KU LEUVEN

HOGESCHOOL-UNIVERSITEIT BRUSSEL HUB

FACULTEIT ECONOMIE EN BEDRIJFSWETENSCHAPPEN

Intern aangestuurde masterproef

A composite indicator for the prediction of the one-year-ahead return of stocks: an application to the Belgian stock market

Michaël HAMELRYCK

Masterproef aangeboden tot het behalen van de graad van
Master of Science handelsingenieur

Promotor: Prof. Dr. Filip VAN DEN BOSSCHE

Copromotor: Prof. Dr. Tom VAN PUYENBROECK

Academiejaar: 2013 – 2014

Verdedigd: juni 2014

FACULTEIT ECONOMIE EN BEDRIJFSWETENSCHAPPEN CAMPUS BRUSSEL (HUBRUSSEL)

WARMOESBERG 26 1000 BRUSSEL BELGIË

KU LEUVEN

HOGESCHOOL-UNIVERSITEIT BRUSSEL

HUB

FACULTEIT ECONOMIE EN BEDRIJFSWETENSCHAPPEN

Masterproef

A composite indicator for the prediction of the one-year-ahead return of stocks: an application to the Belgian stock market

Michaël HAMELRYCK

Abstract

Even when financial markets are bullish, many investors struggle to achieve decent returns. In this paper, we analyse different financial statement variables of respectively 35, 30 and 38 traditional stocks quoted on the Euronext Brussels segment in a three year period (2011-2013). The main goal is to rank these stocks in a relative way in order to reflect their one-year-ahead expected return. We develop a composite indicator (CI) where subjective judgements are reduced to a minimum, as this is one of the main critics of composite measures. Therefore, we apply six normalization methods, four weighting schemes and three aggregation systems to assess the influence of the choices we have made during the entire development process of the CI. This should enable us to select a justifiable method. The results demonstrate that flexible weighting schemes, such as data envelopment analysis, are amongst the best performers in terms of predicting stock returns. Additionally, we show that fundamental analysis (FA) is still useful to investors, even in turbulent stock markets.

Keywords: stock return, fundamental analysis, composite indicator, data envelopment analysis, Belgian stock market

JEL-code: C43, C58, D53, G11, G12, G17, G32

1 Introduction

Every year, "*De Tijd*", a Belgian economic newspaper, publishes a guide for investors with relevant information about the most important Belgian stocks. According to the authors, this guide should enable investors to construct a decent stock portfolio. If this were true, then why are only a few investors able to develop an investment strategy that yields consistent profits (Lipe, 1998)? Clark-Murphy and Soutar (2004) find an important reason for this phenomenon. They conclude that investors are making investment decisions based on emotional grounds. Small investors tend to buy stocks that have risen sharply, believing that these stocks will continue to perform well. However, the chance increases that these stocks will actually underperform in the next periods. In the mean time they are also neglecting less exciting, undervalued stocks that have more growth potential in terms of return. This naivety of small investors has already been noticed in 1988 by Andreassen (cited in Lipe (1998)) and is clearly still relevant. Hence, we try to develop a tool for investors which assures that investment decisions are based on rational grounds. We will do this by constructing a composite indicator (CI) that ranks stocks according to their expected one-year-ahead return.

We start our research by examining the data in the guide published by *De Tijd*. Here, we find more than 20 indicators that should be predictive for future stock returns. Including all these indicators in our investment decisions makes it rather unmanageable to select the best stocks. Moreover, different questions arise in this decision process. Which indicators are relevant? How can we compare stocks with each other? Is every indicator equally important? For example, do we choose a stock with a low profit margin, low growth perspectives and a low valuation or do we prefer a stock with high profit margins, high growth perspectives and a dazzling valuation? If for instance 40 indicators are considered instead of only three basic ones, this problem becomes even more complicated. Hence, this is the ultimate motivation for the construction of a CI, an aggregate measure which allows evaluating stocks with the blink of an eye, even if we incorporate as many as 40 indicators.

This study contributes to different fields of research. First, it contributes to the literature of financial economics by examining whether fundamental analysis (FA) can still predict returns in turbulent stock markets (see Curtis, 2012 for an overview). Second, this research complements the existing literature on the most important indicators in fundamental and financial statement analysis. Additionally, a survey that was carried out by Richardson, Tuna en Wysocki (2010) clearly shows that insufficient research has been performed in bundling different accounting attributes in one composite index in order to predict profit and return. Moreover, both academics and individual investors state that more empirical research is needed to predict the fundamental value of a company (Richardson et al., 2010). Hence, according to this survey, our research seems very relevant. Of course, our main goal remains to analyse whether or not a composite measure can yield a profitable trading strategy.

The remainder of this paper is organized as follows. Section 2 will start with a literature review of FA, the selection of indicators that are predictive for stock returns and the use, construction and relevance of CI's. Subsequently, Section 3 presents the methodological framework that will be adopted when developing the CI. Section 4 discusses the results of this research. To end this paper, Section 5 will give some concluding remarks and suggestions for future research.

2 Predicting stock returns

This literature review consists of three important parts necessary to develop a framework for the construction of CI's. First, we will start by giving a more in-depth review of the usefulness of FA, one of the building blocks of our study. Second, an assessment of the most important and frequently applied indicators in FA will be presented. Indeed, the main goal is to select the most widely accepted indicators that are predictive for future returns. Third, we will discuss whether CI's are an appropriate tool for fundamental investment strategies and stock selection. To conclude this section, we will highlight the debate between believers and non-believers of CI's or aggregate measures.

2.1 Fundamental analysis

The existing literature presents many models trying to predict the return of stocks. In 1952, Markowitz developed a portfolio theory in which a set of efficient portfolios was defined. Then, equilibrium models such as the Capital Asset Pricing Model (Sharpe, 1964) and the Arbitrage Pricing Theory (Ross, 1976) were developed. Another category of prediction models is the efficient market theory which states that there is so much competition between investors that stock prices reflect all relevant information at any moment (Fama, 1965). Hereafter, different valuation models arose, such as the models developed by Feltham and Ohlson (1995) and the Real-Option-Based valuation model (Zhang, 2000), where the price of a stock is predicted by investigating financial variables (see Callen, 2013 for a complete overview). Our research is situated within this last category. The technique that will be used is called fundamental analysis (FA) and is often referred to as financial statement analysis.

FA is an important building block of our study and can be described as a method that connects financial data from public sources, such as the annual account and balance sheet of a firm, with the true value of that company (Lev & Thiagarajan, 1997; Abad, Thore, & Laffarga, 2004; Samaras, Nikolaos, & Zopounidis, 2008; Grimm, 2012). According to Abad et al. (2004) this takes place in two subsequent steps. First, a predictive link has to be established in order to predict future profits or cash flows of a company (Chen & Zhang, 2007; Jiang & Penman, 2013). Second, a valuation link has to be constructed so expected profit can be compared with the price of the stock. Hence, the first question that investors should ask themselves is: "Which financial variables do I need to make a good estimation of future growth?". The next question that arises is: "Which financial ratios determine whether a stock becomes a buy opportunity?". Both questions have been a subject of academic research for years (Jiang & Penman, 2013).

Ball and Brown (1968), Ou and Penman (1989), Chen and Zhang (2007), Samaras et al. (2008) and many others showed that by implementing a fundamental strategy, the true or intrinsic value of a company can be estimated. If investors are rational, stock prices reflect this intrinsic value. Hence, fundamental values should be correlated with the stock price of that company. In other words, when a stock is currently trading under (respectively above) its intrinsic value, we call this an undervalued (respectively overvalued) company. Undervalued companies can be seen as buy opportunities because we expect that in the long term, stock prices will evolve to their fundamental value. The opposite is true for overvalued companies (Samaras et al., 2008). Therefore, a fundamentalist sees stock prices as speculative because, at some moments in time, they can differ from their intrinsic value (Ball & Brown, 1968; Jiang & Penman, 2013).

Hence, FA seems a very useful tool, but has it also been proved to be successful in the past? Ou and Penman (1989), Holthausen and Larcker (1992), Lev and Thiagarajan (1993), Abarbanell and Bushee (1997), Setiono Strong (1998), and Charitou and Panagiotides (1999) are only a few of many studies proving that by developing a trading strategy based on FA, it is possible to achieve higher returns. However, Van Caneghem, Van Campenhout and Van Uytbergen (2002) showed that the fundamental models developed by Ou and Penman (1989) and Holthausen and Larcker (1992) do not yield consistent abnormal returns when applied to the Belgian stock market. The study of Curtis (2012) complements this by demonstrating that during the time period 1979-1993, accounting variables and stock prices moved together. Yet, Curtis (2012) also showed that in the subsequent period (1994-2008), which was characterized by bubbles, temporary crashes and major financial crises, this relation between accounting variables and prices was not found. Further research should be carried out in order to get more insight in whether this new pattern is structural or not.

2.2 Selection of indicators

FA looks interesting for predicting stock returns. The main question however remains which indicators should be taken into account to make a decent FA of a company. In order to select the most important indicators, we have chosen to consider both academic literature and 'wisdom' of some stock gurus (such as Warren Buffet, Kenneth L. Fisher and Peter Lynch). Additionally, we have checked which indicators are important to individual investors. In the following paragraphs, one can find a summary of the literature reviewed. A list of the indicators that will eventually be used in our study is available in Appendix A.

Matsumoto, Shivaswamy and Hoban (1995, cited in Delen, Kuzey, & Uyar, 2013) carried out a survey to investigate the perception of financial analysts about the ability of different financial variables to predict stock returns. Their conclusion was that growth ratios were seen as the most important ones, followed by valuation ratios and profitability ratios. However, there was no clear consensus among the financial analysts. Atsalakis and Valavanis (2009), Delen et al. (2013) and Jiang and Penman (2013) complement this by stating that there is no generally accepted list of financial ratios that has to be applied when evaluating the financial health of a company. Ho and Wu (2006) for example used 59 ratios in their analysis, Delen et al. (2013) investigated the influence of 26 financial ratios, Van Caneghem et al. (2002) added a list of 52 ratios, while Cinca, Molinero and Larraz (2005) only mentioned 16 ratios (also see Atsalakis & Valavanis, 2009 for an extensive overview). Atsalakis and Valavanis (2009) even mention cases where only two variables were evaluated. Most handbooks and studies however assume that 20 to 30 of the most commonly used ratios are sufficient to evaluate a firm as it becomes very unpractical to consider all available ratios.

One of the most cited papers in the field of financial statement analysis is Ou and Penman (1989). In their analysis, they conclude that there are different variables that are predictive for future profits and therefore also for future returns. Some of these variables are: change in current and quick ratio, change in dividend, return on equity (ROE), solvability, turnover to assets ratio, return on assets (ROA), gross profit margin, EBITDA margin, operational margin and net profit margin. However, their results were not consistent for all time periods considered. Other papers, such as Morton and Shane (1998), Sneed (1999), and Sharma and Sharma (2009), used Ou and Penman (1989) as their building block. Another widely cited paper for the selection of financial indicators is Lev and Thiagarajan (1993). While Ou and Penman (1989) opted for statistical methods to derive significant variables, Lev and Thiagarajan (1993) performed a guided search, roughly the same method as we have used for the selection of variables for our CI. Lev and Thiagarajan (1993) searched the Wall Street Journal and Barron's Value Line for professional remarks about financial variables that explain future profit and stock prices. This guided search approach resulted in a set of twelve variables. Different studies such as Abarbanell and Bushee (1997, 1998) and Al-Debie and Walker (1999) implemented these twelve variables in their research. For an in-depth review of the most cited papers discussing important variables in FA, we refer to the excellent work of Ramnath, Rock and Shane (2008) and Richardson et al. (2010).

As mentioned before, growth ratios are seen as very important in different studies if one wants to predict stock returns. Additionally, there is overwhelming evidence that 'value stocks', i.e. stocks with a low valuation (measured by the price to earnings ratio and price to book ratio), on average outperform 'growth stocks', i.e. stocks with a high valuation accompanied by high growth prospects (e.g. Basu, 1997; Fama & French, 1992; Chan & Lakanishok, 2004; Arnott, 2005). Another factor that is known to have an impact on return is the size of the company. Small companies tend to have more growth opportunities as it is easier to grow at a pace of 4% a year. Large companies have to achieve an absolute growth that is much higher in order to achieve this 4%. Arnott (2005) and Bodie, Kane and Marcus (2007) are only two of many studies confirming this phenomenon.

In addition to academic literature, we should also take into account what private and professional investors value as they are the targeted audience for the CI developed in this paper. Moreover, Shleifer and Summers (1990) showed that individual investors are also able to influence the market. Hence, it is important that we add indicators to our model that are meaningful in the decision making process of investors. Clark-Murphy and Soutar (2004) explore which attributes can influence the choice of buying a particular stock. Their results demonstrate that investors are not interested in speculation and that the majority are long-term investors. When selecting stocks, financial measures such as the price to earnings ratio and the dividend yield are very important. Yet, these measures are less decisive than management performance and recent movements of the stock price.

2.3 The scope of composite indicators

The goal of a CI is to reveal relative positions, in this case between stocks. It aggregates different individual indicators into one relative metric which eases the comparison of the cases that have to be evaluated. Brand, Saisana, Rynn, Pennoni and Lowenfels (2007) for example, compare the efficiency of alcohol controls in OECD countries, showing that CI's are useful in identifying trends and drawing attention to potential problems (Nardo, Saisana, Saltelli, Tarantola, Hoffman, & Giovanni, 2008). Accordingly, CI's are mainly developed when a policy instrument has to be constructed or when the performance of for example countries has to be compared with a specific benchmark (see Bandura, 2008 and JRC, 2014 for an overview of the most recent publications on CI's). The use of CI's is also encouraged when multidimensional concepts such as competition, industrialization, happiness, world peace and literacy have to be measured (Nardo et al., 2005, 2008; JRC, 2014). Hence, because the evaluation and ranking of stocks is a multidimensional process, we presume that a CI could be an appropriate tool. Additionally, the variety of CI's also reflects their value in summarizing complex problems and initiating discussions (Cherchye et al., 2006). Although the use of CI’s is increasing, they remain a subject of controversy as different subjective judgements, such as the selection of the weighting scheme, have to be made.

Other examples demonstrating the usefulness of CI’s are: the Technology Achievement Index (e.g. Saisana, Tarantola, & Saltelli, 2005; Cherchye et al., 2006; Nardo et al., 2005, 2008), the measurement of nature-based tourism in Cuba (Pérez, Guerrero, González, Pérez, & Caballero, 2013), monitoring the vegetation in Australia (Pert, Butler, Bruce, & Metcalfe, 2012), the Environmental Performance Index (Rogge, 2012), and the measurement of the efficiency of health systems (Smith, 2002).

Clearly, the application of CI’s is very widespread. However, CI’s have not drawn much attention in the area of predicting stock returns. Yet, a CI could clearly help investors with the construction of their portfolio. All factors that one wants to take into account could be put together in one single number, making comparisons across stocks very easy. In other words, it becomes possible to select the most promising stocks with the blink of an eye. The existing literature presents two major schools of techniques that enable the transformation of financial statement information into a robust prediction of future stock returns. The first school makes use of different econometric techniques such as logit/probit models (e.g. Ou & Penman, 1989; Holthausen & Larcker, 1992; Stober, 1992; Greig, 1992; Bernard, Thomas, & Wahlen, 1997; Setiono Strong, 1998; Charitou & Panagiotides, 1998; Beneish, Lee, & Tarpley, 2001; Van Caneghem et al., 2002) or regression analysis (e.g. Lev & Thiagarajan, 1993; Sloan, 1996; Abarbanell & Bushee, 1997; Mironiuc & Robu, 2013). The second school is based on neuro(fuzzy) techniques (see Atsalakis & Valavanis, 2009 for an overview of more than 100 related articles). Even though these studies purvey a valuable contribution to the academic field of financial economics, they all lack one very important factor: relative comparisons. Investment decisions have to be made in a relative way. By simply looking at one stock, one cannot conclude whether this is a good investment or not (Edirisinghe & Zhang, 2007). A CI solves this issue as it results in a relative ranking of the stocks.

Moreover, CI’s have been employed for decades in their simplest form. Mohanram (2005) for example, assigns a 0/1 score to six different financial variables for each stock. Hence, the sum of these six scores results in a CI. Likewise, the well-known magazine Business Week, ranks stocks based on three criteria: return on equity, profit growth and turnover growth. The first criterion gets a weight of 0,5, while the other two criteria each get a weight of 0,25. The weighted sum of these criteria gives the CI (Yu & Kim, 2009). The CI’s constructed by Mohanram (2005) and Business Week have already proven their value. However, in this study we would like to implement some more advanced methods in order to derive a robust CI. The studies that have the most common ground to the research that will be performed here are those of Edirisinghe and Zhang (2007) and Samaras et al. (2008). Samaras et al. (2008) rank stocks according to their utility functions, while Edirisinghe and Zhang (2007) construct a Generalized-DEA model for portfolio optimization using FA.

2.4 Debate on composite indicators

The application of CI’s has been encouraged by among others, the European Commission. One of their main arguments is that by aggregating variables, CI’s are able to summarize the core of different complex or multidimensional problems (European Commission, 2013). Additionally, we notice that the amount of CI’s has increased exponentially over the years (Nardo et al., 2005, 2008). Despite this, we also see that some academics question the usefulness and robustness of composite measures. Hence, separate indicators are still utilised as a base for decision making, even when multidimensional concepts have to be evaluated. Table 1 gives a brief summary of the main advantages and disadvantages of CI’s. For a more in-depth discussion, we refer to Saisana and Tarantola (2002), Sharpe (2004) and Nardo et al. (2005, 2008).

Advantages	Disadvantages
• Can summarize complex and multidimensional problems • Are more easy to interpret than a bunch of separate indicators • Make it possible to compare and rank multidimensional concepts • Enable the communication with a bigger audience (civilians, companies, media, investors, …) • Make it possible to include more information within the existing size limit	• May send misleading policy messages, here investment decisions, if poorly constructed or misinterpreted • May lead to simplistic conclusions • Are often not transparent enough • The selection of indicators and weights remains subjective and could be the object of dispute • May disguise serious failings in some dimensions if the construction is not transparent

Table 1: Advantages and disadvantages of CI's (Saisana & Tarantola, 2002).

3 Methodology

During recent years, clear methodological frameworks for the construction of CI's have been developed. One of the most widely applied and well documented handbooks is Nardo et al. (2005, 2008). In this handbook, the authors define ten steps that have to be followed in order to construct a transparent and robust CI. We will implement this framework in our research to ensure the robustness of our indicator. A graphical representation of these steps is presented in Figure 1. There are of course other frameworks that could have been followed. Dobbie and Dail (2013) for example, choose to only adopt five of these steps. Most frameworks however look very similar as the construction of CI's has become very standardized. Note that the link with other indicators (Step 9 in Figure 1) will not be considered in our research as this is less relevant in the context of FA.

1. Developing a theoretical framework
2. Data and variables selection
3. Imputation of missing data
4. Multivariate analysis
5. Normalization of the data
6. Weighting and aggregation of indicators
7. Uncertainty and sensitivity analysis
8. Looking back at the data
9. Link with other indicators
10. Visualization of results

Figure 1: Schematic overview of the methodology used to construct a CI (based on Nardo et al., 2008).

3.1 Theoretical framework

A decent theoretical framework is the building block of every CI (Nardo et al. 2005, 2008). Without a clear definition of the concept that has to be measured we will never obtain useful results. In our study, we want to define a CI that ranks stocks according to their one-year-ahead expected return.

Hence, the most promising stocks have to be selected, based on rational rather than on emotional grounds, as is usually the case now (Clark-Murphy & Soutar, 2004). The trading strategy that will be applied is a passive buy and hold strategy. This means that we buy a portfolio with the most promising stocks (e.g. three, five, seven or ten stocks) on the first trading day of the year. Subsequently, we sell this portfolio on the last trading day of that same year. We will thus not try to improve our return by trading on other days, often referred to as active trading. The buy and hold strategy limits our transaction costs and makes the interpretation of the results easier. Additionally, as shown earlier (see Section 2.3), other stock selection methods are criticized because they do not compare stocks in a relative way. Therefore, we will also focus on this part when developing the CI.

