA has improved between the current and prior year;
  - One point if the accruals (ROA – CFO) are negative.
- **Leverage, liquidity and source of funds:**
  - One point if the leverage ratio fell in the year preceding portfolio formation;
  - One point if the current ratio has improved between the current and prior year;

\(^3\) \(CFO = \frac{\text{Cash flow from operations}}{\text{Total assets}}\)
- One point if the firm did not issue common equity in the year preceding portfolio formation.

- Operating efficiency:
  - One point if the gross margin ratio has improved between the current and prior year;
  - One point if the asset turnover ratio\(^4\) has improved between the current and prior year.

### 3.2.3 Asness et al. (2013) – Quality Minus Junk

Asness et al. (2013) synthesize a series of findings in the asset price literature and define a quality stock using a set of dimensions derived from Gordon’s growth model. The latter is used to calculate the intrinsic value of a stock on the basis of the stock’s expected future dividends and takes the following form:

\[
P = \frac{D_1}{(K - G)} \tag{3.8}
\]

Where:

- \(D_1\) denotes the expected dividend per share in one year;
- \(K\) denotes the required rate of return;
- \(G\) denotes the dividend growth rate.

The authors rewrite this equation (3.8) by scaling prices by book value in order to improve the stationarity over time and in the cross section:

\[
\frac{P}{B} = \frac{1}{B} \cdot \frac{D_1}{(K - G)} = \frac{\text{profit}}{B} \cdot \frac{D_1}{(K - G)} = \frac{\text{Profitability} \cdot \text{Payout}}{\text{Safety - Growth}} \tag{3.9}
\]

The four right-hand side dimensions (Profitability, Payout, Safety and Growth) lie at the core of their quality definition and encompass no less than 21 measures. Consequently, the quality score is far less parsimonious than other previously seen definitions but the authors favour a

\[^4\text{Asset turnover ratio} = \frac{\text{Total sales}}{\text{Total assets}}\]
robust analysis where a specific measure does not drive the explanatory power of quality on price. Among the 21 measures included in the model appear many of the return-based anomalies related to outperformance of high-quality stocks and described in section 3.1.

The computation of the score starts with the calculation of intermediary scores for each quality dimension; that is, Profitability, Growth, Safety and Payout. These intermediary scores consist of the average of their respective underlying measures’ z-scores. The quality score is then obtained by averaging the four intermediary scores. This methodology is further explained in section 4.3. The computation of intermediary scores is performed as follows:

\[
Profitability = z(\text{gpoa} + \text{roe} + \text{roa} + \text{cfoa} + \text{gmar} + \text{acc})
\]

(3.10)

Where:

- gpoa denotes gross profits over asset;
- roe denotes return on equity;
- roa denotes return on assets;
- cfoa denotes cash flow over assets;
- gmar denotes gross margin;
- acc denotes accruals.

\[
Growth = z(\Delta \text{gpoa} + \Delta \text{roe} + \Delta \text{roa} + \Delta \text{cfoa} + \Delta \text{gmar} + \Delta \text{acc})
\]

(3.11)

Where:

- \(\Delta\) denotes five-year growth.

\[
Safety = z(\text{bab} + \text{ivol} + \text{lev} + \text{o} + \text{a} + \text{evol})
\]

(3.12)

Where:

- bab denotes low beta;
- ivol denotes low idiosyncratic volatility;
- lev denotes low leverage;
- o denotes Ohlson O-score;
- a denotes Altman Z-score;
- evol denotes low ROE volatility.
\[ Payout = z(z_{eiss} + z_{diss} + z_{npop}) \] (3.13)

Where:

- \( eiss \) denotes equity net issuance;
- \( diss \) denotes debt net issuance;
- \( npop \) denotes total net payout over profits.

The quality score is finally obtained by averaging the four intermediary scores:

\[ Quality = z(Profitability + Growth + Safety + Payout) \] (3.14)

On the basis of this Quality score, Asness et al. (2013) construct a QMJ (Quality Minus Junk) factor following Fama and French (1992, 1993) methodology which is further explained in section 4.4. More precisely, this portfolio is long the top 30% high-quality stocks and short the bottom 30% low-quality stocks. Overall, they show that their QMJ factor deliver significant risk-adjusted returns; has negative market, value and size exposure; has positive alpha; and generates high returns during market downturns which is in line with the concept of flight to quality.

Another interesting fact about the QMJ factor concerns its relation with the Size factor which has been strongly challenged lately. Asness, Frazzini, Israel, Moskowitz and Pedersen (2015) identify, among others, the fact that the size premium has a weak historical record, varies significantly over time, notably weakening after its discovery by Banz (1981), is not present for measures of size not relying on market prices, is concentrated among microcap stocks, mainly resides in January, is weak internationally, and is subsumed by proxies for illiquidity. Nevertheless, Asness et al. (2015) show that, controlling for the quality of a firm resurrects a stronger and more stable size premium while rejecting the above-mentioned statements. More specifically, since large firms tend to be high-quality firms and small firms tends to be low-quality firms, and given the fact that high-quality stocks tend to outperform junk stocks, one can conclude that the standard Size factor is fighting a headwind because of the poor quality of small stocks. Said differently, when comparing stocks of similar quality, smaller stocks significantly outperform larger ones on average, but the standard size effect hence suffers from a size-quality composition effect (Asness et al., 2013, 2015).
3.3. Industry

While already unclear among academicians, the diversity of views about the definition of a quality stock is even more complex when considering the industry. As seen in section 3.1, while some ratios have been explicitly related to quality – e.g. accruals (Sloan, 1996) or gross profit over assets (Novy-Marx, 2013) –, academicians tend to link many return-based anomalies to quality even if the link is not always straightforward. This leads to a large panel of quality measures, some of which are eventually too complex or unusual for practitioners (e.g. external financing or Ohlson’s O-score).

Table 1 shows a sample of quality definitions that can be found in the industry. It can be first stated that industry definitions tend to be multidimensional. Indeed, all of them display at least 3 metrics. Moreover, common patterns are observable – i.e. Profitability (ROE and ROA), Leverage and Earnings quality (accruals and earnings volatility). This is in line with Kyosev et al. (2016) which also point out the recurring appearance of the growth theme through their research of fund prospectuses, index methodologies and research notes. Furthermore, they group the leverage and earnings volatility under the stability label. Even though recurring dimensions are identified, the fact remains that definitions are globally different, which suggests that practitioners are not on the same wavelength when it comes to Quality investing.

From a performance perspective, Kyosev et al. (2016) show that more academic quality definitions such as gross profitability, accruals and net stock consistently have higher predictive power for future stock returns than the more popular ones used in the industry (see appendix 7.1). One may assume that this finding is related to some extent to what Beaver (1966) identifies as self-defeating popularity, stating that the most popular ratios will become the most manipulated by management, for purpose of window dressing, in such a way that destroys their utility. These discrepancies between academics and practitioners raise the issue of whether funds which offer exposure to the quality factor can use the quality returns documented by academic studies as expectation.

Interestingly, they note that such discrepancies are not specific to the Quality factor and were present for other factors as well just after their public dissemination. For instance, regarding the value factor as defined by Fama and French (1993), DiBartolomeo and Witkowski (1997) show that a significant number of mutual funds deviated from their classification, and Cooper et al. (2005) argue that funds included “value” in their name even though their positions did not
match the implied investment style. Finally, Kyosev et al. (2016) argue that these discrepancies between academics and practitioners became smaller over time as a result of a learning effect in financial markets where metrics evolve and converge towards the ones best predicting stocks returns. Therefore, they expect this reasoning to hold for quality since the factor is relatively young. All these key findings lead us to the empirical part of the thesis which, broadly speaking, aims at developing a generally accepted Quality definition that shifts practitioners towards more academic measures or at least towards less popular ones.

<table>
<thead>
<tr>
<th>Issuer</th>
<th>Quality definition</th>
</tr>
</thead>
</table>
| MSCI Quality Index (2013b)                  | • ROE  
• Leverage (debt to equity) 
• Earnings variability                                                              |
| S&P Quality Index (2017)                    | • ROE  
• Leverage (debt to equity) 
• Accruals                                                                            |
| FTSE Quality Index (2014)                   | • ROA  
• Change in asset turnover  
• Leverage (operating cash flow to total debt)  
• Accruals  
• ROA-GARP                                                                          |
| BNP Paribas Quality Index (2017)            | • ROE (accounts for 20%)  
• Free cash flow to total asset (30%)  
• Gross profitability to book value (20%)  
• Accruals (30%)                                                                    |
| GMO (2004)                                  | • Profitability  
• Leverage  
• Earnings volatility                                                                |
| Cambridge Trust Company (2015)              | • ROE  
• Stable or increasing profit margins, preferably from an already-high base  
• Stable or increasing balance sheet capacity  
• Free cash flow generation                                                          |

Table 1: Sample of quality definition from industry
4 Empirical part

4.1 Survey of practitioners

The first step in the development of a generally accepted definition consists in collecting feedbacks from practitioners. To do so, a survey has been developed on Qualtrics and submitted to a panel of rigorously selected practitioners occupying positions close to the investment decisions and the equity research. Globally, the panel is mainly made up of practitioners occupying Portfolio Managers and Equity Research Analysts positions. Out of approximately 75 respondents, 51 have been retained as sufficiently relevant for further processing. By and large, these respondents come from 6 countries (mainly Belgium and Luxemburg but also UK, France, Switzerland and Austria) and represent no less than 30 different institutions (mainly asset management firms, banks and independent equity research providers).

Basically, the survey was addressed to the practitioners with the objective of collecting the three following outputs:

1. The dimensions to include in the quality definition;
2. The measures capturing the underlying dimension;
3. The relative weights attributed to those dimensions;

Before analysing the results, it is worth mentioning that the main quality definition derived from the survey has been constructed by only considering the quality dimensions which are selected for inclusion by a large majority of respondents. Hence, one can see in the following section that the retained dimensions are supported by minimum 77% and up to 92% of respondents. Contrary to the selection of Quality dimensions which is simply subject to a binary choice, the probability that opinions converge to a large extent to a single underlying measure is far lower given the variety of measures submitted for selection. Therefore, in case the most solicited measure of a given dimension does not reach a significant majority, one or more measures are added in order to reflect the dispersion in views while ensuring that a significant majority of respondents supports at least one of the measures finally selected for inclusion. One
can see, in section 4.2, that the sole dimension which finally required such adjustment is the Growth dimension.

Globally, given the large variety of Quality definitions existing among academicians and practitioners, one can admit that, if each dimension included in the definition is accepted and advocated by at least 72% of the population and their measures supported by at least 67% (excepted for the higher indecisiveness relative to the Growth dimension), the resulting quality definition takes a step forward in reconciling the diversity of views.

### 4.2 Analysis of survey results

#### 4.2.1 Identification of sub-factors valued by practitioners

In order to identify the quality dimensions that matter in the industry, the practitioners have first been asked to select the dimensions they would include in their own quality definition. Basically, the six dimensions previously discussed in chapter 3 are proposed; that is, the four dimensions derived from Asness et al. (2013) – i.e. Profitability, Growth, Safety and Payout – plus Earnings Quality (proxied by accruals) and R&D intensity (proxied by price-to-research ratio).

Among the six dimensions under study, four are highly valued by practitioners and appear as quasi-mandatory for the inclusion in the quality definition. These are Profitability (92% of the panel), Safety (92%), Earnings Quality (87%) and Growth which still collects 77%. On the other hand, the R&D Intensity dimension (50%) displays a mitigated success while the Payout (33%) is even worse. The latter result appears rather surprising; indeed, one would expect that more than one third of the investors pay attention to the Payout when considering a quality investment. Nevertheless, this observation from the industry is cross-checked with the last draft of the “Quality Minus Junk” of Asness et al. (2017). In this last version of the model, the authors remove the Payout dimension and only treat the Profitability, Growth and Safety dimensions as detailed in section 3.2.3. The reasoning behind this removal is that the payout ratio does not affect price in their quality model since they consider a frictionless economy in which Modigliani-Miller theorem holds (Asness et al., 2017). Overall, the dimensions which are selected for inclusion in the Quality definition are Profitability, Earnings Quality, Safety and Growth due to the large majority of respondents supporting them.
4.2.2 Identification of sub-factors valued by practitioners

Once a given dimension is considered relevant and selected for inclusion in the quality definition, practitioners are asked to select 1 or 2 measures that, according to them, best capture the associated dimension. Since practitioners are likely not to adhere to the ratios suggested by academicians, ratios selection was made visible if and only if the underlying sub-factor is selected. That way, we avoid the case where the respondent rejects a dimension that he would have normally chosen if not restricted to academic measures.

Globally, most of the measures submitted for selection are derived from the quality definition of Asness et al. (2013). Indeed, since this quality model is poorly parsimonious, it offers a good basis for measures selection. However, a series of adjustments are performed to make the survey even more exhaustive and well in line with other academic research. Hence, the ROIC advocated by Greenblatt (2006) is added in the profitability measures. Indeed, this ratio represents an interesting alternative to the ROE and ROA, and has been very influential in recent years. Growth metrics are maintained but the time window has been shortened from 5 years to 3 years in order to reduce the amount of data lost due to growth computation (see section 4.4). Low accruals which appears as a profitability measure in the model of Asness et al. (2013) is requalified under the Earnings quality dimension. The reason being that accruals is a measure of cash generation and has probably been incorporated into the profitability measure to fit one of the four dimensions derived from Gordon’s growth model. Finally, the dividend yield, whose unexplained absence in the model is quite surprising, has been added to the Payout measures. Indeed, the dividend yield is undoubtedly one the most popular measure of Payout and is supported by both academicians and practitioners implementing an investment.
strategy referred to as Relative Dividend Yield strategy. Overall, all the measures are synthetized in Table 2 below:

<table>
<thead>
<tr>
<th>Profitability</th>
<th>Growth</th>
<th>Safety</th>
<th>Earnings Quality</th>
<th>R&amp;D Intensity</th>
<th>Payout</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPOA</td>
<td>∆GPOA</td>
<td>Low Beta</td>
<td>ACC</td>
<td>PRR</td>
<td>EISS</td>
</tr>
<tr>
<td>ROA</td>
<td>∆ROA</td>
<td>Low IVOL</td>
<td>ACC</td>
<td>PRR</td>
<td>DISS</td>
</tr>
<tr>
<td>ROE</td>
<td>∆ROE</td>
<td>Low Leverage</td>
<td>ACC</td>
<td>PRR</td>
<td>EISS</td>
</tr>
<tr>
<td>CFOA</td>
<td>∆CFOA</td>
<td>O-Score</td>
<td>ACC</td>
<td>PRR</td>
<td>EISS</td>
</tr>
<tr>
<td>GMAR</td>
<td>∆GMAR</td>
<td>Z-Score</td>
<td>ACC</td>
<td>PRR</td>
<td>EISS</td>
</tr>
<tr>
<td>ROIC</td>
<td>∆ACC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: List of Quality dimensions and underlying measures

4.1.2.1 Profitability measures selection

When it comes to profitability (see Figure 3), ROIC reaches an interesting majority and is therefore selected as the measure of Profitability. It is interesting to note that ROIC gets twice as many votes as ROE. Investors have thus a marked preference for the alternative proposed by Greenblatt (2006), which allows to reduce distortions due to the level of debt and the applicable tax rate by using EBIT rather than net income, and to take account of all the capital needed to conduct the firm’s business (see section 3.1.1.2). ROE (33%) and ROA (15%) collects lower results partly due to the preference of practitioners for ROIC. For their part, top line profitability measures are struggling to convince investors. Indeed, GMAR only reaches 21% while Novy-Marx’s GPOA is selected by one single respondent. Hence, whereas GPOA is largely recognized in the literature for its relevance and interesting results, practitioners do not seem to realize the benefits of this measure.
With the almost unanimous support of the panel, the leverage (76%) stands out as the safety measure to include in the quality score (see Figure 4). This result is not so surprising given the popularity of the ratio and its utmost importance in credit analysis. On the other hand, Altman’s Z-score (11%) and Ohlson’s O-score (4%) obtain poor results. This may be explained by their higher complexity, the fact they encompass a series of ratios that overlap other dimensions or simply the fact they compete with the very popular leverage. Low idiosyncratic volatility (17%) and Beta (28%) do slightly better but are still far from competing with the low leverage measure.

While profitability and safety show a clear trend towards a single underlying measure, results are more contrasted when it comes to the growth dimension (see Figure 5). Even so, the ΔCFOA (51%) breaks away and emphasizes the importance of growing cash generation for investors. Astonishingly, the ΔROE does not really benefit from the absence of ΔROIC in the measures submitted. On the contrary, top line measures like GMAR (31%) and especially GPOA (21%)
are more valued in the growth context. Since CFOA does not reach an obvious consensus, GMAR is added as a second measure for growth, leading the cumulated percentage to 74% (82% - 8% overlap), meaning that 74% of the respondents are satisfied with at least one of the two measures. Interestingly, we have here one top line (GMAR) as well as one bottom line (CFOA) measures to capture growth.