The final composite tool has to possess all following properties (Samaras et al., 2008):

(1) Useful for both academics and practitioners;
(2) Stocks should be ranked in a correct and transparent way and anticipate the problems of portfolio construction (as discussed in Section 1);
(3) Implementable in real world applications;
(4) In agreement with investors' individual preferences and needs;
(5) Flexible across objectives, over stocks and through time.

Additionally, the following notations will be used frequently in the remainder of the paper:

- x_{qs} the value of indicator q (q=1,…,Q) for stock s (s=1,…,S);
- I_{qs} the normalized value of x_{qs};
- $\bar{x}_q$ the mean value of indicator q (q=1,…,Q);
- $Direction_q$ the correlation of indicator q (q=1,…,Q) with expected return (i.e. 1 or -1);
- S the amount of stocks in a specific year, D the amount of dimensions, and N_d the amount of indicators in dimension d (d=1,…,D).
- Note that in our study, D = 7, Q = 39 and S = 30, 35 or 38 depending on the year we want to analyse.

Other notations will be explained throughout the paper. Note that we could have added a time suffix t in the notations. However, as we wanted to simplify the equations, we have decided not to do so.

3.2 Data and variables selection

Based on the literature review in Section 2.2, we have selected 33 indicators (see Appendix A for an overview). This set of indicators is certainly not exhaustive, but it represents the most commonly used indicators in FA. After examining strategies of world's most famous investors (such as Warren Buffet, Kenneth L. Fisher, Peter Lynch and some Belgian analysts that helped constructing the stock guide), we decided to add six extra indicators, namely: profit growth acceleration, (expected) payback time, difference in return on assets, expected EBITDA growth and net cash per share (Cunningham, 1998; De Tijd, 2013). Adding these variables to those derived from our literature review, we get a total of 39 individual indicators for the construction of the CI (see Figure 2).

The data necessary for constructing the CI consists of different financial accounting ratios for Belgian firms quoted on the Euronext Brussels segment. The accounting information was taken from the stock guide (De Tijd, 2011, 2012, 2013) and from annual reports of the companies, as not all necessary information was available in the guide. We always focused on the consolidated financial statements because it is the consolidated information, and not the information for the parent or single firm, that is taken into account when valuing a company (Abad et al., 2004).

Growth opportunities are important for future performance (as discussed in Section 2.2). Because this is not directly measurable, we will use growth forecasting by analysts as a proxy (Chen & Zhang, 2007). A clear disadvantage of this proxy is that when a company reports turnover or profit that is much higher or lower than these estimates, the stock price will react quite heavily. Hence, if analysts' predictions are too optimistic, as suggested by Boudt, De Goeij, Thewissen, and Van Campenhout (2012), this may induce a bias in our CI. An alternative considered by Sharma and Sharma (2009), is the extrapolation of past growth rates into the future. We have chosen not to adopt this approach as we think that analysts respond better to new market situations. Additionally, the track record of the company's management is also difficult to quantify. We calculated mean growth rates (both turnover and profit) for the last three years and the variance of dividends and profit during the last five years as a proxy, as we want to identify steady growing companies without large fluctuations in financial variables. Moreover, we also added the sales to assets ratio and the sales to equity ratio to estimate management performance (Samaras et al., 2008).

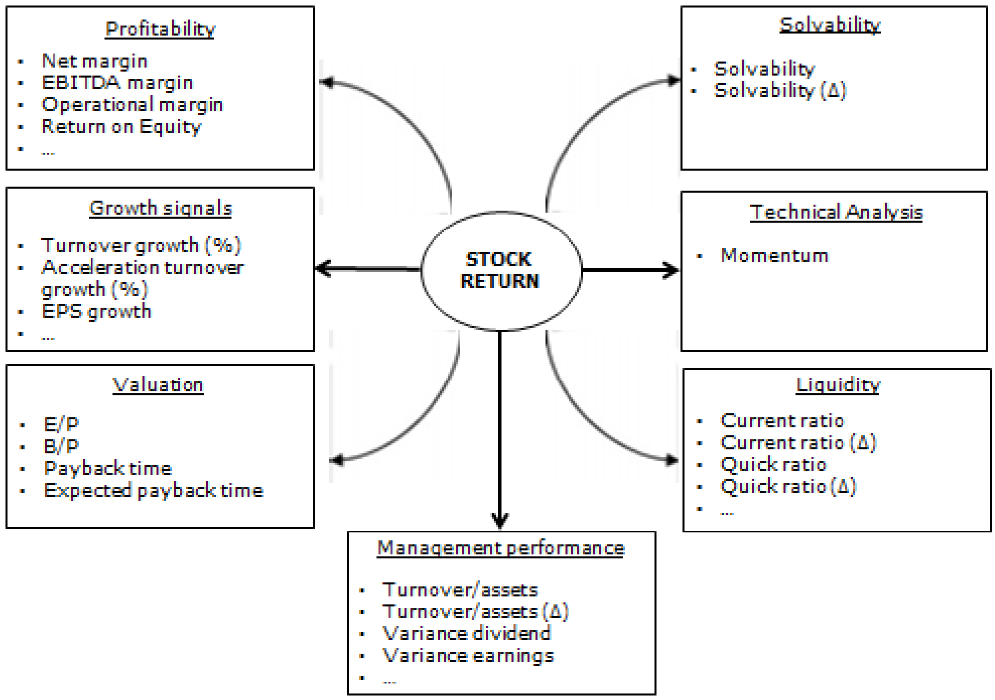


Figure 2: Schematic overview of indicators that estimate stock return.

Strictly speaking, technical analysis does not belong to the field of FA. However, we see that in some studies on FA (e.g. Jegadeesh & Titman, 1993; Clark-Murphy & Soutar, 2004; Yu & Kim, 2009; Chen et al., 2010) momentum, an indicator of the technical analysis dimension, is also considered. Therefore, we added this variable to our list.

Ideally, if we want to construct a transparent and robust indicator, we require a population of companies that is homogenous in terms of common practices. This means that if we compare stocks, the same variables should be predictive for one-year-ahead returns. However, the value of banks, biotech companies and holdings is calculated in a different way than what we are used to for traditional (manufacturing and services) companies. For biotech companies for example, the rate of cash burn and the development of products have a major impact on the value of the company. Turnover and profit are only to a lesser extent important. Hence, it becomes complicated to compare biotech companies with traditional companies, as their evaluation does not consist of the same input variables. Therefore, we only evaluated the traditional companies and eliminated all banks, biotech companies and holdings from our analysis, like among others Kallunki, Martikainen, and Martikainen (1998), Giner and Reverte (2006), Samaras et al. (2008), Chen et al. (2010), and Mironiuc & Robu (2013) did. The number of traditional companies in the stock guide was 36 in 2011, 34 in 2012 and 39 in 2013. Sneed (1999) suggests dropping firms with negative equity (one case in 2012 and 2013) as this would result in meaningless indicators. In addition to this, stocks with a stock price under one euro (one case in 2011 and 2012) were also omitted from the data as these stocks would be a major subject to speculation. Hence, their price evolution would not be value relevant. Finally, after dropping firms with lacking analysts' predictions or other important variables that could not be estimated (two cases in 2012), we ended up with 35 companies in 2011, 30 companies in 2012, and 38 companies in 2013.

Note that the selection process of individual indicators remains quite subjective as there does not exist one definitive set of fundamental variables (Nardo et al., 2005, 2008; Delen et al., 2013; Jiang & Penman, 2013). Different questions arise. How many indicators do we consider in our analysis? Which indicators do we take into account? Yet, we have tried to limit the subjectivity when selecting indicators by implementing both financial analysts' opinions and a literature review (as discussed in Section 2.2.). A list of all 39 indicators that will be used as input in our analysis is provided in Appendix B. For a schematic overview, we again refer to Figure 2. Additionally, we refer to Appendix C for a complete overview of all stocks that will be considered in this paper.

3.3 Imputation of missing data

Missing data is one of the most hindering problems when developing robust CI's (Nardo et al., 2005, 2008). Fortunately, in the context of financial variables, missing variables is not a frequent problem as all annual reports and balance sheets consist of the same compulsory numbers. The only missing data we had to account for were missing analysts' predictions. We identified two such cases: Atenor and Banimmo, both in 2012.

Numerous methods are available to impute or correct for missing data. According to Nardo et al. (2005, 2008), three general methods are used in practice: (1) case deletion, (2) single imputation, and (3) multiple imputation. In the case of Atenor, the only missing value was the projected turnover for the year to come. We accounted for this missing projection by substituting it by the mean turnover in the last two years (Nardo et al., 2005, 2008). The case of Banimmo was trickier. No projections were available at all for this company and as no EBITDA numbers were reported, we decided to delete the case from our analysis.

3.4 Multivariate analysis

3.4.1 Accounting for outliers and non-normality

Outliers could have the power to be influential for our results. Specifically, they could have a misleading impact when we normalize our data in the subsequent step. As we do not want this to happen, we account for outliers in two different ways. First, we make some adjustments through Box-Cox transformations and winsorization. Second, we choose normalization methods that are not affected by outliers. One disadvantage of these normalization methods, such as ranking and categorical scales, is that they omit the absolute level of information, which is often criticized (Nardo et al., 2008). Both techniques will be performed separately in our analysis.

The first step is to identify the outliers. One way is to standardize all values according to the following equation (1), where σ_q is the standard deviation of indicator q:

$$I_{qs} = \frac{x_{qs} - \bar{x}_q}{\sigma_q} \tag{1}$$

If the standardized value was larger than three, we labelled this case as an outlier. Another way is to calculate the skewness and kurtosis of the indicator. When an indicator had a skewness larger than two and simultaneously a kurtosis larger than 3,5, this indicator needed extra treatment before the main analysis could start (Jacobs, Smith, & Goddard, 2004; Saisana, 2012).

In about half of the indicators we had to deal with problematic cases or non-normality, i.e. indicators with either high standardized values or high values for both skewness and kurtosis. Therefore, we tried different methods to account for this problem. The first method is widely used in statistics and is suggested by Nardo et al. (2008) and Saisana (2012). Here, we accounted for problematic values through Box-Cox transformation or winsorization. Our preference goes to Box-Cox transformation as this affects all the cases in the indicator and does not require manual modifications. For variables such as market capitalization it is very straightforward which transformation should be performed. Because some companies have an extremely high market capitalization compared to others (e.g. AB InBev in our analysis), we need to move this indicator 'down the ladder of powers' (Fox, 2008). Hence, an LN Box-Cox transformation is very often applied in practice (e.g. Arnott, 2005; Atsalakis & Valavanis, 2009; Chen et al., 2010). Indicators C7 and G5 were both adjusted by an LN Box-Cox transformation, as the data was suited for this correction. For additional information on how these transformations are performed, we refer to Fox (2008).

The main problem however was that for all other problematic indicators (i.e. except C7 and G5 that have already received a Box-Cox transformation), different cases with negative values occurred. Therefore, we had to solve this problem with another technique called winsorization. Winsorization is a very simple technique that substitutes the most extreme value in a problematic indicator by the second most extreme value (Dixon & Yuen, 1974). The main critique of winsorization is that some level of absolute information is lost as the largest (or smallest) value becomes equal to the second largest (or second smallest) value. However, Saisana (2012) also uses this technique to deal with outliers. According to his guide, winsorization can be used when the amount of outliers per indicator remains small (roughly 5%). Hence, we first applied a Box-Cox transformation on all indicators with non-negative values (i.e. C7 and G5) and if this was not possible we performed a winsorization. After making these adjustments, no problematic indicators remained. Additionally, we checked whether the correlations between indicators remained roughly the same before and after adjusting for problematic values. As almost no fluctuations occurred, we decided to continue the analysis with this improved dataset.

3.4.2 Grouping information on individual indicators

Based on the accounting literature, like for example Edirisinghe and Zhang (2007), Samaras et al. (2008), Wang and Lee (2008), and Sharma and Sharma (2009), we have grouped our individual indicators together in several dimensions, as listed in Table 2.

Dimension	
Profitability	A
Solvability	B
Growth signals	C
Technical Analysis	D
Valuation	E
Liquidity	F
Management performance	G

Table 2: Dimensions or the highest hierarchical level grouping indicators.

We did not perform any statistical procedures, such as principal component or factor analysis, to confirm these dimensions for two reasons. First, these dimensions are quite arbitrary, and even when statistical tests would suggest that for example the operational margin should belong to the liquidity dimension, this would not accord with common accounting practice. Second, our sample is too small compared to the number of indicators that we consider in this study. Hence, results will not have known statistical properties. As a rule of thumb, Nardo et al. (2008) mention a 3:1 ratio of cases to variables, while others even speak of a 5:1 ratio. Because we only have between 30 and 38 cases each year and consider 39 indicators, performing a factor analysis would not result in useful statistics.

However, we still have to pay attention to interrelations between indicators (Nardo et al., 2005, 2008). If two indicators are highly correlated (i.e. a correlation higher than 0,92), there is a chance that they measure the same underlying construct. Table 3 for example shows that indicators A5 and A7 are highly correlated. As we do not want these highly correlated variables to dominate stock scores due to some sort of double counting, we have to make some adjustments (Jacobs et al., 2004). One possibility is to omit one of the two variables, as they measure almost the same concept. We have however chosen to give each variable a weight of 0,5 instead of 1, following Saisana (2012). Additionally, the squares in Table 3 represent the correlations between indicators in each dimension. If the dimensions have been chosen correctly, we should see that correlations are higher in the squares (i.e. indicators belonging to the same dimension are more correlated). Table 3 demonstrates that this is the case for most dimensions.

3.4.3 Grouping information on stocks

Cluster analysis is often used to group information on stocks based on their similarity on several individual indicators (Nardo et al., 2005, 2008). Our first goal is to gain some deeper insight into the structure of the dataset. Moreover, we perform this cluster analysis so we can reflect on the final results of our CI. The main advantage of cluster analysis is that we do not have to make any assumptions about the underlying distribution of the data (Norusis, 2011). Additionally, cluster analysis can also be executed for small datasets, as is the case in our study. Like Nardo et al. (2005, 2008), Norusis (2011), and Saisana (2012), we have decided to perform this step before normalizing the indicators.

We want to make a distinction between value and growth stocks, as is often done by fund managers. In our literature review, we already referred to the importance of this distinction (see Section 2.2). Value stocks will achieve low scores on dimension C (growth signals) and will perform well on dimension E (valuation). For growth stocks, exactly the opposite is true. As suggested for small datasets where the amount of clusters is known, we have selected *k*-means as our clustering method (Norusis, 2011). This analysis was performed in SPSS.

3.5 Normalization

Normalization is often required as not all indicators have the same measurement units (Nardo et al., 2005, 2008). Because numerous normalization methods are used in practice, it is important to select a normalization method that is suitable for the phenomenon that we are trying to measure. How robust are these methods against outliers? Do we want to keep the absolute level of information? Should we compare the companies with a benchmark? All these questions need to be addressed first if we want to justify the normalization method applied during the development of the CI. Additionally, the selection of the normalization scheme should take the properties of the data, such as highly skewed indicators, into account (Freudenberg, 2003; Nardo et al., 2005, 2008). In our case, we have chosen to select six different normalization methods, all with their own advantages and disadvantages. These methods will be briefly discussed in the remainder of this section. The goal is to investigate the influence of the normalization scheme on the final results of our CI. For a summary with frequently used normalization methods, we refer to Freudenberg (2003), Jacobs et al. (2004), Nardo et al. (2005, 2008), and Atsalakis and Valavanis (2009).

3.5.1 Ranking

Due to its simplicity, ranking is often utilised to normalize data (e.g. Jencks, Huff, & Cuerdon, 2003). The results are very easy to interpret and are not affected by outliers. However, this method is also highly criticized because it cannot evaluate performance in absolute terms as all information on levels is lost (Nardo et al., 2008). As an example, consider Table 4 below. Company C achieves the best score (26), closely followed by company A (25), leaving company B far behind (5). Although there is a huge gap between the scores of companies A/C and company B, this is not reflected in the ranking. Moreover, when we solely look at this ranking, we have no idea whether company B is heavily or only slightly underperforming the other companies.

	A1	A2	A3	A4	A5	A6	A7	B1	B2	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	D1	
A1	**1,00**	0,48	**0,59**	**0,62**	0,07	**0,69**	0,11	0,47	0,18	-0,09	-0,21	-0,19	-0,17	0,19	0,22	0,26	-0,26	0,19	-0,03	0,21	…
A2	0,48	**1,00**	**0,86**	0,25	-0,18	0,25	-0,18	0,15	0,03	-0,20	-0,23	-0,07	-0,18	-0,10	0,02	0,09	0,03	-0,13	-0,35	0,11	…
A3	**0,59**	**0,86**	**1,00**	0,46	-0,16	0,50	-0,18	0,36	0,06	0,02	-0,10	0,12	-0,04	0,11	0,17	0,24	0,19	-0,03	-0,11	0,19	…
A4	**0,62**	0,25	0,46	**1,00**	0,12	***0,95***	0,12	0,27	0,17	-0,06	-0,08	0,06	-0,07	0,21	0,42	0,18	-0,07	0,14	0,32	0,05	…
A5	0,07	-0,18	-0,16	0,12	**1,00**	0,18	***0,98***	0,04	0,21	0,02	-0,17	-0,31	0,03	-0,17	-0,01	-0,11	-0,28	0,02	0,18	-0,35	…
A6	**0,69**	0,25	**0,50**	***0,95***	0,18	**1,00**	0,19	0,49	0,20	0,02	-0,04	0,09	-0,06	0,30	**0,53**	0,15	-0,05	0,21	0,39	0,06	…
A7	0,11	-0,18	-0,18	0,12	***0,98***	0,19	**1,00**	0,07	0,26	0,02	-0,16	-0,35	-0,04	-0,12	0,05	-0,09	-0,34	0,07	0,20	-0,34	…
B1	0,47	0,15	0,36	0,27	0,04	0,49	0,07	**1,00**	-0,10	0,07	0,04	0,19	0,06	**0,51**	0,34	0,01	0,15	0,50	0,22	0,17	…
B2	0,18	0,03	0,06	0,17	0,21	0,20	0,26	-0,10	**1,00**	-0,12	0,04	-0,21	-0,49	-0,13	0,31	0,18	-0,22	-0,23	0,32	0,02	…
C1	-0,09	-0,20	0,02	-0,06	0,02	0,02	0,02	0,07	-0,12	**1,00**	**0,67**	0,37	0,35	0,50	0,14	-0,01	0,37	0,24	0,25	0,12	…
C2	-0,21	-0,23	-0,10	-0,08	-0,17	-0,04	-0,16	0,04	0,04	**0,67**	**1,00**	0,22	-0,04	0,40	0,04	-0,17	0,41	0,02	0,22	0,21	…
C3	-0,19	-0,07	0,12	0,06	-0,31	0,09	-0,35	0,19	-0,21	0,37	0,22	**1,00**	**0,61**	**0,59**	0,46	0,01	**0,72**	**0,51**	0,27	**0,56**	…
C4	-0,17	-0,18	-0,04	-0,07	0,03	-0,06	-0,04	0,06	-0,49	0,35	-0,04	**0,61**	**1,00**	0,29	0,03	0,12	**0,52**	**0,55**	0,04	0,25	…
C5	0,19	-0,10	0,11	0,21	-0,17	0,30	-0,12	**0,51**	-0,13	0,50	0,40	**0,59**	0,29	**1,00**	**0,56**	0,09	0,49	**0,79**	0,49	**0,59**	…
C6	0,22	0,02	0,17	0,42	-0,01	**0,53**	0,05	0,34	0,31	0,14	0,04	0,46	0,03	**0,56**	**1,00**	0,18	0,22	0,41	**0,61**	0,28	…
C7	0,26	0,09	0,24	0,18	-0,11	0,15	-0,09	0,01	0,18	-0,01	-0,17	0,01	0,12	0,09	0,18	**1,00**	0,11	0,21	0,07	0,24	…
C8	-0,26	0,03	0,19	-0,07	-0,28	-0,05	-0,34	0,15	-0,22	0,37	0,41	**0,72**	**0,52**	0,49	0,22	0,11	**1,00**	0,37	0,18	0,41	…
C9	0,19	-0,13	-0,03	0,14	0,02	0,21	0,07	0,50	-0,23	0,24	0,02	**0,51**	**0,55**	**0,79**	0,41	0,21	0,37	**1,00**	0,29	**0,51**	…
C10	-0,03	-0,35	-0,11	0,32	0,18	0,39	0,20	0,22	0,32	0,25	0,22	0,27	0,04	0,49	**0,61**	0,07	0,18	0,29	**1,00**	0,07	…
D1	0,21	0,11	0,19	0,05	-0,35	0,06	-0,34	0,17	0,02	0,12	0,21	**0,56**	0,25	**0,59**	0,28	0,24	0,41	**0,51**	0,07	**1,00**	…
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
Weights within principle	1	1	1	**0.5**	**0.5**	**0.5**	**0.5**	1	1	1	1	1	1	1	1	-1	1	1	1	1	

Table 3: Extract of the correlation matrix for the set of indicators (see Appendix B) of year 2012. Cases in bold indicate a correlation larger than 0,5. Cases showing high correlation (larger than 0,92) are in italic.

	Indicator value	Ranking
Company A	25	2
Company B	5	3
Company C	26	1

Table 4: Example of ranking as normalization method.

The equation (2) that is applied for ranking is very straightforward:

$$I_{qs} = \text{Rank}\left(x_{qs}\right) \tag{2}$$

3.5.2 Standardization (Z-scores)

Standardization normalizes all indicators to a common scale where the mean is zero and the standard deviation is one (Saisana et al., 2005; Nardo et al., 2005, 2008). To achieve this, equation (3) is utilised:

$$I_{qs} = \left(\frac{x_{qs} - \bar{x}_q}{\sigma_q}\right) * Direction_q \tag{3}$$

Hence, indicators with extreme values have a larger effect on the final score of the composite measure. Also note that by taking the direction of the indicator into account in the equation, all indicators will 'point in the same direction'.

3.5.3 Min-Max

When using Min-Max normalization, all indicators receive values between zero and one by subtracting the minimum value of the indicator from the actual value of each stock, while dividing it by the range of indicator values (Saisana et al., 2005; Nardo et al., 2005, 2008). Equation (4) represents this normalization method. Here, $\min(x_q)$ represents the minimum value of indicator q, while $\max(x_q)$ represents the maximum value of indicator q.

$$I_{qs} = \left(\frac{x_{qs} - \min(x_q)}{\max(x_q) - \min(x_q)}\right) * Direction_q + 0{,}5 * \left(1 - Direction_q\right) \tag{4}$$

Hence, when the normalized value is zero, stock s is the worst performer on indicator q. The opposite is true when the normalized value is one. The main problem with Min-Max normalization is that outliers heavily influence the outcomes. However, as all indicator values are within an identical range of [0, 1], this normalization method simplifies the interpretation of our results. Moreover, because no negative values occur, Min-Max is often applied when statistical techniques such as Data Envelopment Analysis have to be performed (Nardo et al., 2008; Saisana, 2012).

3.5.4 Distance to a reference

The distance to a reference normalization measures the relative position of an indicator to a chosen reference point. This reference point could be the top or worst performer, as well as the mean or median value of an indicator. Furthermore, an external benchmark can also be utilised. In this study, we have chosen to use the mean value of the indicator as our reference. Hence, stocks achieving better scores than the average will receive a value larger than zero. The equation (5) that we have applied is:

$$I_{qs} = \left(\frac{x_{qs} - \bar{x}_q}{\left|\bar{x}_q\right|}\right) * Direction_q \tag{5}$$

We have chosen not to select the top or worst performer as a benchmark as the presence of outliers could heavily influence the results.

3.5.5 Categorical scales

Categorical scales can be both numerical and qualitative and are not sensitive to outliers. A qualitative scale is for example: 'high growth', 'medium growth' or 'low growth'. In our study, we have chosen to calculate the percentile of every indicator value, as suggested by Nardo et al. (2005, 2008). This percentile is then used as the new normalized value, which is again very easy to interpret (see equation (6)). However, as was the case with ranking, some level of absolute information is lost. For another example of percentile ranking normalization, we refer to Prathap (2012) who adopts this method to develop new energy indicators.

$$I_{qs} = \text{Percentile}\left(x_{qs}\right) * Direction_q + 0{,}5 * \left(1 - Direction_q\right) \tag{6}$$

3.5.6 Indicators above or below the mean

A last normalization method that is considered in our analysis is displayed in equation (7). Here, we assign a value of one to every observation that outperforms the mean by 20%. If the observation performs worse than -20% of the mean, it obtains a value of minus one. All observations within the interval of -20% and 20% around the mean, get a value of zero. These boundaries are arbitrarily chosen.

$$I_{qs} = \begin{cases} 1 * Direction_q & \text{if } 0{,}2 < r \\ 0 & \text{if} - 0{,}2 \leq r \leq 0{,}2 \\ -1 * Direction_q & \text{if } r < -0{,}2 \end{cases} \qquad \text{with } r = \frac{x_{qs} - \bar{x}_q}{|\bar{x}_q|} \tag{7}$$

When applying this normalization method, all information on absolute levels is lost. However, it is not sensitive to outliers and the interpretation is very straightforward.

3.6 Weighting and aggregation

Probably the most important and influential step in the construction of our CI is the weighting and aggregation of the different individual indicators. Regardless of which method we choose to assign weights, it always remains a subjective choice (Nardo et al., 2005, 2008). As discussed in Section 2.4, this is one of the main critiques of CI's. Because it is not possible to completely eradicate this subjectivity, we will choose different weighting schemes that involve no additional subjective judgements such as for example expert opinions. Hence, we still address this problem, although no complete solution can be offered. As a result, we have selected the following weighting schemes: equal weighting (Section 3.6.1), benefit of the doubt (BOD) (Section 3.6.2), random weights (Section 3.6.3) and the Copeland rule (Section 3.6.4). In all these methods, subjectivity is reduced to a minimum. We will mainly focus on the BOD and Copeland rule approaches as well as on a combination of both methods.

Other weighting schemes such as budget allocation or analytic hierarchy process could have been applied but they induce much more subjectivity (Hermans, Van den Bossche, & Wets, 2008). Additionally, other problems make it more difficult to use these methods. According to Nardo et al. (2005, 2008), budget allocation is very suitable when experts have to give their opinion about 10 to 12 indicators. It becomes rather unmanageable if one has to divide 100 points over nearly 40 indicators, as is the case in our study. The same arguments hold for the analytic hierarchy process. This however does not mean that these methods could not achieve good results in terms of return. Another weighting scheme that is often utilised when constructing CI's is factor analysis. Unfortunately, the same arguments that were given in Section 3.4.2 are still valid whereby factor analysis is not suitable for our data. If for example a larger financial market, such as the New York Stock Exchange (NYSE) would have been analysed, factor analysis could be more attractive as a weighting method.

After the selection of an appropriate weighting scheme, we still have to aggregate the individual indicators. In order to be able to discuss the influence of different aggregation systems we subsequently apply linear or arithmetic aggregation (Section 3.6.5), geometric aggregation (Section

3.6.6) and summation of rankings (Section 3.6.7). Again, this is only a small selection of all aggregation methods that are available. For a more in-depth review, we refer to Nardo et al. (2005, 2008).

3.6.1 Equal weighting

An often occurring misunderstanding is that equal weighting implies assigning no weights at all. Most of the time equal weights means that all variables are worth the same in the CI (Nardo et al., 2008). Today, most CI's rely on this weighting scheme (e.g. CEC, 2004; Yale, 2005), despite Hermans et al. (2008) conclude that this method should only be utilised when no other weighting method yields valid results or when all indicators are uncorrelated or highly correlated. Moreover, indicators are often grouped into dimensions, as is the case in our study (see Section 3.4.2). Hence, if we assign equal weights to every individual indicator this will induce unequal weighting on the dimension level (Hermans et al., 2008). Dimensions with more individual indicators will unintendedly receive higher weights in the CI. As this could result in an unbalanced structure, we impose equal weights across dimensions and equal weights within each dimension as shown by equation (8). Additionally, another problem occurs when individual indicators are highly correlated. Indeed, this can lead to double counting in our CI, again inducing an imbalanced structure. In Section 3.4.2 we have however adjusted the weights of highly correlated indicators whereby the problem of double counting is mostly solved (Freudenberg, 2003; Nardo et al., 2008).

$$w_q = \left(\frac{1}{D}\right) * \left(\frac{1}{N_d}\right) \qquad ; \forall q \in D_d \tag{8}$$

In equation (8), w_q is the weight of indicator q in the CI. Clearly, indicator q needs to be part of a dimension d as depicted in the second part of equation (8). As an example, consider the following data from our study:

Indicator	Dimension	1/*D*	1/N_f	W_{F3}
F3	F	1/7	1/6	1/42

Table 5: Example of equal weighting.

In Table 5 we determine the weight of indicator F3, which is a member of dimension F that consists of six individual indicators. By multiplying 1/7 and 1/6 we obtain a weight of 1/42 for indicator F3 in the CI. Note that if indicator A4 was chosen as an example, we had to adjust the weights due to its high correlation with indicator A6 (see Section 3.4.2).

3.6.2 Benefit of the doubt analysis (BOD)

BOD or benefit of the doubt analysis is based on data envelopment analysis (DEA), which was developed by Charnes, Cooper and Rhodes (1978). It is used for evaluating the relative efficiency of decision making units, or stocks in our case, where efficiency is calculated as the ratio of the weighted sum of outputs to the weighted sum of inputs (Nardo et al., 2005, 2008; Hermans et al., 2008). While calculating the relative efficiency of stocks, weights are determined in such a manner that the efficiency score for each stock is maximized. Hence, the DEA-model chooses optimal weights, satisfying all restrictions. No other set of weights can be found that would yield a higher composite score (Hermans et al., 2008). Because the results are only influenced by the stocks in the dataset, the method applied here is solely about relative comparisons, which is particularly relevant for stock picking (see Section 2.3). Abad et al. (2004) for example investigate whether DEA can be utilised to rank stocks according to their fundamental variables. Halkos and Tzemeres (2012) on the other hand, apply DEA to evaluate different industries. Moreover, relative ranking has already been used extensively to compare the efficiency of banks, given their fundamental ratios (e.g. Halkos & Salamouris, 2004; Ho & Wu, 2006; Avkiran & Morita, 2010; Avkiran, 2011). According to Web of Science, there were already more than 170 papers in 2011 that combined DEA and banks (Avkiran, 2011).

Based on the general DEA model for indexes, we can translate the original input oriented DEA-model to the CI's framework, as proposed by Cherchye et al. (2006). In this model, each indicator is referred to as an output while inputs are no longer considered. The basic linear programming problem choosing the optimal weights for each stock can be written as follows:

$$CI_s = \max_{w_{qs}} \sum_{q=1}^{Q} I_{qs} * w_{qs}$$

$$\text{Subject to} \tag{9}$$

$$\sum_{q=1}^{Q} I_{qk} * w_{qk} \leq 1 \qquad \forall\, k = 1, \dots, S$$

$$w_{qk} \geq 0 \qquad \forall\, q = 1, \dots, Q; \quad \forall\, k = 1, \dots, S$$

The result of the linear problem (9) is a set of optimal weights w_{qs} for every indicator q and stock s that maximizes the CI score CI_s for each stock. The first constraint, often called the normalization constraint, states that no stock can achieve a composite score higher than one. This boundary is of course arbitrary but by limiting the CI value to one, future interpretations are a lot easier. CI scores will thus range between zero and one, with higher scores indicating a better performance. In our study, a good performance is indicated by sound fundamental variables resulting in a higher expected return (see Section 2.1). Hence, a stock with a composite score of one is the top performer of the group. The second constraint states that all weights have to be non-negative.

This method is very different from other weighting schemes for two reasons. First, a separate maximization model is constructed for every stock which results in stock-specific weights, opposite to a uniform set of weights for other weighting schemes. Second, in most BOD models, no normalization is needed as DEA can handle raw values (Cherchye et al., 2006; Hermans et al., 2008). However, as there are many problematic cases in our data, such as negative values and outliers, we still normalized our data with the Min-Max method (Cherchye et al., 2006). This method is very suitable for DEA as all negative values are filtered from the data while still keeping the absolute level of information. Sarkis (2002), Nardo et al. (2008), and Saisana (2012) also suggest performing a Min-Max or mean normalization prior to the DEA. Additionally, Sarkis (2002) and Saisana (2012) mention that negative and zero values induce difficulties when the DEA is conducted. In order to solve this problem, they make some suggestions. One suggestion is to add a significantly high constant to all values so no negative values occur anymore. However, as we often work with percentages, this method could lead to unbalanced results as a vast amount of absolute information will be lost. Again, this is a justification for the normalization prior to the DEA. Another suggestion is to adjust all zero values to a very small constant after normalization. Therefore, we have chosen to replace all zero values by 0,01 in our data.

As it becomes difficult to compare stock rankings based on different sets of weights, we will mainly consider models in which the sum of all composite scores has to be maximized or models in which cross-efficiency will be compared (Cherchye et al., 2006; Hermans et al., 2008; Nardo et al., 2008; Saisana, 2012). In our study, we subsequently incorporate four different extensions to the basic DEA model in equation (9). Additionally, our model is slightly more complicated as we will perform a two-stage DEA. We first start by explaining this concept in more detail. Second, we present the four extensions that will be made to the basic model.

The idea of a two-stage DEA is very straightforward and is easy to implement. We have based our methodology on Abad et al. (2004). As we have two hierarchical levels in our data, namely dimensions and individual indicators, we perform a DEA on each level. First we execute a DEA on the indicator level, within each dimension. This means that we try to maximize the dimension score for every stock, assigning higher weights to indicators on which the stock achieves a better score, while satisfying all constraints. The result is a set of optimal dimension scores for every stock. Hence, if there are no restrictions, the model will generally simply assign the maximum weight to the highest indicator score. Second, we utilise these optimal dimension scores as input in a second DEA, which in turn assigns optimal weights to every dimension. By applying a two-stage approach, visibility is guaranteed during the whole process as it will be much easier to demonstrate which dimension and which indicators are mainly responsible for the CI score. Another advantage of this method is that we can implement the output of the first stage in other weighting schemes, such as the Copeland rule (see Section 3.6.4). Thus, model (9) can be rewritten as follows:

Stage 1:

$$DI_{ds} = \max_{w_{ps}} \sum_{p \in D_d} I_{ps} * w_{ps}$$

$$\text{Subject to} \qquad (10)$$

$$\sum_{p \in D_d} I_{pk} * w_{pk} \leq 1 \qquad \forall\, k = 1, \dots, S; \quad \forall p \in D_d$$

$$w_{pk} \geq 0 \qquad \forall\, k = 1, \dots, S; \quad \forall p \in D_d$$

Stage 2:

$$CI_s = \max_{v_{ds}} \sum_{d=1}^{D} DI_{ds} * v_{ds}$$

$$\text{Subject to}$$

$$\sum_{d=1}^{7} DI_{dk} * v_{dk} \leq 1 \qquad \forall\, k = 1, \dots, S; \quad \forall\, d = 1, \dots, D$$

$$v_{dk} \geq 0 \qquad \forall\, k = 1, \dots, S; \quad \forall\, d = 1, \dots, D$$

In the first stage, DI_{ds} or the optimal dimension score for dimension *d* and stock *s* is calculated by assigning weights w_{ps}, with *p* (*p*=1,…,*Q*) an indicator in dimension *d*, to every normalized value I_{ps}. For example, the first dimension consists of seven indicators. Therefore, we maximize the weighted sum of all seven indicators in this dimension. This process is then repeated for every dimension and every stock which results in a *S*x7 matrix. This matrix is subsequently implemented as input in the second stage where basically the same process is repeated. Hence, we calculate the composite score by maximizing the weighted sum of the optimal dimension scores DI_{ds} (retrieved from stage 1) and their optimal weights v_{ds} for each dimension *d* and stock *s*. Model (10) will serve as our basis scenario from which we will depart when constructing additional models. We will list these models briefly in the remainder of this section.

Model 1

The first model that we consider is exactly the same as model (10), i.e. with no restrictions on the weights, except the non-negativity constraint. However, if we apply this classical DEA method (i.e. the optimal score for every stock), about 90% of all stocks achieves the maximum score of one (for a more in-depth discussion of this phenomenon we refer to Rogge, 2012). Most stocks perform well on at least one indicator whereby the maximum weight is assigned to this indicator. As we cannot conclude anything from a ranking where almost every stock has the same score, we have chosen to measure the mean cross-efficiency of the CI scores for every stock (Saisana, 2012). This means that we let every stock choose his optimal weights and when optimal weights have been determined, we calculate the average score across all stocks. Based on this method, every set of optimal weights will contribute (1/*S*)*100% to the final composite score.

Model 2
In our second model, we introduce restrictions on weights as we do not want all weight to be given to one single indicator. This step normally involves gathering expert opinions about the boundaries of the weight for every indicator. As we want to reduce subjectivity in the construction of our CI, we have chosen not to use expert opinions. Hence, we have developed some absolute restrictions that should enforce weights to be larger than zero. In the first stage we added the following restriction (11):

$$\frac{1}{3*N_d} \leq w_{pk} \leq \frac{3}{N_d} \qquad \forall\, d = 1, \dots, D; \quad \forall\, p \in D_d; \quad \forall\, k = 1, \dots, S \tag{11}$$

As there are for example seven indicators in the first dimension, every indicator's weight is restricted to [1/21; 3/7] in this dimension. Hence, we still leave room for large variations between weights, while avoiding that one indicator will dominate the complete dimension score (Cherchye et al., 2006).

Moreover, based on the literature review performed in Section 2.2 we have an idea about the relative importance of each dimension. We can use this information in order to demonstrate that for example dimension A is more important than dimension B, etc. Hence, we do not impose numerical bounds on weights, only softer ordinal bounds. Therefore, we have added the following ordinal restriction in the second stage:

$$v_{4k} \leq v_{6k} \leq v_{2k} \leq v_{1k} \leq v_{7k} \leq v_{5k} \leq v_{3k} \qquad \forall\, k = 1, \dots, S \tag{12}$$

Again, the mean cross-efficiency measure will be calculated. In restriction (12), v_{1k} corresponds to dimension A, v_{2k} to dimension B, and so on. By imposing restrictions (11) and (12), we are more confident that we will be able to highlight differences in our stock's ranking. However, Cherchye et al. (2006) propose to impose restrictions on pie-shares (i.e. the product of the indicator value and the indicator weight) instead of on weights directly. Hence, we introduce restrictions on pie-shares in the next model so the importance of every subindicator can be evaluated.

Model 3
In our third model, we impose restrictions on pie-shares rather than on weights themselves. This approach allows for a clear representation of our composite scores as 'pies'. Here, the size of the pie represents the CI score of a particular stock, while the pie-shares indicate the contribution of every indicator/dimension to its composite score. We implement two different pie-restrictions for each of the two DEA stages (Cherchye et al., 2006). In the first stage, we impose a proportional share restriction (13), while for the second stage we will again impose an ordinal restriction (14) on top of the problem that was formulated in (10). We also state that no dimensions' pie share can be smaller than 1/21 (i.e. 1/3*D*) times the sum of all pie shares of a particular stock (i.e. the size of the pie of the CI score). Hence, we ensure that all dimensions will be represented in the CI. We have also incorporated a limit to a dimensions pie-share as otherwise, the results would be skewed towards stocks performing extremely well on dimension C.

$$\frac{1}{3*N_d} \leq \frac{I_{pk} * w_{pk}}{\sum_{p \in D_d} I_{pk} * w_{pk}} \leq \frac{3}{N_d} \qquad \forall\, d = 1, \dots, D; \quad \forall\, p \in D_d; \quad \forall\, k = 1, \dots, S \tag{13}$$

$$\begin{aligned}\frac{1}{21} * \sum_{d=1}^{D} DI_{dk} * v_{dk} &\leq v_{4k} * DI_{4k} \leq v_{6k} * DI_{6k} \leq v_{2k} * DI_{2k} \leq v_{1k} * DI_{1k} \leq v_{7k} * DI_{7k} \leq v_{5k} * DI_{5k} \\ &\leq v_{3k} * DI_{3k} \leq \frac{3}{7} * \sum_{d=1}^{D} DI_{dk} * v_{dk} \quad \forall\, d = 1, \dots, D; \quad \forall\, p \in D_d; \quad \forall\, k = 1, \dots, S\end{aligned} \tag{14}$$

Note that we no longer calculated cross efficiency measures in this model as this would not yield any logical results.