Regarding Earnings quality, given the preponderance of Sloan’s accruals theory in the quality investing literature, it has been decided to submit only the ACC measure in the framework of the survey. Therefore, when respondents have been offered the choice to select Earnings quality for inclusion in their quality definition, they were explicitly informed that the dimension would be captured by ACC.

Even though the small percentage of votes obtained by the Payout dimension (see Figure 6) does not allow it to integrate the quality definition, it is worth mentioning that the dividend yield (65%), missing in the model of Asness et al. (2013), is actually by far the most valued measure of Payout. On the other hand, debt issuance (6%) and equity issuance (0%), whose link with Payout is eventually less straightforward than for the dividend yield, appear almost irrelevant. Like GPOA in the Profitability dimension, practitioners do not seem to be ready or sufficiently informed to shift their definition towards purely academic Quality definition such as external financing.
4.2.3 Weighting of the different dimensions

In an attempt to understand the relative importance attributed by practitioners to each dimension, they have been asked to distribute 100 points among the aforementioned dimensions; that is, Profitability, Growth, Earnings Quality, Safety/Credit Risk, R&D intensity and Payout. Without surprise, the four dimensions which are selected for inclusion within the quality definition – i.e. Profitability (26.9%), Earnings Quality (21.4%), Growth (20.3%) and Safety (16.7%) – collect the higher weights while Payout and R&D Intensity lag behind. One can note that the Safety dimension that is almost unanimously selected for inclusion (92%) collect the lowest weight among the selected dimensions. This may suggest that practitioners cannot conceive that a quality stock exhibits poor creditworthiness but do not expect that high creditworthiness would be achieved at the expense of profitability and growth prospects.
4.3 Determination of scores

4.3.1 Quality scores

The analysis of practitioners’ feedbacks regarding relevant dimensions and measures allows to develop a first definition of a quality stock. Therefore, according to practitioners, a quality stock is a stock that performs well on this set of characteristics:

- Profitability measured by ROIC;
- Safety measured by LEV;
- Earnings Quality measured by ACC;
- Growth measured by CFOA and GMAR.

The dimensions and associated measures being known, the next step consists in quantifying this quality in order to make quality comparisons between stocks possible. To determine the quality score associated with the aforementioned quality definition, the same methodology as Asness et al. (2013) has been used. The latter consists in combining the z-scores of the different measures included in the quality definition. It appears important to specify that the z-score computed here has nothing whatsoever to do with Altman’s Z-score. To avoid any confusion, a lower case z will be used.

Calculating z-scores is actually a widely used method of standardizing a variable in order to combine it with other variables which may have a different scale or a different unit of measurement. Since it has a mean value of 0 and a standard deviation of 1, the value of the z-score reflects how many standard deviations a given value lies from the mean and can be both positive and negative. The z-score is computed as follows:

\[ z_\alpha = \frac{x_\alpha - \mu}{\sigma} \quad (4.1) \]

Where:

- \( z_\alpha \) denotes the z-score for a given security;
- \( x_\alpha \) denotes the variable under consideration for a given security;
- \( \mu \) denotes the mean of the variable in a given universe;
- \( \sigma \) denotes the standard deviation of the variable in a given universe.

The methodology is applied the following way:
1) First of all, each variable – i.e. ROIC, LEV, ACC, ΔCFOA, ΔGMAR – is converted into ranks.

2) Then, a z-score is computed for each variable using ranks. Consequently, replacing in the formula (4.1), \( x_\alpha \) becomes the rank of the variable for a given security, \( \mu \) the mean of the ranks and \( \sigma \) the standard deviation of the ranks. In practice, one can observe that the conversion in ranks is often use with the aim of winsorizing⁵ the outlier variable values and thus ensuring that the mean value is less affected by extreme values. This winsorizing method is used in both MSCI and S&P quality indices methodology. By computing z-scores on ranks, even the distance between variables values is standardized.

3) The score of each dimension is computed. If the dimension is measured by a single metric, its score is equal to the z-score of this metric. Else, the score of the dimension is equal to the average of the individual z-scores. Overall, the score of each dimension is computed as follows:

   a. Profitability 1 = \( z_{ROIC} \)
   b. Safety = \( z_{LEV} \)
   c. Earnings Quality = \( z_{ACC} \)
   d. Growth 1 = \( \frac{(z_{\Delta CFOA} + z_{\Delta GMAR})}{2} \)

4) The quality score is finally computed by averaging the scores of the four dimensions:

\[
Quality 1 = \frac{Profitability1 + Safety + Earnings Quality + Growth1}{4} \tag{4.2}
\]

In order to compare the performance of an investment strategy based on this quality definition with some alternatives, two additional definitions are proposed:

The second quality definition is developed with the aim of analysing how a larger number of measures impacts the performance of a quality investment strategy based on its implied quality score in comparison with the main quality score – i.e. Quality 1. This definition is therefore less parsimonious but also less generally accepted since additional measures, which have been selected for inclusion, are less agreed by practitioners. To do so, two profitability measures are added to ROIC: CFOA to have a focus on cash generation and GMAR to have a top-line profitability measure. Regarding the growth dimension, ROE is also added as a third measure. Hence, like the profitability dimension, growth is measured by a top-line metric (ΔGMAR),

⁵ Winsorization is the process of transforming statistics by reducing the impact of extreme data values.
one focusing on cash flow generation (ΔCFOA) and another acting as an alternative to ROIC (ΔROE). In parallel, safety and earnings quality measures are kept identical due to their high preponderance in the survey. Overall, the score of each dimension is computed in the following way:

\[ \text{Profitability 2} = \frac{(z_{ROIC} + z_{CFOA} + z_{GMAR})}{3} \]

\[ \text{Safety} = z_{LEV} \]

\[ \text{Earnings Quality} = z_{ACC} \]

\[ \text{Growth 2} = \frac{(z_{CFOA} + z_{GMAR} + z_{ROE})}{3} \]

The resulting quality score is computed in the same way as Quality 1, by averaging the scores of the four dimensions:

\[ \text{Quality 2} = \frac{\text{Profitability 2} + \text{Safety} + \text{Earnings Quality} + \text{Growth 2}}{4} \]  \hspace{1cm} (4.3)

The third quality definition is simply a weighted version of Quality 1 using the weights attributed by the practitioners to the different dimensions (see section 4.1.3). Since the Payout and R&D intensity dimensions are not considered in the quality definition, the weights attributed to the four other dimensions are increased to be expressed on a 100% basis. This last quality score is computed as follows:

\[ \text{Quality 3} = 0.315 \times \text{Profitability 1} + 0.167 \times \text{Safety} + 0.251 \times \text{Earnings Quality} + 0.238 \times \text{Growth 1} \]  \hspace{1cm} (4.4)

\subsection*{4.3.2 QARP score}

In order to analyse the benefits of using the quality definition developed to condition a value strategy, a QARP score is determined following a methodology similar to the one used for quality scores. In order to capture the value component of QARP, BM ratio is used. Here, the value component of QARP consists of the z-score of the BM ratio calculated on ranks:

\[ \text{Value} = z_{BM} \]  \hspace{1cm} (4.5)

The QARP score is then computed as the average of Quality 1 and Value:

\[ \text{QARP 1} = \frac{\text{Quality 1} + \text{Value}}{2} \]  \hspace{1cm} (4.6)
4.4 Portfolio construction and analysis methodology

The methodology used to construct portfolios based on the previously developed Quality definitions follows the method used by Fama and French (1992, 1993) and focus on U.S. stocks from the New York Stock Exchange (NYSE). Subsample of the CRSP/Compustat merged database spanning from July 2001 to December 2015. This database provides the historical matching of CRSP (Center for Research in Security Prices) stock data with Compustat fundamental data. In order to process data and simulate portfolios, the database is imported into SAS Studio University Edition.

The first step consists in calculating all the ratios which are selected by practitioners in the framework of the survey. The definition of each ratio, using CRSP/Compustat data items, is provided in Appendix 7.2. Once computed and added to the database, ratios are further processed and combined to obtain quality dimensions scores and final quality scores as detailed in section 4.3. It is worth mentioning that, given the period covered by the data sample (2001-2015), the choice has been made to compute growth measures over a 3-year timeframe instead of the 5-year window proposed by Asness et al. (2013) in order to reduce the amount of data lost because of growth measures computation by two years. Besides, since ∆CFOA requires one more year for the computation of the change in working capital, four years are finally truncated from the subsample leading to a final sample spanning from July 2005 to December 2015.

Then the quality scores and the underlying dimension scores and measures are used to replicate Fama and French portfolio construction using SAS open source codes from Wharton Research Data Services (Palacios & Vora, 2009) as a basis. Hence, for the case of a Quality score, the corresponding QMJ factor is constructed using 6 value-weighted portfolios formed on Size and Quality – i.e. Small Quality (SQ), Small Medium Quality (SM), Small Junk (SJ), Big Quality (BQ), Big Medium Quality (BM) and Big Junk (BJ). The size breakpoint is the median NYSE market for U.S. securities and the quality breakpoints are the 30th and 70th percentile. The value-weighted resulting portfolio, QMJ, hence holds (shorts) stocks in the top (bottom) 30% by quality rank. Besides, portfolios are rebalanced each year at the end of June, using accounting data from the fiscal year ending in the previous calendar year. More precisely, QMJ factor is computed as follows:
\[ QMJ = \frac{1}{2} (SQ + BQ) - \frac{1}{2} (SJ + BJ) \] (4.7)

In addition to the computation of QMJ 1, QMJ 2, QMJ 3 and QARP factors, separate sub-portfolios based on the four components of Quality as well as individual measures are constructed similarly.

Regarding the portfolio performance analysis, basic statistics, tests for location and normality, correlation and regression are performed using SAS Studio while Sharpe ratios, Sortino ratios, value at risk, expected shortfalls and maximum drawdown are computed on Microsoft Excel. In order to compute regression on Fama and French three-factor, Carhart four-factor and Fama and French five-factor models, SMB, HML, UMD, RMW, CMA, Rm-Rf (market factor) and risk free rate are retrieved from Kenneth R. French Data Library. Furthermore, the data relative to the QMJ factor of Asness et al. (2013) are retrieved from AQR library.

### 4.5 Performance measures

Before proceeding any further, it might be appropriate to specify the risk statistics (i.e. value at risk, expected shortfall and maximum drawdown) and risk-adjusted statistics (i.e. Sharpe ratio and Sortino ratio) used for the assessment and comparison of portfolios.

#### 4.5.1 Sharpe ratio

The Sharpe ratio is one of the most popular measures of the risk-adjusted return which is computed as the average return earned in excess of the risk-free rate per unit of volatility (captured by the standard deviation).

\[
\text{Sharpe ratio} = \frac{\text{Expected portfolio return} - \text{Risk free rate}}{\text{Portfolio standard deviation}}
\] (4.8)

#### 4.5.2 Sortino ratio

The Sortino ratio is another risk-adjusted statistic that serves as an interesting alternative to the Sharpe ratio which does not distinguish between downside and upside volatility and thus
penalizes both equally. Hence, considering a positively skewed distribution, the high outliers, which are in fact welcomed by the investors, increase the denominator to a greater extent than the numerator leading to a lower Sharpe ratio. In this case, the performance is in fact achieved with less risk than the Sharpe ratio would lead to expect. The rationale behind the Sortino ratio is that the upside volatility is beneficial for the investment and should therefore not be included in the risk calculation. As a consequence, the Sortino ratio only uses the downside volatility in the denominator instead of the total standard deviation.

\[ \text{Sortino ratio} = \frac{\text{Expected portfolio return} - \text{Risk free rate}}{\text{Portfolio downside deviation}} \]  

(4.9)

Where the portfolio downside deviation is computed as follows:

\[ \text{Downside deviation} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\text{Min}(0, x_i - rf_i))^2} \]  

(4.10)

Where:

- \( N \) denotes the total number of returns;
- \( x_i \) denotes the \( i^{th} \) return;
- \( rf_i \) denotes the risk free rate applicable at the \( i^{th} \) period.

### 4.5.3 Value at risk

Value at risk (VaR) measures the amount of potential loss that could happen in a portfolio over a given time period. Hence, there are three significant parts to VaR:

- A confidence level which is typically fixed at 95 or 99%;
- A time period (monthly in the case of this study);
- A potential loss amount (eventually expressed as a loss percentage) related to the aforementioned timeframe and confidence level.

For instance, if a portfolio has a one-month 3.5% VaR at a confidence level of 0.99, it means that there is a 0.01 probability that the portfolio will fall in value by 3.5% or more over a one-month period.

In the framework of this thesis, two methods are used to compute the value at risk. The first one, also referred to as historical method, uses the actual cumulative distribution function of
historical returns data and then assumes that history will repeat itself (see Figure 8 – left-hand side). On the other hand, the variance-covariance method makes the assumption that returns are normally distributed and, hence, requires only two inputs – i.e. the mean return and the standard deviation (see Figure 8 – right-hand side).

Figure 8: Illustration of VaR - Historical method (left) vs. Variance-Covariance method (right)

Source: Investopedia

4.5.4 Expected shortfall

The expected shortfall, also referred to as conditional value at risk, is a risk measure that assesses the tail end of the distribution of loss and thus serves as an extension of value at risk. More specifically, the expected shortfall represents the expected loss on the basis of the x% worst occurrences. In the framework of the thesis, a 5% level is fixed so that the expected shortfall calculates the expected return on the portfolio in the worst 5% cases.

4.5.5 Maximum drawdown

The maximum drawdown is a measure of downside risk over a given timeframe. It measures the maximum loss from a peak to a trough of the portfolio returns over a specified time period and is expressed as a percentage of the peak value.

\[
\text{Maximum drawdown} = \frac{\text{Trough value} - \text{Peak value}}{\text{Peak value}}
\]
4.6 QMJ portfolio analysis

4.6.1 Analysis and comparison of QMJ 1 vs. QMJ 2 and QMJ 3

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
<th>Skew</th>
<th>Kurt</th>
<th>Normality test p-value</th>
<th>t-stat p-value</th>
<th>Sign test p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>QMJ 1</td>
<td>126</td>
<td>-0.045528</td>
<td>1.42834</td>
<td>-4.40368</td>
<td>-0.11842</td>
<td>3.52558</td>
<td>0.04215</td>
<td>0.56641</td>
<td>0.7590</td>
<td>0.72228</td>
<td>0.5331</td>
</tr>
<tr>
<td>QMJ 2</td>
<td>126</td>
<td>-0.057972</td>
<td>1.89014</td>
<td>-6.94315</td>
<td>-0.12185</td>
<td>5.71114</td>
<td>0.12239</td>
<td>2.07355</td>
<td>0.0013</td>
<td>0.73122</td>
<td>0.5331</td>
</tr>
<tr>
<td>QMJ 3</td>
<td>126</td>
<td>-0.035917</td>
<td>1.32469</td>
<td>-3.16958</td>
<td>-0.11831</td>
<td>3.71849</td>
<td>0.15587</td>
<td>0.20306</td>
<td>0.5476</td>
<td>0.76137</td>
<td>0.4228</td>
</tr>
</tbody>
</table>

Table 3: Basic statistical measures, normality test and tests for location on QMJ 1-2-3

Using the UNIVARIATE procedure on SAS, basic statistical measures and tests for location are computed (see Table 3) with the monthly returns of each long-short QMJ portfolios derived from the three Quality definitions developed in section 4.3.1. Since literature argues that high-quality stocks tend to outperform low-quality stocks, one should expect that these portfolios earn average abnormal returns. On the contrary, statistics do not go in that direction. Indeed, the monthly returns of the QMJ 1 portfolio, derived from our main quality definition (i.e. Quality 1), exhibits a negative mean (-0.045). The Student’s t-test is used to test the null hypothesis that the mean of the returns equals 0. Since the p-value (0.72) associated is larger than the 0.05 level of significance, the null hypothesis cannot be rejected. In other words, the QMJ 1 portfolio exhibits an average return which is not significantly different from 0 and one can observe that its cumulative returns over the decade do not lead to a clear trend (see Figure 9). This statement is also true for the QMJ 2 and QMJ 3 portfolios which exhibits p-values of 0.53 (tested with nonparametric Sign test due to non-normality of returns) and 0.76 respectively.

As is reasonably logical to expect, the three portfolios are highly correlated (see Appendix 7.3). Indeed, QMJ 1 and QMJ 3, whose underlying quality definitions vary only in terms of measures weighting, have a Pearson correlation coefficient of 0.95 (adjusted R-square of 0.91). Despite the fact that QMJ 2 includes additional measures in its quality definition, the correlation with QMJ 1 remains significant with a Pearson correlation coefficient of 0.91 (adjusted R-square of 0.83).
In terms of risk-adjusted returns, each portfolio exhibits both negative Sharpe and Sortino ratios, meaning that their respective returns are lower than the risk-free rate (see Table 4). From a risk perspective, one can note that the additional measures of Quality 2 make QMJ 2 much more risky than QMJ 1 and QMJ 3. Indeed, its standard deviation (1.89), VaR and expected shortfall (-4.26%) are well above the levels of QMJ 1 and QMJ 3. Besides, its max drawdown (24.8%) is almost twice as high as QMJ 3 (14.3%). Finally, one can observe that risk measure values of QMJ 1 and QMJ 3 lie in a similar range.