Model 4
In our last model, we abandon all restrictions on weights in the first stage, except for the non-negativity constraint. This time, we make an adjustment in our objective function itself. As critics would argue that it is unfair to rank stocks based on different sets of weights, we introduce a model where the sum of the indicator values of all stocks in a specific year has to be maximized, i.e. based on the same set of weights for all stocks (Hermans et al., 2008). In models 1 and 2, we have provided an answer to this problem by measuring cross-efficiency, rather than the classical DEA diagonal. As we no longer consider different weights for different stocks, we have to rewrite model (10). Note that again, no subjective choices have to be made during this process as neither expert opinions nor absolute boundaries have to be determined. In the second stage however, we have still chosen to add a soft ordinal restriction similar to restriction (12). Provided that we want to maximize the sum of the composite scores while keeping one set of weights, the following model (15) can be constructed, with w_p the weight of indicator p ($p \in D_d$) and v_d the weight of dimension d.

Stage 1:

$$\sum_{s=1}^{S} DI_{ds} = \max_{w_p} \sum_{s=1}^{S} \sum_{p \in D_d} I_{ps} * w_p$$

Subject to (15)

$$\sum_{p \in D_d} I_{pk} * w_p \leq 1 \qquad \forall\, k = 1, \dots, S; \quad \forall\, p \in D_d$$

$$w_p \geq 0 \qquad \forall\, p \in D_d$$

Stage 2:

$$\sum_{s=1}^{S} CI_s = \max_{v_d} \sum_{s=1}^{S} \sum_{d=1}^{D} DI_{ds} * v_d$$

Subject to

$$\sum_{d=1}^{D} DI_{dk} * v_d \leq 1 \qquad \forall\, k = 1, \dots, S; \quad \forall\, d = 1, \dots, D$$

$$\tfrac{1}{21} \leq v_{4k} \leq v_{6k} \leq v_{2k} \leq v_{1k} \leq v_{7k} \leq v_{5k} \leq v_{3k} \leq \tfrac{3}{7} \qquad \forall\, k = 1, \dots, S$$

Like, Cherchye et al. (2006) and Hermans et al. (2008) we have opted for the DEA Excel Solver program in order to solve all four models. The results will be presented in Section 4.2.2.

3.6.3 Random weights

Assigning random weights to indicators is not often used in practice. However, when deriving thousands of portfolios based on different sets of weights, it becomes very interesting to investigate the changes in composite rankings. If for example, a particular stock's ranking remains in the range of [1, 4] for every set of weights, one can conclude that this stock should certainly be a top performer. Additionally, when the range of rankings is extremely large, one can say that the choice of weighting scheme has a huge influence on the CI. Moreover, if we can demonstrate that ranking intervals remain roughly the same when other normalization methods were applied, we could suggest that the choice of normalization method is of less importance. Therefore, we will mainly utilise the results of this random weighting strategy as a reference study in order to initiate a discussion in Section 4.

The methodology for constructing random sets of weights is very simple. First, a number from a uniform random distribution is determined for every indicator. Second, we rescale these numbers (i.e. weights) in such a way that they sum up to one (Saisana, 2012). Third, we aggregate the indicators with their corresponding weights using both arithmetic and geometric aggregation (see Section 3.6.5 and 3.6.6). Equation (16) summarizes this method for arithmetic aggregation and equation (17) for geometric aggregation. In both equations, δ_q is a uniformly distributed variable between 0 and 1 for every indicator q.

$$CI_s = \sum_{q=1}^{Q}\left(I_{qs} * \frac{\delta_q}{\sum_{q=1}^{Q}\delta_q}\right) \qquad \forall\ s = 1,\ldots,S \tag{16}$$

$$CI_s = \prod_{q=1}^{Q}\left[I_{qs}^{\left(\frac{\delta_q}{\sum_{q=1}^{Q}\delta_q}\right)}\right] \qquad \forall\ s = 1,\ldots,S \tag{17}$$

We repeat this process 10.000 times, resulting in 10.000 sets of weights and composite rankings. After these Monte Carlo simulations, we calculate ranking intervals, composite scores and returns, in addition to median values of these outcomes. These measures will serve as a reference point for all other weighting schemes and will allow us to assess their performance. Again, we have accounted for highly correlated indicators by adjusting their weights (see Section 3.4.2). This method was repeated for different normalization methods.

3.6.4 Copeland rule

Although the Copeland rule is not frequently displayed in CI manuals, it is certainly a useful weighting technique. The Copeland rule starts by comparing stocks pairwise (Saisana, 2012). For each comparison, all indicator weights w_q favouring stock 1 before stock 2 are added up (see equation (18)). The main purpose is to collect 'evidence' that stock 1 is better than stock 2. As we need to compare all stocks, *S**(*S*-1) comparisons have to be made. Hence, the result is an *S*x*S* matrix, often called the outranking matrix where the main diagonal (i.e. the comparison between a stock and itself) is put to zero by definition (Nardo et al., 2008; Saisana, 2012).

$$C_{sk} = \sum_{q=1,\ s\neq k}^{Q} \frac{\left[w_q * \left(1 + Direction_q\right) * \mathrm{Sign}\left(x_{qs} - x_{qk}\right)\right]}{2} \qquad \forall\ s,k = 1,\ldots,S \tag{18}$$

Subsequently, we apply the Copeland rule to this outranking matrix. When the majority of indicators prefers stock *s* before stock *k* (*k*=1,…,S), a value of +1 is assigned to this case (i.e. the value in the outranking matrix is larger than 0,5). The opposite is true when the majority of indicators indicates that stock *s* performs worse than stock *k*, receiving a value of -1 (i.e. the value in the outranking matrix is smaller than 0,5). If a tie between two stocks occurs, a score of zero is assigned to this case (i.e. the value in the outranking matrix equals 0,5). We obtain the Copeland composite score by adding all scores in a row.

Some important remarks can be made about the Copeland rule. First, one of the main reasons why we have chosen to include this method in our study is because, again, it reduces the subjectivity in our decision process as it does not involve any data normalization nor does it require expert opinions. Moreover, it compares indicator scores in a relative way, which is highly recommended in the context of stock selection (Edirisinghe & Zhang, 2007; see Section 2.3). Second, Saisana (2012) suggests that in general, some compensability is desired at lower hierarchical levels (i.e. indicators within dimensions), and less desired at higher hierarchical levels (i.e. dimensions). Therefore, the Copeland rule will be mostly applied to aggregate dimension levels and to a lesser extent to aggregate indicators within each dimension in order to leave some room for compensability. Hence, we think that a combination of BOD and the Copeland rule would optimize our CI's. A third important remark is that critics argue that too much emphasis is put on the quantity of times that a stock outperforms another stock rather than on their magnitudes (Saisana, 2012). However, when the Copeland rule can be combined with another method, such as BOD, this issue is partly addressed.

3.6.5 Linear aggregation

According to Nardo et al. (2008), linear or arithmetic aggregation is most often performed in practice. This aggregation method corresponds with the following equation (19):

$$CI_s = \sum_{q=1}^{Q} I_{qs} * w_q \qquad \forall\, s = 1, \dots, S \tag{19}$$

$$\text{with} \sum_{q=1}^{Q} w_q = 1 \text{ and } 0 \le w_q \le 1 \qquad \forall\ q = 1, \dots, Q$$

Due to its simplicity, this method is frequently subject to criticism (see Nardo et al., 2008 for a more in-depth discussion). One of the main critics is that linear aggregation is preferential independent, which means that all bad scores can be compensated by a good score on one or more indicators. This implies that neither synergy nor conflicts may occur in the CI. However, this issue seems less relevant for the selection of stocks as we think that in the context of FA, a certain degree of compensability is desirable. Despite this, there is always a chance that stocks performing high on most indicators will achieve synergy effects which can lead to higher returns.

3.6.6 Geometric aggregation

It could be that investors dislike the idea that a bad score on one indicator can be compensated by a good score on another one. Nardo et al. (2008) mention that weights are often seen as importance coefficients rather than as trade-offs between indicators. This implies that greater weight should be given to indicators which are considered more significant in the context of the particular CI (Freudenberg, 2003).

$$CI_s = \prod_{q=1}^{Q} \left(I_{qs}\right)^{w_q} \qquad \forall\, s = 1, \dots, S \tag{20}$$

$$\text{with} \sum_{q=1}^{Q} w_q = 1 \text{ and } 0 \le w_q \le 1 \qquad \forall\ q = 1, \dots, Q$$

Notation (20) can be linearized by taking logarithms such that the model becomes tantamount to the linear aggregation model in order to utilise this aggregation method for the BOD weighting. However, as Cherchye et al. (2006) discuss, the interrelationship of a logged form of the geometric aggregation has not yet been analysed. Therefore, like Cherchye et al. (2006), Nardo et al. (2008), and Saisana (2012), we will only apply the linear aggregation method when we want to aggregate weights resulting from the BOD approach. Additionally, like Saisana (2012) did in his analysis, we have also made some adjustments to the data in order to simplify the logarithmic calculations (i.e. adding a constant to all indicators so that neither negative nor zero values would occur anymore).

3.6.7 Summation of ranks

Undoubtedly, one of the simplest techniques is to calculate the rank on each indicator and add those ranks together for each stock (Nardo et al., 2008). The following equation (21) is used to calculate the CI:

$$CI_s = \sum_{q=1}^{Q} \text{Rank}(x_{qs}) \qquad \forall\ s = 1, \dots, S \tag{21}$$

$\text{Rank}(x_{qs})$ is equivalent to the Borda-ranking method (Nardo et al., 2008; Saisana, 2012), where the Borda-rank equals the amount of stocks considered minus the rank of a particular stock on a specific indicator. Equation (21) can also be adapted so the summation of indicators that are below and above the mean can be calculated (see Nardo et al., 2008; see Section 3.5.6).

3.7 Uncertainty and sensitivity analysis

Different methods are available to assess the robustness of CI's (Saltelli et al., 2008). We have tried to limit subjective choices during the construction of the CI, yet several subjective judgements have insurmountably slipped in. Uncertainties in our CI are mostly linked to the following factors:

(1) Including or excluding different indicators in the CI (e.g. adding Beta to the model);
(2) The transformation of indicators (e.g. dealing with outliers by Box-Cox transformations or winsorization);
(3) The normalization scheme that was chosen (e.g. Min-Max, categorical, ranking, distance);
(4) The selection of a weighting scheme (e.g. equal weights, BOD, Copeland rule, random weights);
(5) The level of aggregation (e.g. indicator or dimension level);
(6) The choice of aggregation method (e.g. arithmetic, geometric).

All these subjective choices can influence stocks' scores in the CI. Therefore, we should explore the influence of the decisions we have made during the entire process before interpreting the results (Cherchye et al., 2006). Of course, these sources of uncertainty are inherent to the construction of CI's. Most techniques for performing an uncertainty, sensitivity or robustness analysis are based on Monte Carlo simulations, variance based techniques or visualization (Saisana et al., 2005; Nardo et al., 2005, 2008; Saltelli et al., 2008). In our research we will mainly focus on visualization, incorporating both basic simulations and variations in rankings. Yet, literature presents numerous extensions and other alternative methods that are beyond the scope of this paper.

The average shift in stock's rankings (i.e. $\overline{R}$) is particularly of interest when performing a robustness analysis (see equation (22); Saisana et al., 2005; Saltelli et al., 2008).

$$\overline{R} = \frac{1}{S}\sum_{s=1}^{S}\left|\mathrm{Rank}_{ref}(CI_s) - \mathrm{Rank}(CI_s)\right| \qquad (22)$$

Here, every stock is compared to a reference ranking. As no clear reference is present in our study, we have decided to repeat the calculation of $\overline{R}$ until every ranking has been used as a reference. This results in an unbiased measure. Note that almost the same measure can be calculated in order to assess the differences in ranking for every stock across all different sources of uncertainty (i.e. $\overline{R_s}$ and summation over the amount of CI's).

Additionally, differences in CI values (see equation (23)) are also an important subject of this analysis (Saisana et al., 2005; Saltelli et al., 2008). Why is stock A outperforming stock B? Can we explain this by comparing the values of each indicator (e.g. by looking at the pie-shares of the third BOD model) or is the difference in rank influenced by one of the aforementioned sources of uncertainty? Indeed, we do not want our choices to be too influential when composite scores are derived.

$$D_{AB} = \sum_{q=1}^{Q}(I_{qA} - I_{qB}) * w_q \qquad (23)$$

Using equations (22) and (23) as a starting point, we can conduct Monte Carlo simulations, resulting in (trimmed) uncertainty intervals, an in-depth analysis of rankings and a clear explanation of the influence of each choice (Saisana et al., 2005; Nardo et al., 2008, Cherchye et al., 2006; Saltelli et al., 2008). Moreover, we can also incorporate these results as input in a regression model where the difference in ranking is the variable that has to be explained, while different dummy variables assess the influence of all the decisions made during the construction of our CI. Additionally we will add a very interesting extension to our CI by making it possible for investors to implement their own predictions in the model. Are we for example expecting bull (i.e. increasing) or bear (i.e. decreasing) markets? One way to account for investors' expectations is by adding beta (i.e. the correlation between the stock and the market) to our model, in turn leading to changes in rankings.

As mentioned earlier, variance-based techniques are often used when performing a sensitivity analysis (e.g. Nardo et al., 2005, 2008; Cherchye et al., 2006). The main advantage of these variance-based techniques is that they are 'model free', which is necessary when non-linear models arise due to several layers of uncertainty (Nardo et al., 2008). One example is the first order sensitivity index, represented by equation (24). In this equation, Y is the overall shift in stocks ranking with respect to a benchmark ranking (here equal weights with Min-Max normalization and arithmetic aggregation), $V(Y)$ is the variance of Y, and X_i is the source of uncertainty that we are analysing.[1] Normally, one should also take interaction effects into account, in addition to the higher order sensitivity measure in equation (24).

$$S = \frac{\mathrm{V}_{X_i}(E_{X_{-i}}(Y|X_i))}{\mathrm{V}(Y)} = \frac{V_i}{\mathrm{V}(Y)} \tag{24}$$

As we have decided not to change the set of input variables, the amount of simulations is limited in our analysis, making variance-based sensitivity measures less meaningful. Hence, we will not address these techniques in much detail as this is beyond the scope of our research. The interested reader is referred to Jacobs et al. (2004), Saisana et al. (2005), Nardo et al. (2005, 2008), Cherchye et al. (2006), and Saltelli et al. (2008).

4 Results & Discussion

In this section, we start by presenting the results of the uncertainty and sensitivity analysis. The following section will highlight some key results of our research with a clear focus on visualization techniques. Due to the vast amount of results that are available, we will already discuss them in more detail when the results are presented. Additionally, we will conclude Section 4 with a general discussion about the key results.

4.1 Uncertainty and sensitivity analysis

4.1.1 Uncertainty intervals

Based on 60.000 simulations, varying normalization methods, weights (randomly constructed) and aggregation systems, while keeping the indicators used constant, one can get a general overview of the influence of all these sources of uncertainty (Saltelli et al., 2008). A graphical representation is presented in Figure 3.[2] Figure 3 shows that regardless of weights, normalization and aggregation, Melexis clearly outperforms Hamon. However, it is more difficult to decide whether Sipef should achieve a higher rank than Solvay as there is some overlap in the stocks' ranking interval.

[1] See Nardo et al. (2008) for the complete derivation of this equation.

[2] In order to guarantee the clarity and minimize the size of this paper, we have chosen to take some extractions from the results so a discussion can be initiated. If one would like access to the Excel-files with complete calculations, composite scores and complete rankings, one can obtain them by sending a request to michael.hamelryck@student.kuleuven.be.

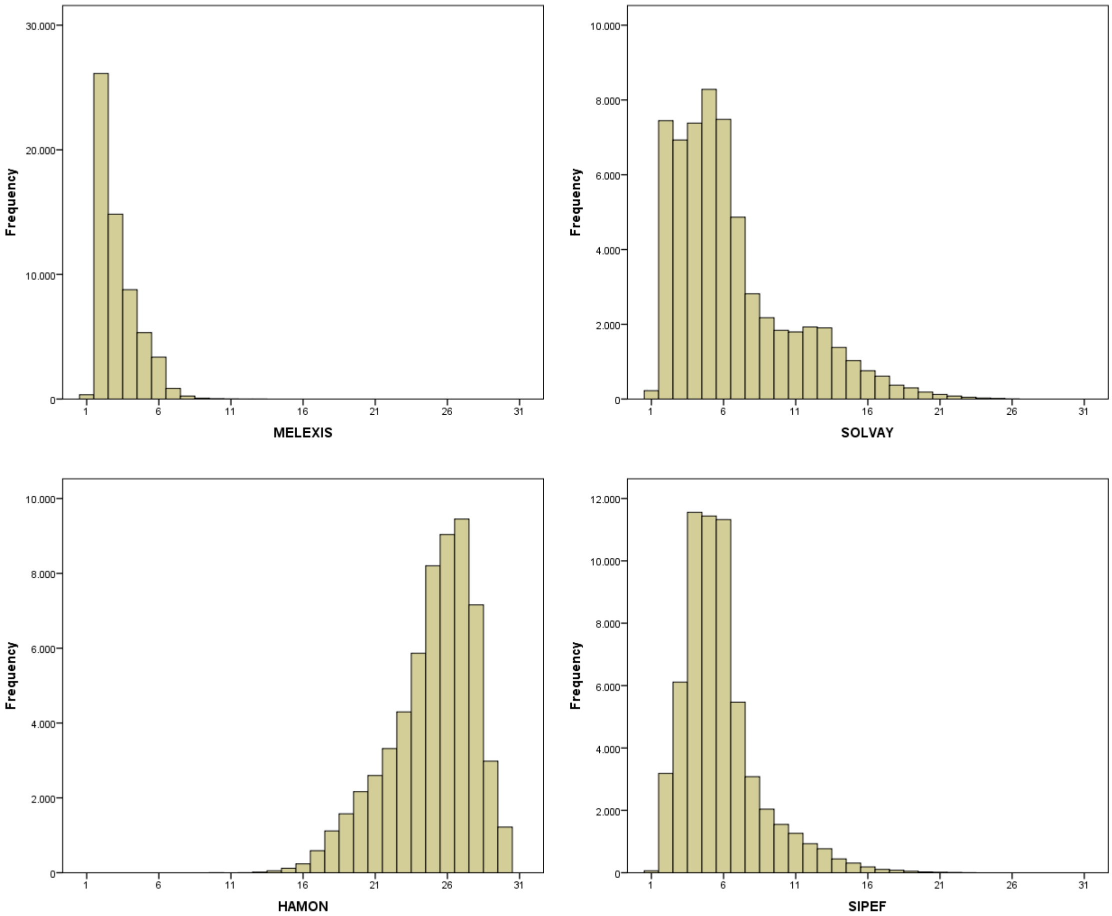


Figure 3: Uncertainty analysis: possible rankings when accounting for uncertainties in weights, normalization and aggregation.

When we choose to calculate uncertainty intervals for specific weighting schemes (i.e. equal weights, BOD, median random weights and Copeland rule), normalization methods (i.e. Min-Max, Z-score, categorical, distance, above or below the mean and ranking) and aggregation systems (i.e. arithmetic, geometric and multi-criteria), Figure 4 can be developed. This figure demonstrates that uncertainty intervals are smaller for top and worst performers than for average performers.