These results show that the adjustments made to the main quality definition in an attempt to eventually observe different returns patterns and trends afterwards are not sufficient. Indeed, one can observe differences in their risk profile but all of them still generate negative returns and have negative Sharpe ratios over the sample period. For this reason, the focus will be primarily put on the main quality definition (Quality 1) and its implied long-short portfolio (QMJ 1).

When it comes to regressing QMJ 1 on the CAPM market factor, Fama and French 3-and-5-factor and Carhart 4-factor (see Table 5), one can first notice that, whatever the model used, QMJ 1 do not generate any significant alpha. Furthermore, QMJ 1 is positively correlated with...
RMW and UMD though not loading significantly on the latter (see Appendix 7.4). This suggests that QMJ 1 is long company that are more profitable. On the other hand, QMJ 1 is negatively correlated with Mkt-rf, SMB, HML and CMA. This suggests that high-quality stocks, in the sense of Quality 1 definition, tend to have lower volatility, and to be larger, more expensive and more aggressive in terms of investments.

4.6.2 Comparison with academic definition: GPOA

In order to compare the performance of QMJ 1 with an academic definition, a factor has been created in the same way using GPOA. Since Novy-Marx (2013) argues that the gross profitability is the notion of Quality that contributes the most to investment performance, GPOA appears as an interesting reference point. At first sight, one can state that GPOA displays a much more interesting pattern regarding cumulative returns over the sample period (see Figure 9). Indeed, while QMJ 1 generates a raw return of roughly -7% over the period, GPOA yields no less than 62%.

![Figure 9: Cumulative returns of QMJ 1 and GPOA](image)

The performance measures argue even more for GPOA. Indeed, thanks to a positive monthly return mean of 0.4%, GPOA exhibits both positive Sharpe and Sortino ratios of 0.17 (0.59 when annualized) and 0.30 respectively contrary to QMJ 1 which exhibits negative values (see Table 6). GPOA exhibits a higher standard deviation (1.77 vs. 1.43) but the VaR increases marginally while the expected shortfall (-3.06%) and maximum drawdown (12.1%) are reduced.
<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std Dev</th>
<th>Sharpe ratio</th>
<th>Sortino ratio</th>
<th>VaR (VC) 95%</th>
<th>VaR (Historical) 99%</th>
<th>Expected shortfall (5%)</th>
<th>Max drawdown</th>
</tr>
</thead>
<tbody>
<tr>
<td>QMJ 1</td>
<td>-0,045</td>
<td>1,43</td>
<td>-0,1</td>
<td>-0,14</td>
<td>-2,39%</td>
<td>-3,37%</td>
<td>-2,28%</td>
<td>-3,11%</td>
</tr>
<tr>
<td>GPOA</td>
<td>0,4</td>
<td>1,77</td>
<td>0,17</td>
<td>0,3</td>
<td>-2,51%</td>
<td>-3,72%</td>
<td>-2,48%</td>
<td>-3,06%</td>
</tr>
</tbody>
</table>

**Table 6: Performance of QMJ 1 vs. GPOA**

Besides, as illustrated in Table 7, GPOA generates significantly positive alphas whatever the asset pricing model used for regression, which QMJ 1 manifestly fails to achieve.

<table>
<thead>
<tr>
<th></th>
<th>α</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAPM alpha</td>
<td>0,0048</td>
<td>3,18</td>
</tr>
<tr>
<td>Fama and French 3-factor alpha</td>
<td>0,0038</td>
<td>3,09</td>
</tr>
<tr>
<td>Carhart 4-factor alpha</td>
<td>0,0039</td>
<td>3,12</td>
</tr>
<tr>
<td>Fama and French 5-factor alpha</td>
<td>0,0027</td>
<td>2,23</td>
</tr>
</tbody>
</table>

**Table 7: Alphas generated by GPOA**

Interestingly, one can note that these results converge towards the findings documented by Kyosev et al. (2016) stated in section 3.3. As a reminder, the authors address the large performance differences between the different quality definitions and argue that academic definitions for quality seems to have significant predictive power while beyond common factors, they do not find significant predictive power for individual industry definitions. Moreover, these results also highlight the fact that, even when shifting practitioners towards more academic quality measures, they tend to build a quality definition providing poor predictive power.

**4.6.3 Analysis of individual dimension of QMJ 1**

In order to assess whether any individual dimension of Quality 1 displays significant predictive power, regressions on the different factors are performed for each dimension, just as in the case of QMJ 1 and GPOA in the previous sections. Moreover, in order to briefly compare the performance of the measures selected to capture the associated Quality dimensions, long-short portfolios have been created on the alternative measures. Their cumulative returns and basic statistics are provided in Appendices 7.5 and 7.6.

One can first observe that neither dimension succeeds in generating alpha (see Figure 10). Besides, one can note that the measures selected by practitioners exhibit poor figures relative to their alternatives. Hence, from the Profitability perspective, ROIC exhibits negative mean...
return and lies far behind the performance of GPOA. Moreover, it is interesting to note the significant difference in returns between GPOA and GMAR which yet are both related to gross profit but not scaled by the same variable – i.e. total assets for GPOA and sales for GMAR.

Regarding Safety, the Altman’s Z-score would have made much better than the leverage with, namely, an average return of 0.27% vs. 0.03%. Finally, the same observation can be made for Growth 1 (containing ΔGMAR and ΔCFOA) which is penalized by the poor performance of the 3-year growth in GMAR. Nevertheless, the Growth measures are still more concentrated than in the case of Profitability.

<table>
<thead>
<tr>
<th></th>
<th>ROIC</th>
<th>LEV</th>
<th>ACC</th>
<th>Growth 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CAPM alpha</strong></td>
<td>α</td>
<td>t-stat</td>
<td>α</td>
<td>t-stat</td>
</tr>
<tr>
<td></td>
<td>0,0006</td>
<td>0,79</td>
<td>0,0015</td>
<td>1,02</td>
</tr>
<tr>
<td><strong>Fama and French 3-factor alpha</strong></td>
<td>0,0005</td>
<td>0,66</td>
<td>0,0008</td>
<td>0,62</td>
</tr>
<tr>
<td><strong>Carhart 4-factor alpha</strong></td>
<td>0,0004</td>
<td>0,61</td>
<td>0,0006</td>
<td>0,45</td>
</tr>
<tr>
<td><strong>Fama and French 5-factor alpha</strong></td>
<td>-0,0001</td>
<td>-0,12</td>
<td>0,0015</td>
<td>1,07</td>
</tr>
</tbody>
</table>

Figure 10: Alphas generated by ROIC, LEV, ACC and Growth 1

4.6.4 Size premium controlling for QMJ 1

As seen in section 3.2.3, the Size premium tends to revive when controlling for junk on the basis of the QMJ factor of Asness et al. (2013). It is thus interesting to assess the extent to which QMJ 1 can make the Size premium stronger. Table 8 shows the regression of the size factor (SMB) on the Fama and French factors (i.e. Mkt-rf, HML and UMD) and controlling for Quality (QMJ 1). The first row displays results of QMJ regressed on the market portfolio over the sample period (July 2005 – December 2015). The alpha from the regression is -3 basis points but is not significantly different from 0, meaning that the CAPM explains the returns to SMB quite well. In the second and third rows, one can note that adding HML and UMD factors does not change the significance of the alpha. In other words, there does not appear to be a reliable Size premium in the presence of these factors (Mkt-rf, HML and UMD). However, contrary to expectations, the alpha does not turn significant when controlling for quality with the QMJ 1 factor. However, one can note that SMB loads significantly and negatively on QMJ 1 and that the R-square is marginally improved. This result is in line with the observation that Quality companies tend to be large and vice versa.
Table 8: Size premium controlling for junk

<table>
<thead>
<tr>
<th></th>
<th>α</th>
<th>β</th>
<th>h</th>
<th>m</th>
<th>q</th>
<th>R-square</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0,0003</td>
<td>0,18</td>
<td>0,02</td>
<td>0</td>
<td>0,03</td>
<td>-0,43</td>
<td>0,17</td>
</tr>
<tr>
<td>(-0,14)</td>
<td>(3,89)</td>
<td>(0,22)</td>
<td>(0,41)</td>
<td>(0,57)</td>
<td>(-2,6)</td>
<td></td>
</tr>
<tr>
<td>-0,0002</td>
<td>0,16</td>
<td>0,03</td>
<td>0</td>
<td>0,04</td>
<td>-0,43</td>
<td>0,13</td>
</tr>
<tr>
<td>(-0,11)</td>
<td>(3,91)</td>
<td>(0,41)</td>
<td>(0,45)</td>
<td>(-2,6)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In order to compare the impact of QMJ 1 on the Size premium with the one of Asness QMJ, the same regression has been performed using QMJ instead of QMJ 1 (see Appendix 7.7). One can observe that, though increasing to 10 basis points, the alpha still remains insignificant. Therefore, one may assume that the sample period makes it difficult to revive the Size premium.