It is clear from both Figure 3 and Figure 4 that top performers can be separated from laggards regardless of which methodological choices are made during the development process. This is particularly relevant in the case of stock selection as we want to identify top performers and are less interested in 'average' performing stocks. Therefore we can already cautiously conclude that the CI is robust for the phenomenon we are trying to measure. We should however interpret the ranking of average performers more carefully as these stocks show some overlapping intervals. This is also confirmed by Figure 5 which shows that the average shift in ranking is 4,4 across all normalization, weighting and aggregation methods.

Our main goal in the remaining of this sensitivity analysis is to identify the main influencers (i.e. normalization, weights or aggregation) of this shift in ranking. The following sections will subsequently assess the influence of adding an indicator (Section 4.1.2), changing the normalization method and therefore also the influence of transforming indicators (Section 4.1.3), choosing another weighting scheme (Section 4.1.4) and finally the influence of the aggregation system (Section 4.1.5).

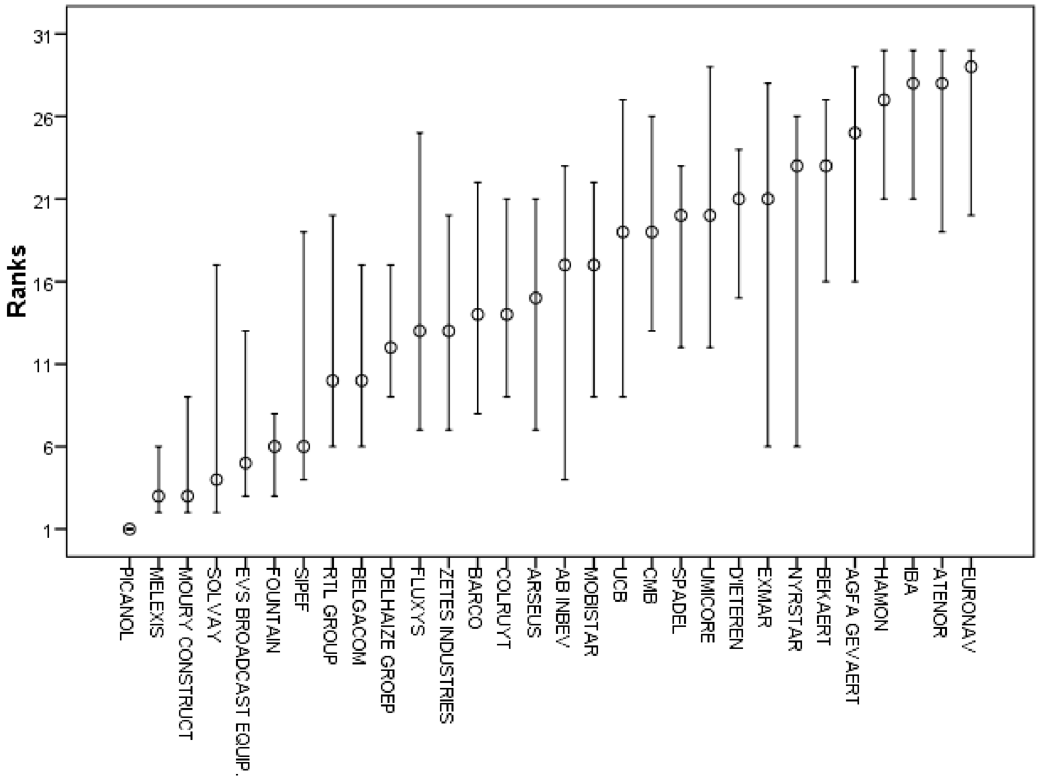


Figure 4: Trimmed uncertainty intervals (90%) around composite rankings for year 2012 for each stock ordered by median ranks. The circle represents the median ranking.

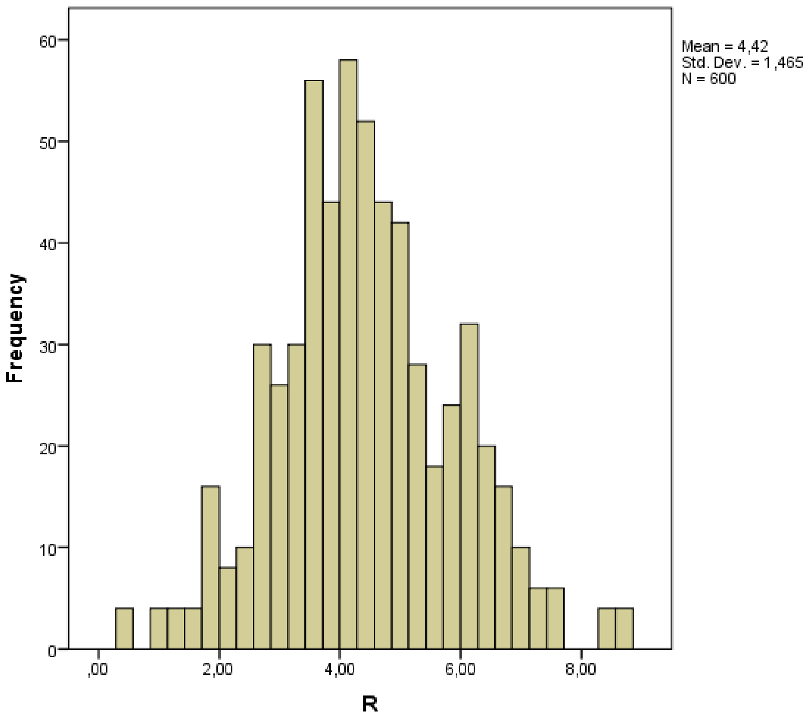


Figure 5: Average shift in ranking across three sources of uncertainty and across all stocks (normalization, weights and aggregation) for year 2012.

4.1.2 Adding Beta as an indicator

By accounting for investors' predictions, rankings and accordingly portfolio returns will change. If an investor is able to predict the trend of the stock market correctly, he can increase his total portfolio return over three years by an average of 18,56% when the top 3 is selected. However, if his prediction is not correct, his total return will drop by an average of 39,53%, punishing the investor very hard. Therefore we suggest only including Beta in the model when the investor is certain about the general market evolution. As Table 6 demonstrates, top ranked stocks' performance worsens while laggards' performance enhances when the prediction was wrong. However, when the investor is successful in his prediction, leaders' performance will ameliorate while laggards' performance will mostly deteriorate.

	2011		2012		2013	
	FALSE	SUCCESS	FALSE	SUCCESS	FALSE	SUCCESS
TOP3	-25,81%	5,69%	-16,81%	7,52%	3,09%	5,35%
TOP5	-16,01%	12,43%	-11,52%	2,66%	-3,66%	3,15%
TOP7	-12,63%	12,84%	-4,70%	3,85%	-2,26%	7,78%
TOP10	-10,94%	7,30%	-2,36%	4,94%	-3,46%	1,13%
LAST10	7,37%	-12,05%	-4,24%	-3,70%	0,08%	-2,02%
LAST7	9,80%	-7,18%	-3,01%	2,48%	-5,53%	-2,03%
LAST5	8,64%	-6,63%	-0,88%	-2,38%	-1,01%	4,72%
LAST3	0,58%	-9,67%	-3,92%	-1,74%	0,60%	10,97%

Table 6: Influence on average return of adding beta to the model for both false and successful predictions for top and last ranked stocks in 2011, 2012 and 2013.

4.1.3 Influence of normalization scheme

Figure 6 shows the differences in ranking for the most important normalization methods. If observations follow a 45° line from the origin, fewer differences in ranking occur, hence assuring that the choice of normalization method is not very influential. The values indicate the average shift in ranking (i.e. $\overline{R}$) when another normalization method was chosen. For example, the average shift in ranking when Z-score was used instead of Min-Max is 0,4, indicating nearly identical rankings. This is confirmed by an almost perfect 45° line. From Figure 6, one can conclude that the choice between Min-Max and Z-score is indifferent, categorical normalization is a decent alternative when we want to account for outliers as information on absolute levels is lost, and that the distance to the mean normalization results in very different rankings. One possible explanation for the differences in distance to the mean normalization is that this method magnifies the impact of outliers in our data. Furthermore, this finding is also confirmed by a correlation analysis. Ranks based on Min-Max normalization are highly correlated with ranks originating from Z-score (r=99,7%) and categorical normalization (r=89,9%). Again, the analysis demonstrates that the correlation with distance to the mean normalization is much smaller (r=60,9%).

Additionally, we can also assess the impact of omitting the absolute level of information of indicators' values. Borda-ranking and indicating whether a score is below or above the mean completely abandon the absolute level of information while categorical normalization still preserves some degree of information. Hence it is logical that the average shift in ranking between Borda-ranking, above/below mean and categorical normalization is very small. Moreover, a more in-depth investigation shows that differences in ranking are also very small when other normalization methods are applied, as leaders remain leaders and laggards remain laggards across all combinations of normalization methods. Again, the distance to the mean normalization is the only misfit in this observation. A graphical presentation as in Figure 6 confirming this phenomenon is given in Appendix D. Based on these results, we are confident that the choice of normalization method is not very influential for our results (except when distance to the mean is used).

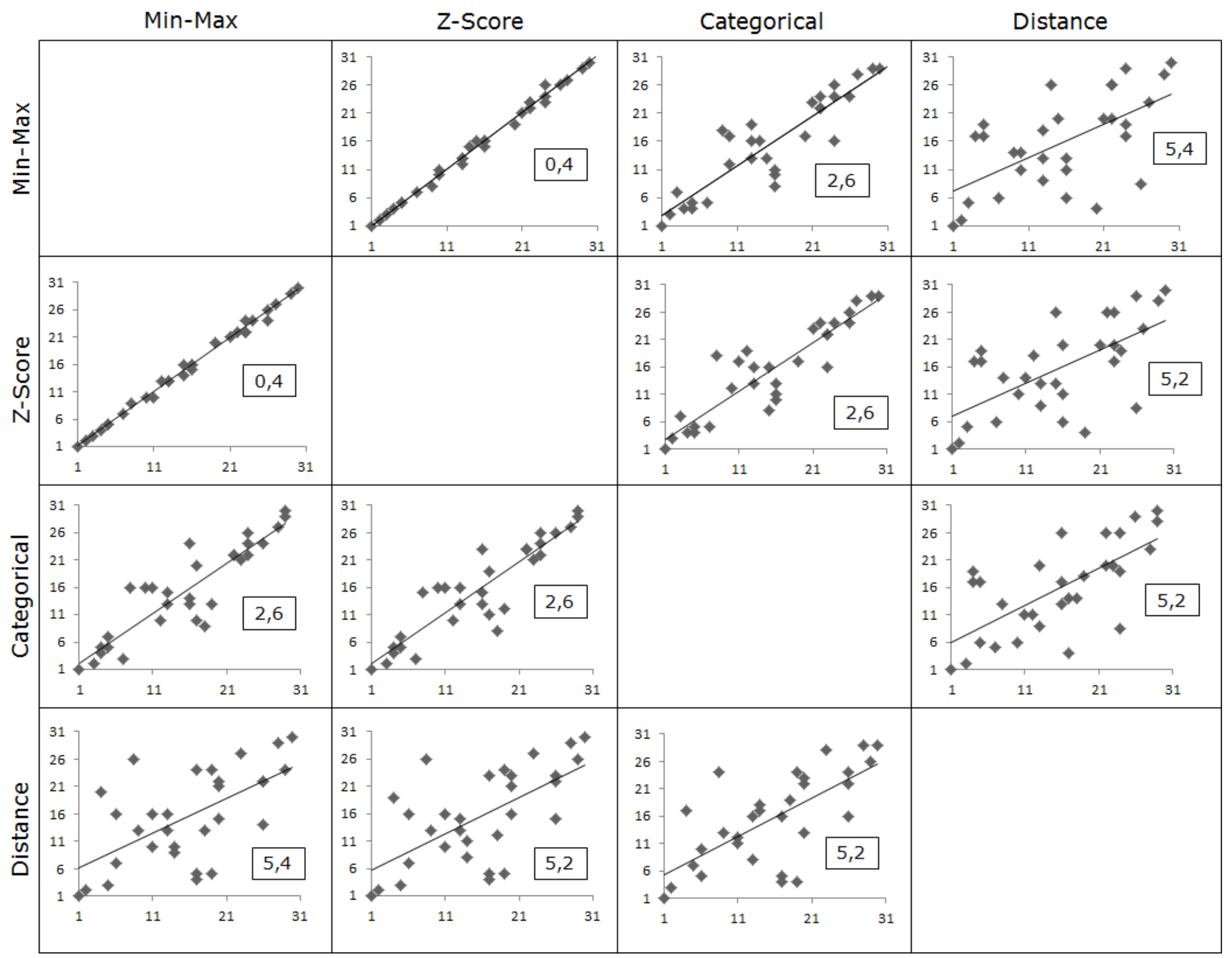


Figure 6: Impact of normalization methods on composite rankings for year 2012 (with median random weights and arithmetic aggregation). The value in the square indicates the average shift in ranking. Each axis represents the rankings of a stock that resulted from the two normalization methods mentioned.

4.1.4 Influence of weighting scheme

When analysing the influence of the weighting scheme, Min-Max normalization and arithmetic aggregation were applied as both methods are compatible with the BOD approach. Figure 7 assesses the difference in ranking between equal, random, BOD (Model 1) and Copeland (without BOD input) weights. The average shift in rankings has increased compared to those in Figure 6, as shown by the values in every quadrant in Figure 7. It is however notable that both leaders and laggards remain roughly the same across all quadrants (indicated by the circles in Figure 7). Again, the most drastic shifts in ranks are situated between average performers (i.e. the area in the middle of each graph). Hence, we call this area ‘noise’ or ‘a grey area’, which is influenced by subjective judgements. Yet, this should not cause any reason to question the robustness of the CI as investors merely want to identify top performers and in some cases also laggards, which remain robust regardless of the weighting scheme chosen. These results are confirmed by a correlation analysis, indicating above average correlations (r≈80%) between equal weights and random or BOD weights (Model 1). Besides, Copeland ranks are highly correlated with ranks derived from equal weights (r=95,4%) while ranks based on median random weights and BOD weights (Model 1) show little similarity (r=52,7%).

Moreover, the average shift in ranking of our four BOD models ranges between 4,9 and 8,4, hence showing that large variations in rankings occur when additional restrictions on weights are imposed (see Appendix E for a visualization). Again, these results are endorsed by a correlation analysis, demonstrating that correlations vary between 22,1% and 69,9%.

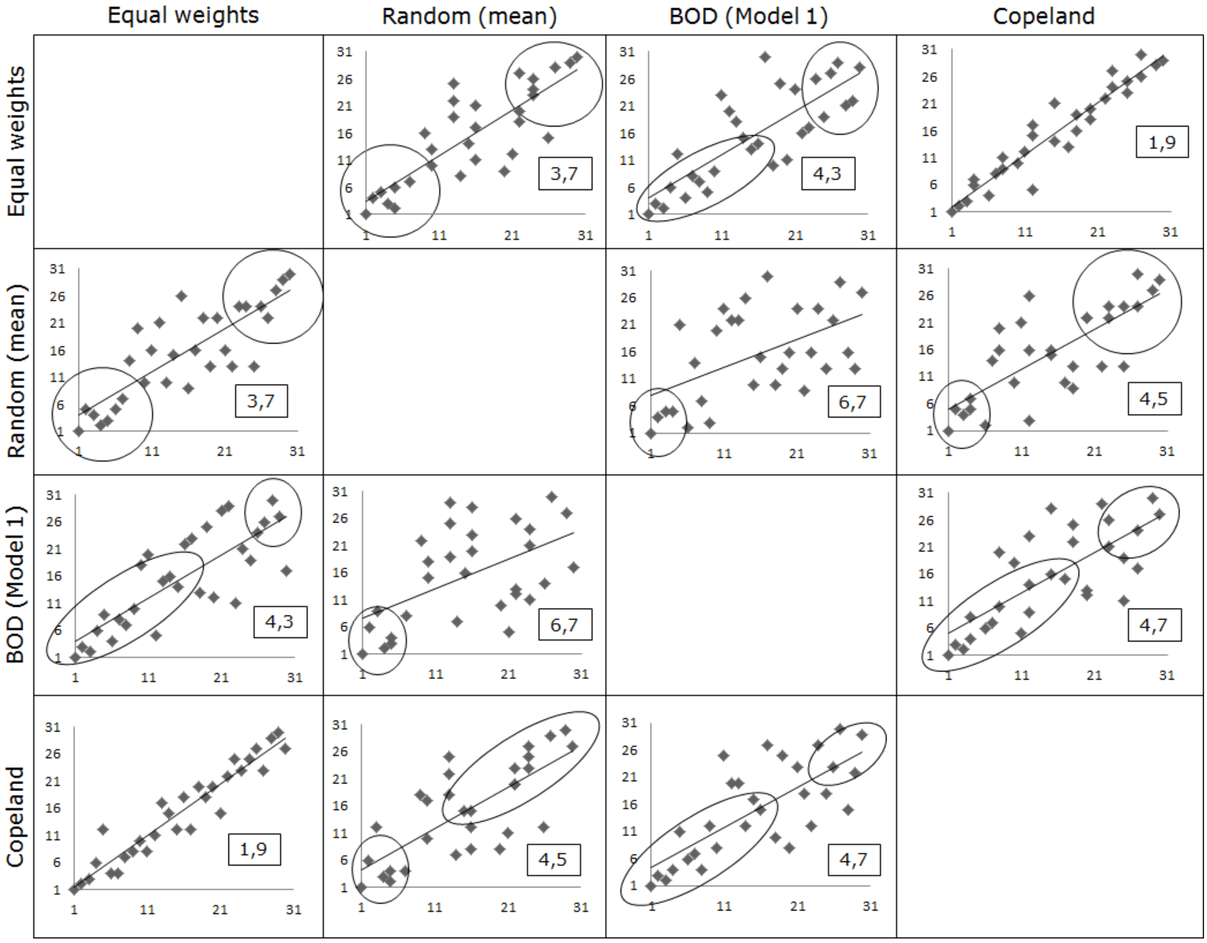


Figure 7: Impact of selection of weighting schemes on composite rankings for year 2012 (with Min-Max normalization and arithmetic aggregation). The value in the square indicates the average shift in ranking. Each axis represents the rankings of a stock that resulted from the two normalization methods mentioned.

4.1.5 Influence of aggregation system

The aggregation system is the final source of uncertainty that is analysed. The choice between arithmetic and geometric aggregation depends on the investor's preferences about whether compensability between indicators is desirable or not. We however think that compensability is recommended to a certain level in the context of FA. Therefore we will only quickly discuss the difference in ranking when geometric aggregation is used. Table 7 shows the average shift in ranking when geometric aggregation was utilised instead of arithmetic aggregation. Again, most shifts in rankings occur with average performers, leaving leaders and laggards mostly untouched.

	EMA	EZA	ECA	RMA	RZA	RCA
EMG	2,87					
EZG		0,87				
ECG			2,73			
RMG				3,70		
RZG					1,63	
RCG						2,70

Table 7: Impact of selection of aggregation system on composite rankings for year 2012 (E=equal weights, R=random weights; M=Min-Max, Z=Z-score, C=categorical; A=arithmetic, G=geometric). The values in the table indicate the average shift in ranking.

4.1.6 Robustness conclusion

Although rankings and composite scores are influenced by different sources of uncertainty, most leaders and laggards remain the same. This is again demonstrated by Table 8 where the average shift in ranking is ordered according to the median ranking of each stock. As it is the main goal of investors to identify top and worst performers, we think that our CI is robust for constructing a stock portfolio. Table 9, where we have regressed different sources of uncertainty (first column) to the difference in ranking, highlights the choice of weighting scheme as the most influential step in our development process. This is also confirmed by analysis of the F-statistics.[3] As a benchmark, we have used equal weights with Min-Max normalization and linear aggregation as these are the most occurring methods in practice. Indeed, Table 9 endorses our findings in the aforementioned sections. The difference in ranking between Min-Max, Z-score and categorical is very small and insignificant, again highlighting that the normalization scheme is not very influential. The highest and most significant coefficients occur when weighting schemes (except for Copeland weights) change, as revealed earlier by Figure 7. Additionally, the higher order sensitivity index states that S = 0,3710 for normalization, 1,56 for weights and 0,06 for the aggregation system, again confirming that the selection of the weights is the most influential.

All rankings should however be analysed carefully as even the smallest change in can induce large differences in return. Depending on the data under study, research comprising larger stock markets and larger time periods should be conducted in order to select the best weighting scheme. Besides, the main reason why average performers' rankings show a lot of variation across all sources of uncertainty is because the composite scores in the 'middle' of the ranking are very close to each other, opposed to larger differences in the top and worst performing segment of the rankings.

Stock (1-15)	$\overline{R_s}$	Median	Stock (16-30)	$\overline{R_s}$	Median
PICANOL	0,0	1	AB INBEV	5,9	17
MELEXIS	1,3	3	MOBISTAR	2,8	17
MOURY	1,7	3	CMB	3,0	19
SOLVAY	2,3	4	UCB	4,7	19
EVS	2,5	5	SPADEL	2,5	20
FOUNTAIN	1,2	6	UMICORE	4,1	20
SIPEF	1,8	6	D'IETEREN	3,0	21
BELGACOM	2,1	10	EXMAR	7,4	21
RTL GROUP	2,9	10	BEKAERT	3,8	23
DELHAIZE GROEP	1,9	12	NYRSTAR	4,1	23
FLUXYS	5,4	13	AGFA	3,3	25
ZETES	2,7	13	HAMON	2,1	27
BARCO	3,8	14	ATENOR	2,6	28
COLRUYT	3,8	14	IBA	1,6	28
ARSEUS	5,7	15	EURONAV	1,8	29

Table 8: Average shift in ranking for all stocks in 2012, across all CI's. The table demonstrates that both top and worst performers ranks are less volatile. The benchmark that was considered in this analysis is Min-Max normalization with equal weights and arithmetic aggregation.

[3] The F-test for the model including all dummy variables is 7,37**.** The model without normalization dummy variables obtains an F-test of 6,57, while the model without weighting dummies results in an F-statistic of 4,56. If we omit the aggregation dummy, an F-value of 7,39 is obtained. All changes in the F-statistic are significant, hence indicating that all choices are influential for our ranks. Yet, these incremental F-tests again highlight the choice of the weighting scheme as the most important source of uncertainty. Further research also showed that no significant changes in F-statistic occurred when only Z-score or categorical normalization were analysed. Indeed, this invigorates our results in Section 4.1.3. Note that the choice of the benchmark can be very influential for these statistics.

	Standardized Beta	t	Sig.
Z-score	-0,076	-1,539	0,124
Categorical	0,056	1,142	0,254
Distance	0,171	3,613	<u>0</u>
Borda	0,149	3,39	<u>0,001</u>
Mean	0,137	3,116	<u>0,002</u>
Random	0,275	6,275	<u>0</u>
BOD1	0,146	3,312	<u>0,001</u>
BOD2	0,198	4,489	<u>0</u>
BOD3	0,267	6,058	<u>0</u>
BOD4	0,243	5,509	<u>0</u>
COP0	0,023	0,526	0,599
COP1	0,077	1,742	0,082
COP2	0,078	1,782	0,075
COP3	0,151	3,43	<u>0,001</u>
COP4	0,104	2,37	<u>0,018</u>
Geometric	0,114	2,512	<u>0,012</u>

Table 9: Significance of coefficients when regressing different sources of uncertainty as dummy variables (first column) to the difference in ranking. Significant coefficients ($p < 0,05$) are underlined. The F-test for the model including all dummy variables is 7,37. See footnote 3 for more information on the F-statistics.

4.2 Visualization of the results

Many visualization techniques are available to represent composite scores (Nardo et al., 2005, 2008; Tarantola, 2008). We will address a few of them during the presentation of our results (e.g. tables, graphs, bar chart decompositions, spider diagrams and pie-shares). However, as most visualization techniques are stock-specific (i.e. a more in-depth analysis of one stock), we have often chosen to include only one figure for illustrational purposes.

4.2.1 Equal weighting

Table 10 represents an extraction of the rankings for year 2012 when equal weights were applied using arithmetic aggregation. As discussed earlier in the sensitivity analysis, rankings remain roughly the same across all normalization methods, except for the distance to the mean normalization.

	Min-Max	Z-score	Distance	Categorical	<u>Median</u>	<u>Interval</u>	<u>Return</u>
PICANOL	1	1	1	1	1	[1, 1]	81,78%
MOURY	2	2	9	2	2	[2, 9]	4,05%
SOLVAY	3	3	8	3	3	[3, 8]	76,59%
MELEXIS	4	4	2	4	4	[2, 4]	24,24%
EVS	5	5	4	9	5	[4, 9]	18,41%
…	…	…	…	…	…	…	…
ATENOR	26	26	24	28	26	[24, 28]	40,47%
BEKAERT	27	27	28	25	27	[25, 28]	-6,98%
HAMON	28	28	22	29	28	[22, 29]	-19,46%
IBA	29	29	29	30	29	[29, 30]	15,93%
EURONAV	30	30	30	27	30	[27, 30]	22,73%

Table 10: Ranks for year 2012 for equal weighting with arithmetic aggregation for different normalization methods. The last two columns represent the median rank and the interval of the ranks across the different normalization methods.

If we take a closer look at the results for the year 2012, Table 11 can be constructed. Here, we see that top-ranked stocks clearly outperform the worst performing stocks, suggesting that we have constructed a sound ranking method. The same is true for the year 2011 but not for all scenarios in year 2013. The underperformance for the year 2013 is due to an extremely good performance of all ship-owners societies (i.e. Euronav, Exmar and CMB), which usually received a bad ranking because of low growth prospects, high debt and poor performance in the past.

2012	Min-Max	Z-Score	Distance	Categorical
TOP3	54,14%	54,14%	49,17%	54,14%
TOP5	41,02%	41,02%	32,75%	37,99%
TOP7	29,45%	29,45%	34,60%	29,80%
TOP10	26,91%	26,91%	33,01%	27,58%
LAST3	6,40%	6,40%	10,56%	12,32%
LAST5	10,54%	10,54%	5,89%	7,18%
LAST7	2,67%	6,63%	6,59%	4,69%
LAST10	10,21%	8,80%	7,90%	5,04%

Table 11: Return of top and last ranked portfolios for 2012 based on equal weighting and arithmetic aggregation for different normalization methods.

After a three year buy and hold strategy with rebalancing every year and the top 3 ranked stocks in our portfolio (based on arithmetic aggregation), we achieve a median total return across all normalization methods of 59,33% (average of 19,78% per year) versus 12,65% for the BEL20 (average of 4,22% per year) and 19,20% for a randomly constructed portfolio (average of 6,40% per year). If we construct a portfolio consisting of the top 5 ranked stocks, we achieve a total return of 46,10% (average of 15,37% per year) versus 19,01% for a randomly constructed portfolio (average of 6,34% per year). When applying geometric aggregation, the results are somewhat different although they follow the same trend. After three years, a total return of 30,12% for a three stock portfolio and 32,56% for a five stock portfolio is achieved. This corresponds to respectively 10,04% and 10,85% on average per year, again outperforming both benchmarks.

To go even further in detail, we could analyse every stock separately. The question that we are addressing is: Why is Picanol achieving a higher score than for example Hamon? One figure that is able to answer this question is a bar chart decomposition presentation as in Figure 8. Here, we analyse the 3 leaders and 3 laggards and their corresponding dimension scores in year 2012. Note that the same presentation could have been made for individual indicators. However, due to the amount of indicators, this would become rather unclear. Again, we conclude from Figure 8 that leaders perform well on most dimensions and that laggards perform badly on most dimensions.

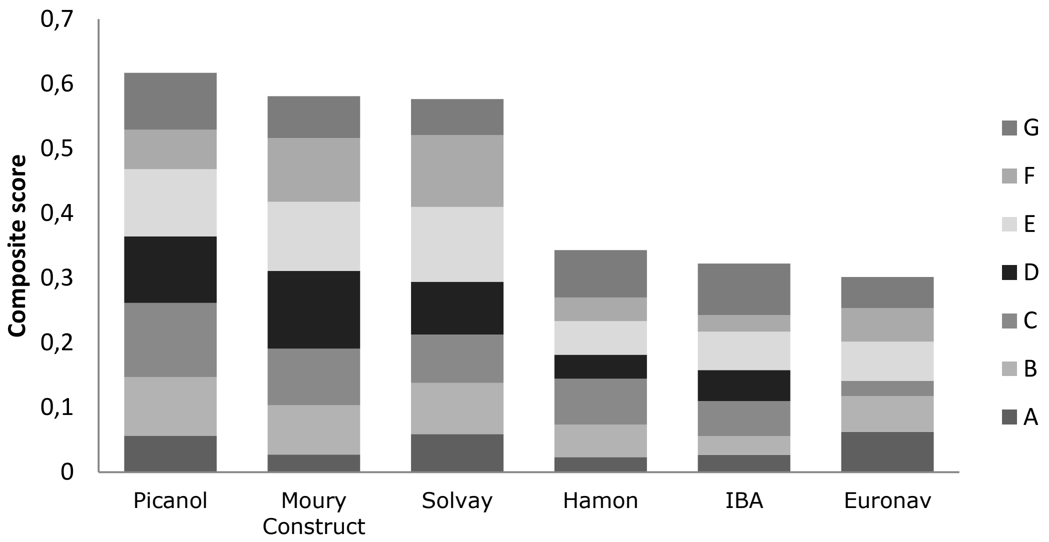


Figure 8: Example of a bar chart decomposition presentation for leaders and laggards in year 2012 (Min-Max normalization and arithmetic aggregation). Here, the sum of all dimensions scores (A to G) is the CI score, listed on the vertical axis.

4.2.2 BOD

Figure 9, shows the intervals of the ranks based on all BOD models. The lines represent the interval (min-max) within which each model's ranking lies. We clearly see that some intervals are very widespread. As we impose more, and more stringent, restrictions, stocks with bad scores on some indicators will be punished. Model 1 for example (i.e. no restrictions on the weights), favours stocks having very high scores on one or more indicators, even though they sometimes perform badly on most indicators. When we impose additional restrictions on weights, these stocks will have to assign higher weights to indicators with lower scores. Hence, stocks achieving average scores on most indicators will benefit. Barco for example (see dashed line in Figure 9) achieves a ranking of 23 in Model 1, as this stock does not achieve very high scores on one particular dimension. However, Barco never performs badly on any dimension, therefore climbing to place 3 in the ranking of Model 3 (i.e. with restrictions on the pie shares).

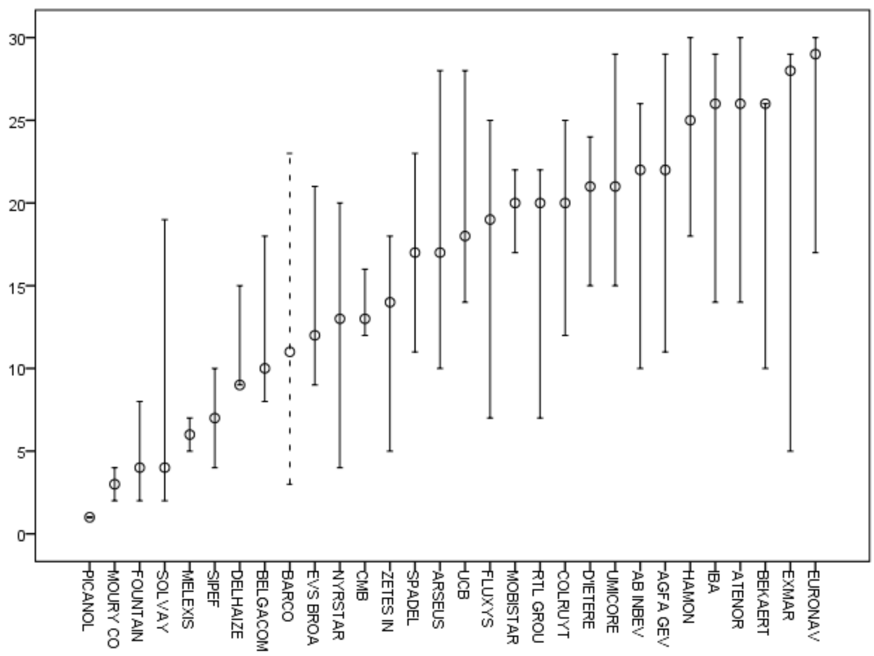


Figure 9: Example of ranking based on BOD for year 2012. The circle represents the median value of each stock's ranking.

One of the main advantages of the BOD approach, and especially from our two-stage approach, is that we can analyse every stock in more detail. Picanol for example, seems a very interesting stock to investigate more thoroughly, as it achieves the highest ranking with every method that we have used so far. We will analyse the pie-shares of all dimensions, as calculated in Model 3. The size of the pie reflects the CI score of that stock, while the size of each pie-share reflects the relative importance of all dimensions in the pie and thus in the CI. These pies are presented in Figure 10. Second, the same process can be repeated in order to analyse subindicators shares in one dimension (i.e. pie-shares after stage 1), which is represented in Figure 11.

Figure 10: Pie chart representation of BOD (dimension level, after stage 2 of BOD Model 3).

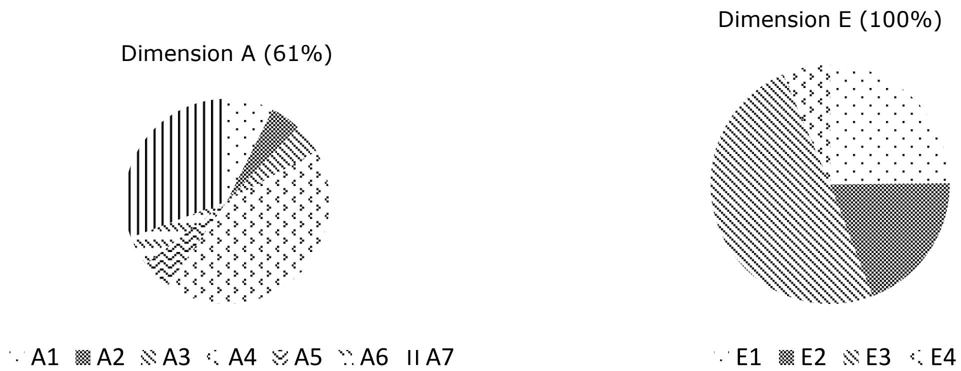


Figure 11: Example of pie chart (after stage 1 of BOD Model 3, for Picanol).

Another interesting presentation is the spider diagram. We will as an example analyse the efficiency of some stocks for each dimension (i.e. their relative performance for each dimension against the leader in that dimension for Model 3 of the BOD approach). To illustrate the value of such a presentation, we again consider Picanol and Hamon, like we did in Figure 10. This should enable us to highlight the differences between the pie-share presentation and the spider diagram, which both represent results in a different way. The spider diagram is depicted in Figure 12, clearly showing why Picanol is the consistent leader in our ranking as it achieves the maximum score on three dimensions (i.e. a score of 100 and thus the best performer in this dimension). Additionally, it also becomes clear why Hamon is always in the lower part of the ranking as it is for example never able to achieve a better score than Picanol on any dimension.

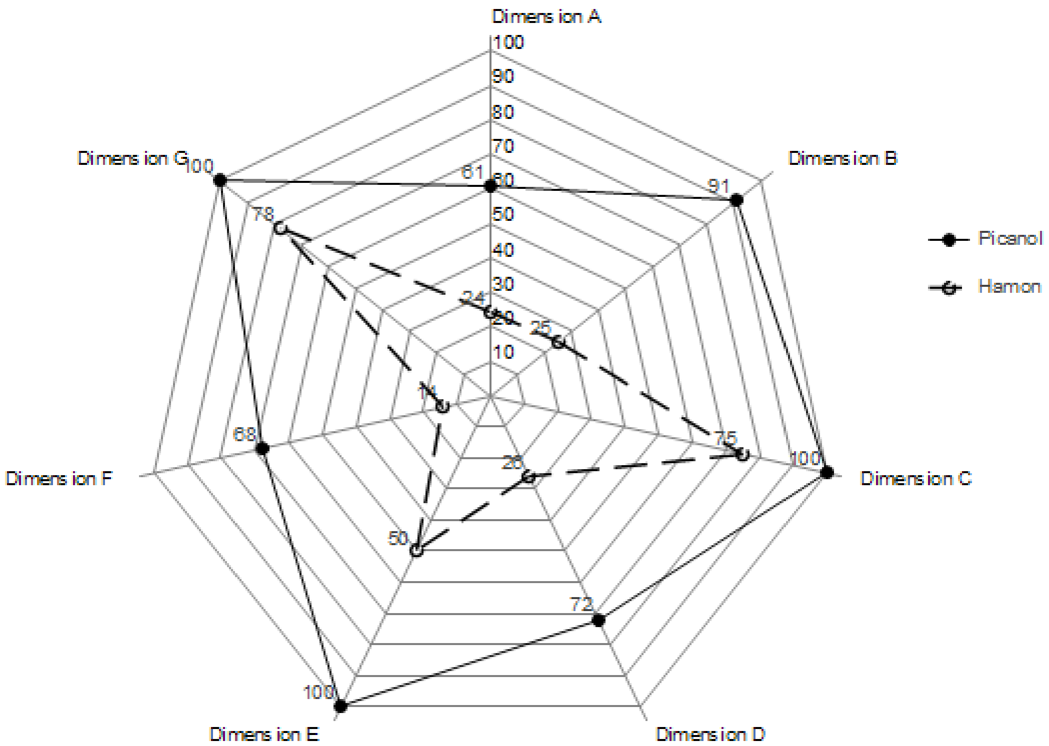


Figure 12: Example of spider diagram decomposition (dimension level, after stage 1 in BOD Model 3).

Until now, we have only presented some results for the year 2012. Hence, as we want to know what our total return would be after three years, we also put our return in a graph and plot it against the return of the BEL20 and the average return of 10.000 random portfolios (see Figure 13). Figure 13 shows that all models outperform the BEL20 after three years and that all but one perform better than the randomly selected portfolios. Models 1 and 4 heavily outperform both benchmarks in all subperiods, while Model 2 performs better in two of the three time periods considered. However, Model 3 is not able to beat the benchmark, although the difference in performance is very small.

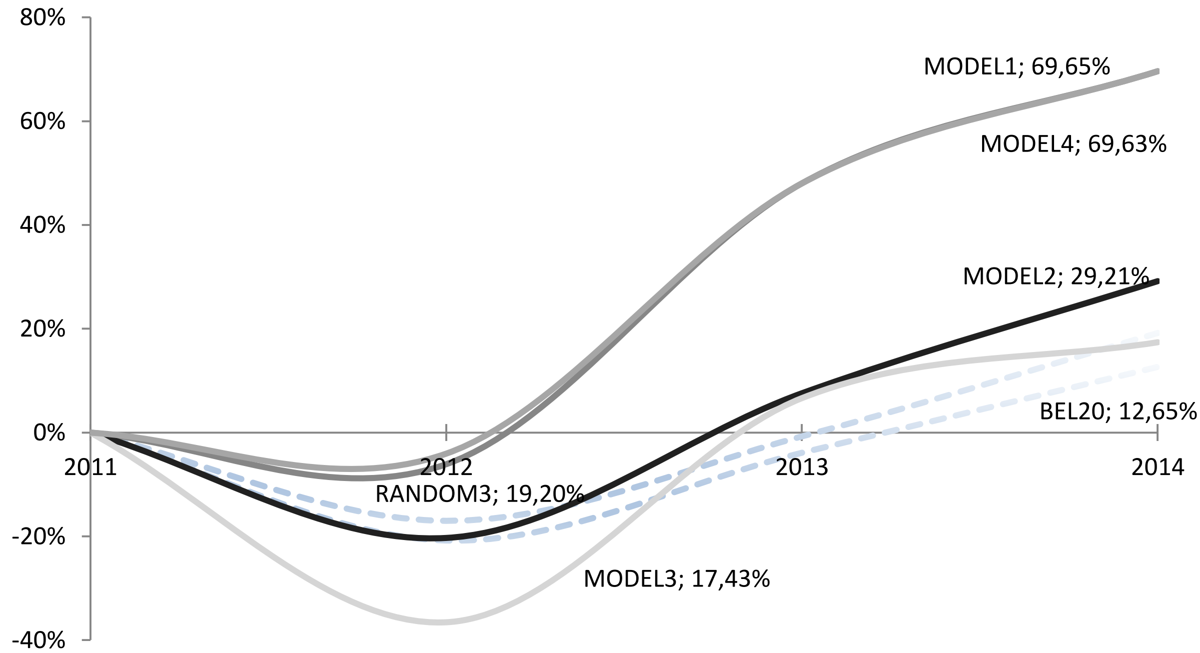


Figure 13: Return of four BOD models with a buy and hold strategy for three years (solid lines), versus BEL20 and average return of 10.000 randomly selected portfolios (dashed lines).