4.6.5 Recession vs expansion periods (NBER)

Although the proposed QMJ 1 portfolio does not show the abnormal returns portrayed by academic definitions, it still displays interesting patterns when looking chronologically at returns (see Figure 11). Indeed, the cumulative return is not flat and experiences an upward trend during the last US business cycle contraction period implied by the last financial crisis and spanning from December 2007 to June 2009 as stated by the National Bureau of Economic Research (NBER), which establishes recession periods through recent decades based on significant decline in economic activity spread across the economy lasting more than a few months, visible in real GDP, real income, employment, industrial production, and wholesale-retail sales.

![Figure 11: Cumulative return of QMJ 1 portfolio](image-url)
When considering separate statistics for both type of economic states, the mean return of expansion periods turns out to be positive and amounts to 0.4673 while de the mean return of contraction periods decreases to -0.1363. While these observations are logically in line with the flight to quality effect documented in section 2.3.3, the group t-test used to compare the means of both groups does not exhibit a statistically significant difference (see Appendix 7.8). However, this statistical significance is complicated by the small size of the recession sample containing only 19 observations. One may thus assume that the Quality 1 definition encompasses safety features that appeal investors during financial market distress. During the contraction period, it is interesting to note the improvement of Sharpe and Sortino ratios which turn positive and reach 0.19 and 0.30 respectively.

A similar upward trend can be observed between February 2011 and September 2011. Even though NBER does not consider any contraction period in the United States at that time, one may assume that the raging European sovereign debt crisis has impacted the U.S. markets. Indeed, the end of this time period is marked, in August 2011, by a sharp drop in stock prices due to fears of contagion of the Eurozone debt crisis and concerns over the slow economic growth of the United States leading to US credit rating downgrade from AAA to AA+.

### 4.7 QARP long-short portfolio analysis

Using the QARP definition presented in section 4.3.2 and the same factor construction methodology as QMJ portfolios, a long-short QARP portfolio is build. That is, a portfolio which evenly combines both quality and value factors by being long stocks which are cheap relative to their quality and short stocks which are expensive relative to their quality.

Before analysing the performance of this portfolio, it is worth noting that the performance of the value factor has been particularly disappointing during the period studied (2005-2015). Figure 12 displays, among others, the cumulative returns of the Fama French US HML over the studied time window based on data retrieved from Kenneth R. French Data Library. Even though the value factor has been one of the most resilient and well accepted factors, one can observe its decline over that decade in the U.S. Nevertheless, this lost decade does not mean that the value factor premium has disappeared.
It is true that some researchers question the sustainability of factor premiums. Indeed, one can fairly assume that the growing awareness of a factor may result in more and more investors embracing the strategy which in turn would push up stocks prices in the most attractive segment of the factor, so that future returns would go down and the factor premium would eventually disappear. Nevertheless, it is worth remembering the risk and behavioural (errors, mispricing, underreaction, overreaction, etc.) drivers of factor returns (see section 2.2) which are not mutually exclusive and limit this convergence phenomenon. For instance, the price of risk is likely to vary through time and eventually likely to fall as the risk premium is more popularized, but if factor return is, at least partly, rational compensation for risk then there is no reason it should ever completely disappear or necessarily fall below a rational level (AQR, 2015). Besides, this convergence of factor spreads is further limited by a series of observations such as the fact that all institutional investors are not able or willing to engage in factor investing strategies or the fact that some factors go against others. In conclusion, as they become more popular, many strategies make a journey to a middle ground – i.e. a place where they are known and still work though perhaps not at the same level as in the past (AQR, 2015).

In the light of these explanations and given the fact that factors are likely to experience significant cyclicalitity over short periods of time (see section 2.3.4), one may attribute the poor performance of the value factor, between 2005 and 2015, to a cyclical low. Therefore, the evolution of the value factor over the past decade reminds us that factors can experience long periods of underperformance and that the ability to fully capture the risk premium associated with factors requires either extraordinary timing or the ability to remain fully exposed through the whole cycle (MSCI, 2016).

![Figure 12: Performance of the value (Fama French US HML). QARP and QMJ 1 factors](image_url)
As stated in section 2.3.4, by conditioning a value investing strategy on the quality of the firms, QARP is likely to capture the best of both worlds while achieving significant diversification benefits. As could be expected, Quality 1 and Value exhibit significantly negative correlation with a Pearson correlation coefficient of -0.51. This suggests that high-quality stocks in the sense of Quality 1 definition tend to be expensive and vice versa. Despite this, QARP does not appear more convincing during the period studied than Quality 1 taken alone. Indeed, with an even worse average return (-0.12% vs. -0.05%), dramatically higher volatility (2.31 vs. 1.43), both negative Sharpe and Sortino ratios, much higher VaR, expected shortfall (-5.35% vs. -3.11%) and maximum drawdown (26.0% vs. 15.5%) (see Table 9), the sole benefit that can be attributed to QARP is to have experienced positive cumulative return much more often, though finally closing the period studied with lower figures (-15.5% vs. -6.8%) (see Figure 12). Yet, QARP makes somewhat better than Value by, namely, slightly improving risk measures (see Table 9). While the weighting of both Quality and Value in a QARP strategy are likely to provide significantly different performances – e.g. Asness et al. (2013) suggest a weight of about 70% on QMJ and the remaining 30% on Value –, the poor performance of both Quality 1 and Value during the studied period made it difficult to expect high performance once combined in a QARP configuration.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std Dev</th>
<th>Sharpe ratio</th>
<th>Sortino ratio</th>
<th>VaR (VC) 95%</th>
<th>VaR (Historical) 95%</th>
<th>Expected shortfall 9%</th>
<th>Max drawdown</th>
</tr>
</thead>
<tbody>
<tr>
<td>QMJ 1</td>
<td>-0.045</td>
<td>1.43</td>
<td>-0.1</td>
<td>-0.14</td>
<td>-2.39%</td>
<td>-3.37%</td>
<td>-2.28%</td>
<td>15.5%</td>
</tr>
<tr>
<td>Value</td>
<td>-0.124</td>
<td>2.37</td>
<td>-0.1</td>
<td>-0.13</td>
<td>-4.02%</td>
<td>-5.63%</td>
<td>-4.13%</td>
<td>26.2%</td>
</tr>
<tr>
<td>QARP</td>
<td>-0.107</td>
<td>2.31</td>
<td>-0.09</td>
<td>-0.12</td>
<td>-3.91%</td>
<td>-5.49%</td>
<td>-4.29%</td>
<td>26.0%</td>
</tr>
</tbody>
</table>

Table 9: Performance and risk measures of QMJ 1, Value and QARP
5 Conclusion

This thesis first aims at highlighting the high complexity of quality investing and the underlying large dispersion in Quality definitions. While, historically, other factors have experienced some marginal discrepancies between measures used by practitioners and academics before converging towards the most effective one, a potential convergence appears much more challenging in the case of Quality. Indeed, Quality definitions even vary significantly among those two groups. Globally, the problem seems related to a fundamental difference in the perceptions of Quality and its purpose before being a debate about the ideal measures to capture the factor.

On the one hand, academicians tend to consider the quality of return performance as the ultimate purpose. Therefore, their Quality definitions contain characteristics that have significant predictive power for future stock returns. As a consequence, it is not surprising that their definitions tend to be reductive since often restrained to one or few measures whose link with Quality is not necessarily straightforward. For instance, associating the quality of a firm with its sole net stock issuance may appear questionable to many, however the implied significant returns here justify the Quality denomination.

On the other hand, practitioners tend to relate Quality to strong fundamentals irrespective of their potential predictive power for future returns. Therefore, their approach is also more comprehensive and pragmatic. Indeed, they constantly use several measures to capture different dimensions of firms’ financial statements. In fact, practitioners tend to look at Quality stocks as the direct opposite of the “junk” term often associated with bonds that are rated noninvestment grade because the issuer is financially distressed and on the brink of failure. In other words, when it comes to Quality, practitioners tend to act as credit analysts assessing the credit risk of a company. An interesting parallel can be drawn here with an empirical analysis of Muscettola and Nacaratto (2013) who, after having defined as “excellent firms” those which exhibit good profitability and solid financial balance, show that a rating model calibrated on bankruptcy (excellence) prediction may be a valuable tool for the selection of firms that, after three years, could turn to a state of excellence (bankruptcy).
As a result of this duality between search for future returns (academics) and search for strong fundamentals at any price (industry), purely academic definitions appear to consistently have significantly higher predictive power. In the face of such situation, this thesis proposes a Quality definition accepted by the majority of a sample of practitioners and based on Quality measures defended in the literature. While attempting to impulse a convergence towards those academic definitions, the hidden purpose was to eventually obtain a new generally accepted Quality definition with strong predictive power for future returns. However, the performance of the portfolio formed on this definition remains poor in comparison to academic definitions. In other words, even when confronted to what exists in the academic side, practitioners stick to their comprehensive vision of Quality.

At this stage, it is appropriate to put forward the idea that a Quality definition should not be considered good or bad on the basis of its return performance. Indeed, no one would consider a firm with high ROIC, low accruals, conservative leverage, and growing profitability and cash flow generation as of bad quality. Likewise, none of the Quality definitions presented in the framework of this thesis, be it derived from academics or industry, encompasses poor fundamental criteria. Nevertheless, their return performances are widely dispersed and statistically significant abnormal returns are far from being a general rule. Hence, this thesis highlights the fact that, when it comes to Quality, the convergence towards a unique Quality definition appears almost utopian and confirms that funds which offer exposure to the quality factor should not use the quality returns documented by academic studies as expectation.

In order to avoid pursuing blindly the returns promised by academicians by using an inappropriate Quality definition, practitioners should learn the definitions that matter from them in order to pick stocks with strong fundamentals that also have significant expected returns. Besides, in order to make factors more transparent so everyone is on the same wavelength, the choice of words should be make after thoughtful consideration. For instance, one can observe that, currently, some strategies, such as defensive equity strategies that promise equity like returns, delivered with lower volatility and smaller drawdowns, tend to be marketed as high quality strategies (Novy-Marx, 2014). Though the defensive label is maybe less appealing than Quality, it remains nonetheless much more concrete and understandable.

Finally, the confusion about Quality is maybe partly due to its own name. Indeed, this very confused term has been used as a catch-all label for a series of very diverse strategies so that it now means everything and nothing. A potential solution would be to consider the labels of the
Quality dimensions – i.e. Profitability, Growth, Earnings Quality, Safety, Payout, etc. – and only select those that match the strategy applied. Hence, from one subjective notion, one obtains a handful of more concrete categories which are individually well understood, for which a generally accepted definition is easier to develop, and which can be combined to capture the different exposures of the strategy. In other words, instead of stubbornly adding definitions under the Quality factor theme, one should maybe deconstruct it for the sake of clarity.


MSCI. (2016). *The Value Factor Marks a Decade of Disappointment*. Retrieved from MSCI: https://www.msci.com/blog-home?p_p_id=extendedlister_WARExtendedlister_INSTANCE_26NOeJbCkF5x&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-blog-search&p_p_col_count=1&_extendedlister_WARExtendedlister_INSTANCE_26NOeJbCkF5x_search=%22lastModification%22:[%22MM%22,%22YYYY%22,%22MM%22,%22YYYY%22].%22keywords%22:%22disappointment%22)#blog-entries


7 Appendix

7.1 International performance of different Quality measures

Figure 13 shows the returns of long-short (long the top 20% stocks and short the bottom 20% stocks) portfolios constructed on the basis of different quality measures. Returns are estimated based on monthly data and then annualized. The sample period is January 1986 – December 2014 for United States, Europe, Japan and Global markets and January 1993 – December 2014 for Emerging markets.

Figure 13: International performance of different Quality measures

Source: Kyosev et al. (2016)
7.2 List of measures definitions

7.2.1 Profitability measures

\[ ROE = \frac{IB}{SEQ} \]  
(7.1)

Where:

- IB denotes income before extraordinary items;
- SEQ denotes stockholders’ equity.

\[ ROA = \frac{IB}{AT} \]  
(7.2)

Where:

- AT denotes total assets.

\[ GPOA = \frac{(REVT - COGS)}{AT} \]  
(7.3)

Where:

- REVT denotes total revenue;
- COGS denotes cost of goods sold.

\[ CFOA = \frac{(IB + DP - \Delta WC - CAPX)}{AT} \]  
(7.4)

Where:

- DP denotes depreciation and amortization;
- CAPX denotes capital expenditures;
- \(\Delta WC\) denotes changes in working capital which is computed as follows:

\[ WC = ACT - LCT - CHE + DLC + TXP \]  
(7.5)

Where:

- ACT denotes current assets;
- LCT denotes current liabilities;
- CHE denotes cash and short-term investments;
- DLC denotes debt in current liabilities;
- TXP denotes income taxes payable.

\[ ROIC = \frac{EBIT}{WC + AT - ACT - INTAN - GDWL} \] (7.6)

Where:

- EBIT denotes earnings before interest and taxes;
- INTAN denotes intangible assets;
- GDWL denotes goodwill.

\[ GMAR = \frac{REV_T - COGS}{SALE} \] (7.7)

Where:

- SALE denotes total sales.

### 7.2.2 Earnings Quality measures

\[ ACC = -\frac{(\Delta WC - DP)}{AT} \] (7.8)

It is worth noting that, since low accruals are associated with quality (see section 3.1.3), a minus is inserted before the calculation so that companies with highest accruals obtain lowest figures for this measure.

### 7.2.3 Safety measures

\[ LEV = -\frac{(DLTT + DLC + MIBT + PSTK)}{AT} \] (7.9)

Where:

- DLTT denotes long-term debt;
- DLC denotes debt in current liabilities;
- MIBT denotes minority interests;
- PSTK denotes preferred stock.

Like ACC, a minus is inserted before the calculation since low leverage is associated with quality (see section 3.1.4.1).
Altman’s Z-score = \frac{(1.2 \times WC + 1.4 \times RE + 3.3 \times EBIT + 0.6 \times ME + SALE)}{AT} \quad (7.10)

Where:

- RE denotes retained earnings;
- ME denotes market value of equity.

### 7.1.4 Growth measures

\[ \Delta ROE = \frac{IB_t - IB_{t-3}}{SEQ_{t-3}} \quad (7.11) \]

\[ \Delta ROA = \frac{GP_t - GP_{t-3}}{AT_{t-3}} \quad (7.12) \]

\[ \Delta GPOA = \frac{GP_t - GP_{t-3}}{AT_{t-3}} \quad (7.13) \]

Where:

- \( \Delta \) denotes 3-year growth;
- GP denotes gross profit (i.e. REVT – COGS).

\[ \Delta GMAR = \frac{GP_t - GP_{t-3}}{SALE_{t-3}} \quad (7.14) \]

\[ \Delta CFOA = \frac{CF_t - CF_{t-3}}{AT_{t-3}} \quad (7.15) \]

Where:

- CF equals (IB + DP - \( \Delta WC \) – CAPX).

\[ \Delta ACC = \frac{MWCPD_t - MWCPD_{t-3}}{AT_{t-3}} \quad (7.16) \]

Where:

- MWCPD equals - (\( \Delta WC \) - DP).
7.3 CORR and REG procedure on QUALITY 1-2-3

<table>
<thead>
<tr>
<th>Table 10: Correlation between QMJ 1-2-3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pearson Correlation Coefficients, N = 126</strong>&lt;br&gt;Prob &gt;</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>QMJ 1</td>
</tr>
<tr>
<td>QMJ 2</td>
</tr>
<tr>
<td>QMJ 3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 11: Regression of QMJ 1 on QMJ 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameter Estimates</strong></td>
</tr>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>QMJ 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 12: Regression of QMJ 1 on QMJ 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameter Estimates</strong></td>
</tr>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>QMJ 3</td>
</tr>
</tbody>
</table>
7.4 Correlation between QMJ 1 and other factors

Figure 14: Correlation between QMJ 1 and other factors
7.5 Cumulative returns of individual measures

Figure 15: Cumulative returns of individual Profitability measures

Figure 16: Cumulative returns of Quality Earnings 1 (ACC)
Figure 17: Cumulative returns of LEV and Altman's Z-score

Figure 18: Cumulative returns of individual Growth measures
### 7.6 Basic statistics of individual long-short portfolios