4.2.3 Random weights

Assigning random weights to indicators does not sound as a much advised method. However, if we calculate 10.000 different sets of randomly selected weights, we can calculate median rankings, returns, and so on, which can lead to very interesting results. Moreover, the main purpose of this method is to implement it as a benchmark. Figure 14 for example gives us an overview of the interval (min-max) of the returns in each year using 10.000 different sets of weights (i.e. a Monte Carlo simulation). Hence, if we plot some of our models determined in the previous sections, we can check whether the performance of our models is closer to the best set of weights or to the worst set of weights in each period according to our simulations. According to Figure 14, Model 4 of the BOD approach closely tracks the performance of the best set of weights from the 10.000 Monte Carlo simulations (i.e. the top 3 with the highest return). When equal weights and geometric aggregation were applied, we achieved a mediocre score. Also note that the total return of an investment strategy based on the selection of the top 3 median ranked stocks after 10.000 random simulations for the weights and Min-Max normalization would have yielded a return of 50,27% after three years, again outperforming both benchmarks.

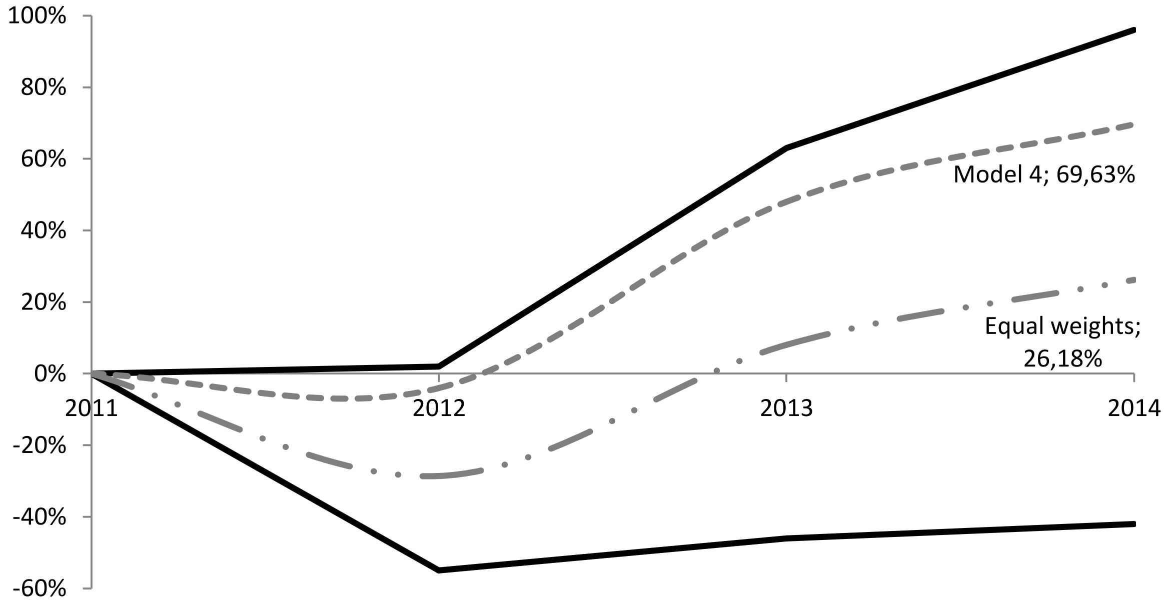


Figure 14: Representation of return for Model 4 (BOD) and equal weighting with geometric aggregation (dashed lines) and the best and worst return of the top 3 ranked stocks after 10.000 Monte Carlo simulations for the selection of weights (solid lines). In all cases, Min-Max normalization was used.

4.2.4 Copeland rule

When we apply the Copeland rule to all indicators in each dimension and subsequently to all dimensions, the rankings that are obtained are very similar to those derived from previous weighting schemes. However, as already discussed in Section 3.6.4, Saisana (2012) suggests that in general, some compensability is desired at lower hierarchical levels. Therefore we have incorporated the outcomes of the first stage of the BOD-approach to determine weights for all indicators within each dimension. Subsequently, the Copeland rule was applied to assign weights to each dimension. Similar to Figure 13, Figure 15 presents the return of this strategy, using all first stage outcomes of the four models derived in Section 3.6.2 as input. Again, one can conclude that the leaders and laggards are very similar to those derived in previous sections, despite major swings in average performers' rankings. Note from Figure 15 that all models, even Model 3 that underachieved the benchmark in Section 4.2.2, perform better than the BEL20 and randomly constructed portfolios. Additionally, one should also notice that although the total return has declined for Model 1 and Model 4, almost all models are able to beat the benchmarks in all periods considered. Therefore we think that in the long run, this method will outperform a portfolio strategy that is solely based on BOD (as considered in Section 4.2.2).

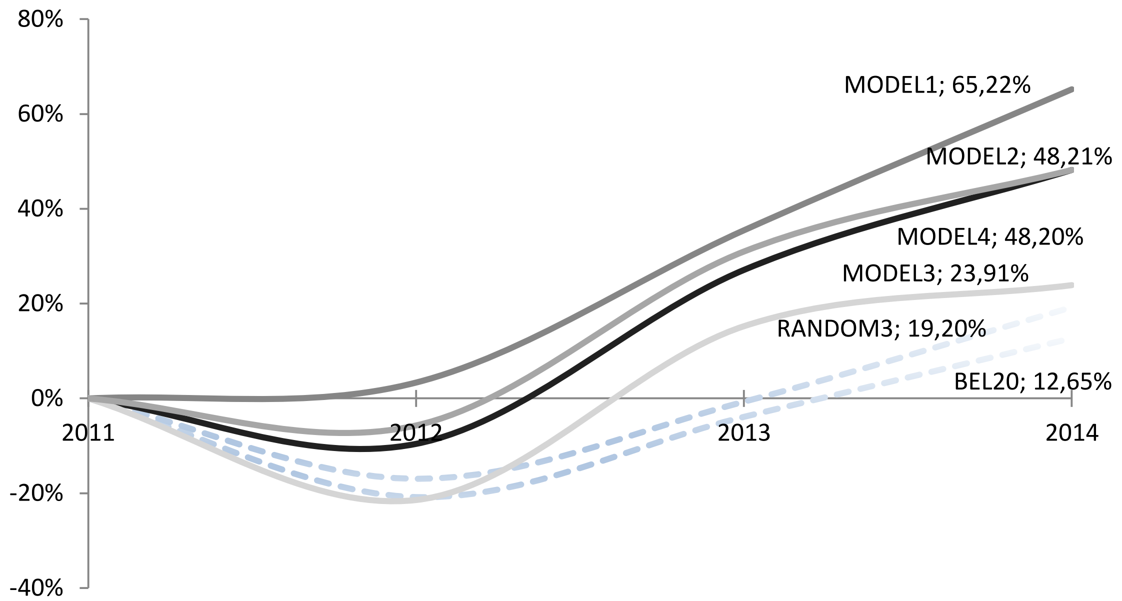


Figure 15: Return of Copeland rule in combination with four BOD models with a buy and hold strategy for three years (solid lines), versus BEL20 and average return of 10.000 randomly selected portfolios (dashed lines).

4.2.5 Summation

We have also considered the summation of rankings according to the Borda rule and the summation of indicators normalized according to whether they outperform the mean or not (see Section 3.5.6). The results indicate that a simple method based on rankings themselves, whereby all absolute information is lost, still yields decent returns. Yet, the results based on indications whether a case is above or below the mean, is among the worst performers considered in this analysis. As we think that both methods are too simplistic, we will not discuss the results in detail here. Moreover, we merely see these results as a reason to engage in FA prior to stock selection, not that assigning weights to indicators is not necessary.

4.2.6 Discussion

As a summary to initiate a more in-depth discussion of the results, Table 12 displays the most important rankings for the Min-Max normalization for all three years. The table should be read horizontally. Note that the same table could have been constructed for other normalization methods and that not all weighting schemes are represented. As can be seen, almost all CI's heavily outperform the benchmarks (i.e. BEL20: 12,65%; RANDOM3: 19,65%). Moreover, these results are still valid when one selects the top 5 or top 7 instead of the top 3 ranked stocks as we did in most examples. Additionally, in 2011 and 2012, laggards in ranking are also among the worst performers in terms of return, yet in 2013 this is not the case. Table 12 should however be interpreted with care as one can conclude that rankings remain the same across all different CI's. Although this is true for leaders and laggards, huge swings can occur when evaluating the ranking of average performers.

One should also note that the impact of the normalization scheme is in most cases indifferent. Moreover, the results showed that softening or abandoning the absolute level of information (i.e. categorical, ranking or above/below mean normalization) does not induce large variations in the ranks. This means that investors could even enhance their portfolio return by simply selecting stocks based on a summation of ranks. If an investor is able to spend more time on his portfolio composition, more advanced weighting techniques such as BOD could enhance the portfolio returns even further. We also suggest to construct more than one CI as the differences in ranks could lead to interesting insights. Furthermore, visualization techniques such as pie-shares, spider diagrams and bar chart decompositions can enable the investor to investigate the most promising stocks more thoroughly. As a CI remains a static representation of a particular situation at a certain time, one should remain cautious with the results. A company can for example publish a press release that could have a significant influence on its stock return, hence potentially resulting in a misleading composite ranking as this press release is not incorporated in our data. One example is that a particular company indicates that it obtained a very high net profit in the previous period due to some exceptional incomes. As this high net profit is not sustainable, it is not advised to base the CI score on this number. To investigate whether this exceptional income is dominating the stock score, we should again investigate the pie-shares, bar chart decompositions or spider diagrams of that stock.

In summary, the results look very promising and could potentially mean that FA is still useful for investors. This accords with the results of among others Ou and Penman (1989), Holthausen and Larcker (1992), Lev and Thiagarajan (1993), Abarbanell and Bushee (1997), Setiono Strong (1998), and Charitou and Panagiotides (1999). Additionally it contradicts the statement of Curtis (2012) which states that in recent time, FA does no longer yield higher returns, as was depicted earlier in the literature review (Section 2.1). However, note two important remarks. First, the time period that was studied by Curtis (2012) differs from ours (1994-2008 versus 2011-2013). Second, our time period is too short to reach any statistical conclusions. More extensive research is needed to confirm these results in other time periods.

Since simple CI's based on equal weighting still yield high returns, it is rather unclear which dimensions are the most predictive for future returns in our study. However, if we analyse our results more thoroughly, we can cautiously point out some remarks. First, when we do not impose any restrictions on the weights in our BOD models, high weights are in general assigned to the following dimensions: growth signals, management performance, valuation, profitability and to a lesser extent solvability. Therefore the ascertainment of Delen et al. (2013), stating that growth signals are the most important in predicting stock returns, followed by valuation and profitability, can be confirmed. Second, Clark-Murphy and Soutar (2004) state that individual investors value the performance of the company's management and the recent movements of the stock when buying a stock. Indeed, management performance is an important aspect in every CI. Yet, the recent movement of the stock price (i.e. momentum) has in most cases a negligible influence on the composite score. As most CI's achieve decent returns, we conclude that individual investors should pay more attention to fundamental ratios rather than to stock price movements. Third, unlike Basu (1977), Fama and French (1992), Chan and Lakonishok (2004), Arnott (2005), and many others, we did not find any support that value stocks are significantly outperforming growth stocks.

Additionally, the results also invigorate the use of a CI in investment decision strategies as almost all CI's outperformed the benchmarks in most periods considered. However, again more research is needed to conclude whether this remains valid for larger stock markets and other time periods. Depending on the preferences of the investor (e.g. compensability of indicators and risk averse), the investor can opt for one of the aforementioned normalization methods, weighting schemes and aggregation systems. Yet, we want to express which method should, according to us, be preferred in the context of portfolio construction. Indeed, the correlation between the ranks of each CI and return should be a key concept in the evaluation of all CI's. Hence, our preference goes to Models 1 and 4 of both the BOD and Copeland approach, not only because they are among the best performers but also because the methodologies accompany the vision of most investors. Moreover, these models were able to closely track the 'ideal' weights derived from our Monte Carlo simulations (see Figure 14). We should again stress the advantages of BOD models compared to other weighting techniques. As we want to design a flexible CI, reducing the static character of our set of weights, flexible weighting schemes such as BOD are necessary. If for example, companies enter or leave our dataset (e.g. due to new stock listings or delistings), the set of weights will change when BOD was applied instead of equal weights. Moreover, when companies' scores change, the optimal set of weights will also change, in turn leading to a more flexible and self-adapting CI. These advantages of BOD models were already highlighted by among others Cherchye et al. (2006) and Hermans et al. (2008).

		Arithmetic						Sum				Geometric	
		Equal weights	Random weights	BOD weights				Copeland rule				Equal weights	Random weights
		Min-Max	Min-Max	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4	Min-Max	Min-Max
2011	TOP 1	SIPEF	SIPEF	SIPEF	RESILUX	DUVEL	SIPEF	TELENET	SIPEF	SIPEF	EVS	SIPEF	SIPEF
	TOP 2	EVS	EVS	RESILUX	AGFA	HAMON	RESILUX	SIPEF	EVS	DUVEL	SIPEF	DUVEL	DUVEL
	TOP 3	DUVEL	DUVEL	VPK	VPK	BEKAERT	EXMAR	EXMAR	DUVEL	BEKAER	EXMAR	BEKAERT	EVS
	LAST 3	IBA	ATENOR	TESSENDER	SAPEC	SAPEC	TESSENDER	SAPEC	ALFACA	IBA	ALFACA	IBA	SAPEC
	LAST 2	ALFACAM	IBA	IBA	IBA	ALFACAM	IBA	ALFACAM	SAPEC	ALFACA	IBA	SAPEC	ALFACAM
	LAST 1	SAPEC	SAPEC	SAPEC	ATENOR	NYRSTA	SAPEC	IBA	IBA	SAPEC	SAPEC	ALFACAM	IBA
2012	TOP 1	PICANOL	PICANOL	PICANOL	PICANOL	PICANOL	PICANOL	PICANOL	PICANOL	PICANO	PICANO	PICANOL	PICANOL
	TOP 2	MOURY	MELEXIS	SOLVAY	FOUNTAI	MOURY	SOLVAY	MELEXIS	MELEXIS	MELEXI	MELEXI	MOURY	MELEXIS
	TOP 3	SOLVAY	EVS	MOURY	MOURY	BARCO	FOUNTAIN	MOURY	MOURY	MOURY	MOURY	MELEXIS	MOURY
	LAST 3	HAMON	HAMON	ARSEUS	EXMAR	ATENOR	UCB	IBA	IBA	NYRSTA	HAMON	AGFA	AGFA
	LAST 2	IBA	IBA	UMICORE	IBA	AGFA	EXMAR	HAMON	HAMON	IBA	UMICOR	ATENOR	ATENOR
	LAST 1	EURONAV	EURONAV	HAMON	EURONA	EURONA	ATENOR	ATENOR	ATENOR	ATENOR	ATENOR	EURONAV	EURONAV
2013	TOP 1	PICANOL	PICANOL	PICANOL	JENSEN	PICANOL	PICANOL	PICANOL	PICANOL	PICANO	PICANO	PICANOL	PICANOL
	TOP 2	VANDEVE	EVS	JENSEN	RECTICE	MOURY	JENSEN	BARCO	BARCO	BARCO	BARCO	BARCO	UMICORE
	TOP 3	BARCO	VANDEVEL	RECTICEL	PICANOL	RECTICE	RECTICEL	DECEUNIN	D'IETER	MOURY	COLRUY	VANDEVE	BARCO
	LAST 3	EXMAR	IBA	LOTUS	EXMAR	EXMAR	EXMAR	SOLVAY	SOLVAY	SOLVAY	EURONA	MOBISTA	NYRSTAR
	LAST 2	NYRSTAR	BEKAERT	CMB	NYRSTA	IBA	BEKAERT	MOBISTAR	IBA	BEKAER	MOBIST	EURONAV	EURONAV
	LAST 1	IBA	NYRSTAR	EURONAV	BEKAERT	MOBISTA	EURONAV	EXMAR	EXMAR	EXMAR	EXMAR	IBA	IBA
	RETURN (2011)	-9,57%	-9,57%	-6,01%	*-20,29%*	*-36,55%*	-4,04%	3,36%	-9,57%	*-21,46%*	-5,70%	*-28,61%*	-12,30%
	RETURN (2012)	54,14%	41,48%	54,14%	27,89%	43,21%	52,07%	32,12%	36,69%	36,69%	36,69%	36,69%	36,69%
	RETURN (2013)	*18,10%*	*18,19%*	21,61%	21,61%	*10,77%*	21,61%	29,74%	21,08%	*8,67%*	*17,21%*	*18,10%*	*18,10%*
	AVG. (TOP 3)	20,89%	16,70%	23,25%	9,74%	*5,81%*	23,21%	21,74%	16,07%	7,97%	16,07%	8,73%	14,16%
	TOTAL (TOP 3)	62,67%	50,10%	69,74%	29,21%	*17,43%*	69,64%	65,22%	48,20%	23,90%	48,20%	26,18%	42,49%

Table 12: Review of the most important methods to construct the CI's. For practical purposes, some companies names have been abbreviated. Returns outperforming both benchmarks are underlined.

4.3 Looking back at the data

Analysis of the clusters defined in Section 3.4.3 exhibits that no specific sectors are consistently in the top or worst performing segment in the ranking. There is however a slight, yet not heavily pronounced, trend in both the top and worst performing segment of the ranking. Both segments often consist of growth stocks. Value stocks are mostly average performers. Indeed, this seems logical as growth stocks will have more extreme prospects and valuations. Additionally, the track records of these growth stocks are more volatile than those of value stocks. We also want to highlight the fact that returns of value stocks were less volatile than those of growth stocks (i.e. less extreme returns, both positive and negative). Hence, if an investor wants to reduce its portfolio risk, he can assign higher weights (e.g. in the BOD models by changing the ordinal restrictions) to dimensions E and G.

5 Conclusions and further research

In this paper, we have developed different CI's that should enable investors to construct a decent stock portfolio. Our literature review (see Section 1 and 2) demonstrated that many investors struggle to develop a profitable trading strategy. The main cause of this problem is that investors are making investment decisions based on emotional, rather than on rational grounds. Hence, one of the main building blocks of our CI was FA, a technique that links financial statement variables to the true or intrinsic value of a company. Based on a guided search approach, we compiled a set of 39 fundamental indicators that have proven to be predictive for stock returns in the past (see Appendix A and B). These indicators were derived for respectively 35, 30 and 38 traditional stocks in a three year period (2011-2013) for stocks quoted on the Euronext Brussels segment (see Appendix C).

The most important step however was the construction of the CI itself. Transparency and reducing subjective judgements were key principles throughout the entire development process. To guarantee the robustness of our CI, we followed the widely adopted framework of Nardo et al. (2005, 2008) as depicted earlier in Figure 1. Additionally, we did not limit our study to the development of one single CI. In order to assess the influence of all subjective judgements that have insurmountably slipped in during the construction of our CI, we have decided to develop numerous CI's varying normalization methods, weighting schemes and aggregation systems. Six normalization methods were chosen to investigate the impact of dropping or softening the level of absolute information of indicators. We applied Min-Max, Z-score, categorical, distance to the mean, (Borda-)ranking and indicating whether an observation is below or above the mean normalization to analyse this impact more thoroughly. Additionally, we selected four weighting schemes: equal weighting, BOD, random weights and the Copeland rule. Equal weighting is often used in practice due to its simplicity, however we demonstrated that this could lead to an unbalanced CI. Hence, we have opted to present a more in-depth analysis of different BOD models. A first advantage of BOD is that weights are only influenced by the restrictions that we impose, the stocks in our dataset and the scores of the stocks on each indicator. Hence, this weighting scheme is flexible and self-adapting, two much advised properties in the context of portfolio construction. This means that when new data is available, weights will be automatically adjusted. Therefore, the static character of the CI is reduced. A second advantage is that BOD determines optimal sets of weights for each stock, resulting in weights that are stock specific. As this could however result in unfair comparisons, we have decided to calculate mean cross efficiencies or to maximize the sum of all CI values.