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
<th>Skew</th>
<th>Kurt</th>
<th>Normality test p-value</th>
<th>t-stat p-value</th>
<th>Sign test p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROIC</td>
<td>126</td>
<td>-0.01330</td>
<td>0.94814</td>
<td>-2.68167</td>
<td>-0.06190</td>
<td>2.47619</td>
<td>0.10209</td>
<td>0.09986</td>
<td>0.62205</td>
<td>0.87516</td>
<td>0.32712</td>
</tr>
<tr>
<td>ROE</td>
<td>126</td>
<td>0.08238</td>
<td>1.75872</td>
<td>-4.84625</td>
<td>0.06292</td>
<td>5.54343</td>
<td>0.18319</td>
<td>0.95525</td>
<td>0.17935</td>
<td>0.59998</td>
<td>0.65618</td>
</tr>
<tr>
<td>ROA</td>
<td>126</td>
<td>0.19001</td>
<td>2.33374</td>
<td>-5.37639</td>
<td>0.15750</td>
<td>8.45984</td>
<td>0.47197</td>
<td>0.84200</td>
<td>0.11540</td>
<td>0.36251</td>
<td>0.53305</td>
</tr>
<tr>
<td>GPOA</td>
<td>126</td>
<td>0.40058</td>
<td>1.76923</td>
<td>-4.15509</td>
<td>0.19475</td>
<td>6.85088</td>
<td>0.59001</td>
<td>1.62280</td>
<td>0.01082</td>
<td>0.01226</td>
<td>0.18122</td>
</tr>
<tr>
<td>GMAR</td>
<td>126</td>
<td>-0.10932</td>
<td>1.40839</td>
<td>-4.89243</td>
<td>-0.19035</td>
<td>4.75033</td>
<td>0.39380</td>
<td>1.64363</td>
<td>0.01363</td>
<td>0.38527</td>
<td>0.04004</td>
</tr>
<tr>
<td>CFOA</td>
<td>126</td>
<td>0.13872</td>
<td>2.02737</td>
<td>-4.08775</td>
<td>-0.10192</td>
<td>6.62244</td>
<td>0.60396</td>
<td>0.76710</td>
<td>0.01692</td>
<td>0.44391</td>
<td>0.53305</td>
</tr>
<tr>
<td>ACC</td>
<td>126</td>
<td>0.04583</td>
<td>1.33574</td>
<td>-4.30818</td>
<td>0.13678</td>
<td>3.75610</td>
<td>-0.24565</td>
<td>1.34389</td>
<td>0.03406</td>
<td>0.70080</td>
<td>0.65618</td>
</tr>
<tr>
<td>LowLev</td>
<td>126</td>
<td>0.03034</td>
<td>1.83839</td>
<td>-9.78816</td>
<td>-0.01005</td>
<td>6.22320</td>
<td>-0.68634</td>
<td>7.02599</td>
<td>0.00000</td>
<td>0.85331</td>
<td>0.92906</td>
</tr>
<tr>
<td>Altman_Z</td>
<td>126</td>
<td>0.26967</td>
<td>2.11006</td>
<td>-4.82606</td>
<td>0.36348</td>
<td>9.07500</td>
<td>0.41161</td>
<td>2.14569</td>
<td>0.00658</td>
<td>0.15390</td>
<td>0.12958</td>
</tr>
<tr>
<td>G_ROA</td>
<td>126</td>
<td>0.07420</td>
<td>1.83236</td>
<td>-3.95084</td>
<td>0.04438</td>
<td>4.85069</td>
<td>0.06673</td>
<td>-0.01281</td>
<td>0.56994</td>
<td>0.65023</td>
<td>0.92906</td>
</tr>
<tr>
<td>G_ROE</td>
<td>126</td>
<td>0.02449</td>
<td>1.70268</td>
<td>-4.64784</td>
<td>0.14812</td>
<td>4.41952</td>
<td>-0.22769</td>
<td>0.44382</td>
<td>0.24481</td>
<td>0.87200</td>
<td>0.42279</td>
</tr>
<tr>
<td>G_GPOA</td>
<td>126</td>
<td>0.08451</td>
<td>1.81026</td>
<td>-4.59234</td>
<td>0.18283</td>
<td>6.70347</td>
<td>0.41116</td>
<td>1.20080</td>
<td>0.09089</td>
<td>0.60118</td>
<td>0.65618</td>
</tr>
<tr>
<td>G_GMAR</td>
<td>126</td>
<td>-0.09190</td>
<td>1.47566</td>
<td>-4.10933</td>
<td>-0.11967</td>
<td>4.15601</td>
<td>-0.01113</td>
<td>0.50778</td>
<td>0.75331</td>
<td>0.48581</td>
<td>0.32712</td>
</tr>
<tr>
<td>G_CFOA</td>
<td>126</td>
<td>0.03627</td>
<td>1.28489</td>
<td>-2.78495</td>
<td>-0.00031</td>
<td>3.14408</td>
<td>0.11563</td>
<td>-0.47462</td>
<td>0.69927</td>
<td>0.75188</td>
<td>1.00000</td>
</tr>
<tr>
<td>Growth 1</td>
<td>126</td>
<td>0.002591</td>
<td>1.48654</td>
<td>-4.96744</td>
<td>0.012884</td>
<td>3.99705</td>
<td>-0.26222</td>
<td>0.60146</td>
<td>0.57101</td>
<td>0.98442</td>
<td>0.92906</td>
</tr>
</tbody>
</table>
7.7 Regression: Size premium controlling for QMJ

| Variable | Label  | DF | Parameter Estimate | Standard Error | t Value | Pr > |t| |
|----------|--------|----|-------------------|----------------|--------|------|---|
| Intercept | Intercept | 1 | 0.10212 | 0.15600 | 0.52 | 0.6033 |
| Mkt_RF | Mkt-RF | 1 | 0.08943 | 0.06095 | 1.47 | 0.1449 |
| HML | HML | 1 | -0.03206 | 0.08747 | -0.37 | 0.7146 |
| UMD | UMD | 1 | 0.06391 | 0.04531 | 1.19 | 0.2364 |
| QMJ | QMJ | 1 | -0.27048 | 0.11258 | -2.40 | 0.0178 |

Table 13: Regression of SMB on QMJ
### 7.8 PROC TTEST and normality checks

#### Tests for Normality (contraction)

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shapiro-Wilk</td>
<td>W</td>
<td>0.9509</td>
</tr>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>D</td>
<td>0.139442</td>
</tr>
<tr>
<td>Cramer-von Mises</td>
<td>W-Sq</td>
<td>0.035878</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>A-Sq</td>
<td>0.283144</td>
</tr>
</tbody>
</table>

#### Tests for Normality (expansion)

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shapiro-Wilk</td>
<td>W</td>
<td>0.99328</td>
</tr>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>D</td>
<td>0.036485</td>
</tr>
<tr>
<td>Cramer-von Mises</td>
<td>W-Sq</td>
<td>0.02154</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>A-Sq</td>
<td>0.164962</td>
</tr>
</tbody>
</table>

#### US Business Cycle Method

<table>
<thead>
<tr>
<th>US_Business_Cycle</th>
<th>Method</th>
<th>Mean</th>
<th>95% CL Mean</th>
<th>Std Dev</th>
<th>95% CL Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contraction</td>
<td></td>
<td>0.4673</td>
<td>-0.4543</td>
<td>1.3889</td>
<td>1.4448</td>
</tr>
<tr>
<td>Expansion</td>
<td></td>
<td>-0.1363</td>
<td>-0.3884</td>
<td>0.1157</td>
<td>1.1594</td>
</tr>
<tr>
<td>Diff (1-2)</td>
<td>Pooled</td>
<td>0.6036</td>
<td>-0.0948</td>
<td>1.3021</td>
<td>1.2609</td>
</tr>
<tr>
<td>Diff (1-2)</td>
<td>Satterthwaite</td>
<td>0.6036</td>
<td>-0.3458</td>
<td>1.5531</td>
<td></td>
</tr>
</tbody>
</table>

#### Equality of Variances

<table>
<thead>
<tr>
<th>Method</th>
<th>Num DF</th>
<th>Den DF</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Folded F</td>
<td>18</td>
<td>106</td>
<td>2.11</td>
<td>0.0196</td>
</tr>
</tbody>
</table>

| Method     | Variances | DF | t Value | Pr > |t| |
|------------|------------|----|---------|------|-----|
| Pooled     | Equal      | 124 | 1.71    | 0.0897 |
| Satterthwaite | Unequal   | 21.126 | 1.32   | 0.2004 |