In our study, we decided to develop a two-stage BOD model. In the first stage, dimension scores were maximized for all indicators in each dimension. Subsequently, these optimized dimension scores were combined into one composite measure in the second stage. The main reasons for this approach were: 1) guaranteeing visibility throughout the entire process for both indicator and dimension optimization, 2) so the output of the first stage could be used as input in other weighting schemes such as the Copeland rule, and 3) enabling the investor to investigate both dimension and indicator importance in the CI more thoroughly. Additionally, we constructed four different BOD models, each with their own restrictions on weights such as absolute or ordinal boundaries and pie-share restrictions.

The random weighting scheme, consisting of 10.000 Monte Carlo simulations, was merely used as a benchmark to assess the performance of other weighting schemes. Finally, the Copeland rule was applied in combination with the output of the first stage BOD models. Hence, we assured compensability at the lower hierarchical levels (i.e. indicators within each dimension) and less compensability at higher hierarchical levels (i.e. dimensions). Subsequently two aggregation systems were applied: arithmetic and geometric aggregation. As we think that compensability is to some extent preferable in the context of stock selection, we have however mainly considered arithmetic aggregation.

Before we interpreted the results of our research, we conducted an uncertainty and sensitivity analysis where different sources of uncertainty were investigated. Here, we concluded that the choice of normalization and aggregation scheme is not very influential for our ranks. However, it became clear that the selection of weighting scheme has a substantial impact on our results. Yet, top ranked stocks remained in the top segment and last ranked stocks stayed in the worst segment across all weighting schemes. As investors merely want to identify leaders and laggards in the ranking, we concluded that the CI's are robust for the phenomenon we are trying to measure.

The results can be summarized in some key findings:

(1) Leaders remain leaders and laggards remain laggards across all CI's considered in the analysis;
(2) Almost all CI's are able to outperform the benchmarks, hence concluding that FA is still useful for investors;
(3) Model 1 and Model 4 of the BOD approach are the best performing weighting schemes with a return of almost 70% after three years which closely tracks the performance of the ideal set of weights (compared to 12,7% and 19,7% for our benchmarks);
(4) The choice of normalization scheme is not very influential, even when we omit the absolute level of information;
(5) Investors can increase their performance by more than 6% per year if they are able to correctly predict the direction of the general stock market.

The main conclusion is however that trading strategies based on FA are able to generate higher returns. We cannot statistically conclude which weighting scheme will outperform stocks in other periods as more research is needed to confirm these results. Yet, the results show that further research would certainly be useful as neither CI's, nor specific BOD and Copeland models have received much attention in academic research comprising the prediction of stock returns. However, the returns of this 'pilot-study' look very attractive to most investors to say the least. We suggest applying the same models for a larger and more liquid market than Euronext Brussels. Additionally, instead of FA, technical analysis or a combination of both could be used. Moreover, we think that the use of quarterly data (instead of yearly in this study) would enhance the performance of the CI. We also strongly believe that the results would be more robust when the data is more homogenous (i.e. only banks, insurance companies or retail companies for example), despite this does not accord with the general conception of diversification in portfolio theory. Beside the financial statement variables, other concepts from the financial economics literature, such as Value at Risk (VaR) or expected shortfall, could be incorporated in the framework of the CI. Additionally, we are also aware of a major drawback of our study, namely that our CI does not account for interrelations between stocks (i.e. correlation). Hence, it could be that the exposure of our portfolio to one specific sector becomes very high, thus unnecessarily increasing our portfolio's risk. Therefore, future research should also take correlations and other aspects of portfolio composition into account. A last point that should be considered is that a one year framework is often too short for FA to be effective. Hence, fundamental investors should broaden their investment period to at least 5 to 10 years.

Acknowledgments
I would like to thank my promotor, Prof. Dr. Filip Van den Bossche, for giving me the opportunity to work on this enthralling subject and for guiding me through this Master's thesis. Additionally, I would also like to point out my regards to Prof. Dr. Chris Verschooten for reviewing my English grammar.

References

Abad, C., Thore, S., & Laffarga, J. (2004). Fundamental analysis of stocks by two-stage DEA. *Managerial and Decision Economics, 25*(5), 231-241.

Abarbanell, J.S., & Bushee, B.J. (1997). Fundamental analysis, future earnings, and stock prices. *Journal of Accounting Research, 35*(1), 1-24.

Abarbanell, J.S., & Bushee, B.J. (1998). Abnormal returns to a fundamental analysis strategy. *The Accounting Review, 73*(1), 19-45.

Al-Debie, M., & Walker, M. (1999). Fundamental information analysis: An extension and UK evidence. *British Accounting Review, 31*(3), 261-280.

Arnott, R.D. (2005). *Disentangling size and value*. Retrieved from http://rallc.com/Our%20Ideas/Insights/Papers/Documents/FAJ_Sep_Oct_2005_Disentangling_Size_and_Value.pdf

Atsalakis, G.S., & Valavanis, K.P. (2009). Surveying stock market forecasting techniques – Part II: Soft computing methods. *Expert Systems with Applications, 36*(3), 5932-5941.

Avkiran, N.K. (2011). Association of DEA super-efficiency estimates with financial ratios: Investigating the case for Chinese banks. *Omega, 39*(3), 323-334.

Avkiran, N.K., & Morita, H. (2010). Predicting Japanese bank stock performance with a composite relative efficiency metric: A new investment tool. *Pacific-Basin Finance Journal, 18*(3), 254-271.

Ball, R., & Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research, 6*(2), 159-178.

Bandura, R. (2008). *A survey of composite indices measuring country performance: 2008 update* (UNDP/ODS Working Paper). Retrieved from http://web.undp.org/developmentstudies/docs/indices_2008_bandura.pdf

Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. *Journal of Finance, 32*(3), 663-682.

Beneish, M.D., Lee, C.M.C., & Tarpley, R.L. (2001). Contextual fundamental analysis through the prediction of extreme returns. *Review of Accounting Studies, 6*(2-3), 165-189.

Bernard, V., Thomas, J., & Wahlen, J. (1997). Accounting-based stock price anomalies: Separating market inefficiencies from risk. *Contemporary Accounting Research, 14*(2), 89-136.

Bodie, Z., Kane, A., & Marcus, A.J. (2011). Investments and Portfolio Management. 9th ed. Singapore: McGraw-Hill.

Boudt, K., De Goeij, P., Thewissen, J., & Van Campenhout, G. (in press). Analysts' forecast error: A robust prediction model and its short term profitability. *Accounting and Finance.* Retrieved from https://lirias.kuleuven.be/bitstream/123456789/409317/1/12HRP16.pdf

Brand, D.A., Saisana, M., Rynn, L.A., Pennoni, F., & Lowenfels, A.B. (2007). *Comparative Analysis of Alcohol Control Policies in 30 Countries.* Retrieved from http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.0040151.

Callen, J.(in press). A selective critical review of financial accounting research. *Critical Perspectives on Accounting,* 2013.

Chan, L.K.C., & Lakonishok, J. (2004). Value and growth investing: Review and update. *Financial Analysts Journal, 60*(1), 71-86.

Chan, L.K.C., Hamao, Y., & Lakonishok, J. (1991). Fundamentals and stock returns in Japan. *Journal of Finance, 46*(5), 1739-1764.

Charitou, A., & Panagiotides, G. (1999). Financial analysis, future earnings and cash flows, and the prediction of stock returns: Evidence for the UK. *Accounting and Business Research, 29*(4), 281-298.

Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(1), 429-444.

Chen, P., & Zhang, G. (2007). How do accounting variables explain stock price movements? Theory and evidence. *Journal of Accounting and Economics, 43*(2-3), 219-244.

Chen, X., Kim, K.A., Yao, T., & Yu, T. (2010). On the predictability of Chinese stock returns. *Pacific-Basin Finance Journal, 18*(4), 403-425.

Cherchye, et al. (2006). *Creating composite indicators with DEA and robustness analysis: The case of the technology achievement index*. Leuven: Catholic University of Leuven and Joint Research Centre.

Cherchye, L., Moesen, W., Rogge, N., & Van Puyenbroeck, T. (2006). An introduction to 'benefit of the doubt' composite indicators. *Social Indicators Research, 82* (1), 111-145.

Cinca, C.S., Molinero, C.M., & Larraz, J.L.G. (2005). Country and size effects in financial ratios: A European perspective. *Global Finance Journal, 16*(1), 26-47.

Clark-Murphy, M., & Soutar, G. N. (2004). What individual investors value: Some Australian evidence. *Journal of Economic Psychology, 25*(4), 539-555.

Cooper, M.J., Gulen, H., & Schill, M.J. (2008). Asset growth and the cross-section of stock returns. *The Journal of Finance, 63*(4), 1609-1651.

Cunningham, L.A. (1998). *The essays of Warren Buffet: Lessons for corporate America.* Retrieved from http://bib.convdocs.org/docs/26/25645/conv_1/file1.pdf

Curtis, A. (2012). A fundamental analysis based test for speculative prices. *The Accounting Review, 87*(1), 121-148.

De Tijd. (2011, January 9). De aandelengids. *Netto (De Tijd)*, p. 1-58.

De Tijd. (2012, January 15). De aandelengids. *Netto (De Tijd)*, p. 1-60.

De Tijd. (2013, January 19). De aandelengids. *Netto (De Tijd),* p. 1-57.

Delen, D., Kuzey, C., & Uyar, A. (2013). Measuring firm performance using financial ratios: A decision tree approach. *Expert Systems with Applications, 40*(10), 3970-3983.

Dixon, W.J., & Yuen, K.K. (1974). Trimming and winsorization: A review. *Statistische Hefte, 15*(2-3), 157-170.

Dobbie, M.J., & Dail, D. (2013). Robustness and sensitivity of weighting and aggregation in constructing composite indices. *Ecological Indicators, 29*(1), 270-277.

Edirisinghe, N.C.P., & Zhang, X. (2007). Generalized DEA model of fundamental analysis and its application to portfolio optimization. *Journal of Banking & Finance, 31*(11), 3311-3335.

Estep, T. (1987). Security analysis and stock selection: Turning financial information into return forecasts. *Financial Analysts Journal, 43*(4), 34-43.

European Commission. (2003). *Third European Report on Science & Technology Indicators 2003.* Brussels.

Fairfield, P.M., & Whisenant, J.S. (2001). Using fundamental analysis to assess earnings quality: Evidence from the center for financial research and analysis. *Journal of Accounting, Auditing & Finance, 16*(4), 273-295.

Fama, E. (1965). Random walks in stock market prices. *Financial Analysts Journal, 21*(5), 55-59.

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

Feltham, G.A., & Ohlson, J.A. (1995). Valuation and clean surplus accounting for operating and financial activities. *Contemporary Accounting Research, 11*(2), 689-731.

Fox, J. (2008). *Applied regression analysis and generalized linear models*. California: Sage Publications.

Freudenberg, M. (2003). *Composite indicators of country performance: A critical assessment.* Paris: OECD publishing.

Giner, B., & Reverte, C. (2006). The risk-relevance of accounting data: Evidence from the Spanish stock market. *Journal of International Financial Management and Accounting, 17* (3), 175-207.

Greig, A.C. (1992). Fundamental analysis and subsequent stock returns. *Journal of Accounting and Economics, 15*(2-3), 413-442.

Grimm, R.C. (2012). Fundamental analysis as a traditional Austrian approach to common stock selection. *The Quarterly Journal of Austrian Economics, 15*(2), 221-236.

Halkos, G.E., & Salamouris, D.S. (2004). Efficiency measurement of the Greek commercial banks with the use of financial ratios: A data envelopment approach. *Management Accounting Research, 15*(2), 201-224.

Halkos, G.E., & Tzeremes, N.G. (2012). Industry performance evaluation with the use of financial ratios: An application of bootstrapped DEA. *Expert Systems with Applications, 39* (5), 5872-5880.

Hermans, E., Van den Bossche, F., & Wets, G. (2008). Combining road safety information in a performance index. *Accident Analysis and Prevention, 40*(9), 1337-1344.

Ho, C.-T., & Wu, Y.-S. (2006). Benchmarking performance indicators for banks. *An International Journal, 13*(1-2), 147-159.

Holthausen, R.W., & Larcker, D.F. (1992). The prediction of stock returns using financial statement information. *Journal of Accounting and Economics, 15*(2-3), 373-411.

Jacobs, R., Smith, P., & Goddard, M. (2004). Measuring performance: an examination of composite performance indicators. Centre for Health Economics. Retrieved from: http://www.york.ac.uk/che/pdf/tp29.pdf

Jaffe, J., Keim, D.B., & Westerfield, R. (1989). Earnings yields, market values and stock returns. *Journal of Finance, 44*(1), 135-148.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implication for stock market efficiency. *Journal of Finance, 48*(1), 65-91.

Jencks, S.F. Huff, E.D., & Cuerdon, T. (2003). Change in the quality of care delivered to Medicare beneficiaries, *Journal of the American Medical Association, 289*(3), 305-312.

Jiang, G., & Penman, S. (2013). A fundamentalist perspective on accounting and implications for accounting research. *China Journal of Accounting Research, 6*(4), 233-245.

JRC. (2014). *JRC publications on composite indicators and multidimensional measures.* Retrieved from http://ipsc.jrc.ec.europa.eu/?id=741

Kallunki, J.-P., Martikainen, M., & Martikainen, T. (1998). Accounting, income components and market-to-book equity ratios: Finnish evidence. *The International Journal of Accounting, 33*(3), 359-375.

Lam, M. (2004). Neural network techniques for financial performance prediction: Integrating fundamental and technical analysis. *Decision Support Systems, 37*(4), 567-581.

Lev, B., & Thiagarajan, S. R. (1993). Fundamental information analysis. *Journal of Accounting Research, 31*(2), 190-215.

Lipe, M.G. (1998). Individual investors' risk judgments and investment decisions: The impact of accounting and market data. *Accounting, Organizations and Society, 23*(7), 625-640.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance, 7*(1), 77-91.

Mironiuc, M., & Robu, M.-A. (2013). Obtaining a practical model for estimating stock performance on an emerging market using logistic regression analysis. *Social and Behavioral Sciences, 81*(1), 422-427.

Mohanram, P.S. (2005). Separating winners from losers among low book-to-market stocks using financial statement analysis. *Review of Accounting Studies, 10*, 133-170.

Morton, R.M., & Shane, P.B. (1998). The information environment and the ability of logit-based financial statement analysis to predict abnormal returns. *Accounting and Finance, 38*(1), 71-89.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., & Giovannini, E. (2005). *Handbook on constructing composite indicators: Methodology and user guide.* France, Paris: OECD publishing.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., & Giovannini, E. (2008). *Handbook on constructing composite indicators: Methodology and user guide.* France, Paris: OECD publishing.

Norusis, M.J. (2011). *Quantitative methods.* Harlow, Essex: Pearson Custom Publishing.

O'Connor, M.C. (1973). On the usefulness of financial ratios to investors in common stocks. *The Accounting Review, 48*(2), 339-352.

Ohlson, D., & Mossman, C. (2003). Neural network forecasts of Canadian stock returns using accounting ratios. *International Journal of Forecasting, 19*(3), 453-465.

Ou, J. A., & Penman, S. H. (1989). Financial statement analysis and the prediction of stock returns. *Journal of Accounting and Economics, 11*(4), 295-329.

Pérez, V., Guerrero, F., González, M., Pérez, F., & Caballero, R. (2013). Composite indicator for the assessment of sustainability: The case of Cuban nature-based tourism. *Ecological Indicators, 29*(1), 316-324.

Pert, P.L., Butler, J. R.-A., Bruce, C., & Metcalfe, D. (2012). A composite threat indicator approach to monitor vegetation condition in the Wet Tropics, Queensland, Australia. *Ecological Indicators, 18*(1), 191-199.

Prathap, G. (2012). Energy indicators and percentile ranking normalization. *Scientometrics, 91*(3), 997-1003.

Ramnath, S., Rock, S., & Shane, P. (2008). The financial analyst forecasting literature: A taxonomy with suggestions for further research. *International Journal of Forecasting, 24* (1), 34-75.

Reinganum, M. (1981). A misspecification of capital asset pricing: Empirical anomalies based on earnings yields and market values. *Journal of Financial Economics, 9*(1), 19-46.

Richardson, S., Tuna, I., & Wysocki, P. (2010). Accounting anomalies and fundamental analysis: A review of recent research advances. *Journal of Accounting and Economics, 50* (2-3), 410-454.

Rogge, N. (2012). Undesirable specialization in the construction of composite policy indicators: the Environmental Performance Index. *Ecological Indicators, 23*(1), 143-154.

Rosenberg, B., Reid, K., & Lanstein, R. (1985). Persuasive evidence of market inefficiency. *The Journal of Portfolio Managament, 11*(3), 9-16.

Ross, S.A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory, 13* (1), 341-360.

Saisana, M. (2012). *A do-it-yourself guide in Excel for composite indicator development*. European Commission, Joint Research Centre, Italy. Retrieved from http://composite-indicators.jrc.ec.europa.eu.

Saisana, M., & Tarantola, S. (2002). *State-of-the-art report on current methodologies and practices for composite indicator development.* Italy, Ispra: Joint Research Centre, European Commission.

Saisana, M., Saltelli, A., & Tarantola, S. (2005). Uncertainty and sensitivity techniques as tools for the quality assessment of composite indicators. *Journal of the Royal Statistical Society, 168*(2), 307-323.

Saltelli et al. (2008). *Global sensitivity analysis.* England, West-Sussex: Wiley.

Samaras, G.D., Nikolaos, F.M., & Zopounidis, C. (2008). A multicriteria DSS for stock evaluation using fundamental analysis. *European Journal of Operational Research, 187*(3), 1380-1401.

Sarkis, J. (2002). Preparing your data for DEA. In Avkiran, N. (Eds.), *Productivity analysis in the service sector with Data Envelopment Analysis.* Retrieved from: www.clarku.edu/~jsarkis/SarkisChapter.doc

Setiono Strong, B.N. (1998). Predicting stock returns using financial statement information. *Journal of Business Finance and Accounting, 25*(5-6), 631-657.

Sharma, M., & Sharma, P. (2009). Prediction of stock returns for growth firms – a fundamental analysis. *Journal of Business Perspective, 13*(3), 31-40.

Sharpe, A. (2004). *Literature review of frameworks for macro-indicators.* Centre for the Study of Living Standards. Retrieved from: http://www.csls.ca/reports/LitRevMacro-indicators.pdf

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

Shleifer, A., & Summers, L. (1990). The noise trader approach to finance. *Journal of Economic Perspectives, 4*(2), 19-33.

Sloan, R.G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review, 71*(3), 289-315.

Smith, P. (2002). Developing composite indicators for assessing health system efficiency. In OECD publishing (Eds.), *Measuring Up: Improving Health System Performance in OECD Countries* (pp. 295-316). Paris: OECD publishing.

Sneed, J. (1999). Estimating earnings forecasting models using fundamental analysis: Controlling for differences across industries. *American Business Review, 19*(10), 17-24.

Stober, T.L. (1992). Summary financial statement measures and analysts' forecasts of earnings. *Journal of Accounting and Economics, 15*(2-3), 347-372.

Tarantola, S. (2008). *European innovation scoreboard: New theoretical advances and visualization tools.* Retrieved from: http://publications.jrc.ec.europa.eu/repository/handle/111111111/11645

Van Caneghem, T., Van Campenhout, G., & Van Uytbergen, S. (2002). Financial statement information and the prediction of stock returns in a small capital market: The case of Belgium. *Brussels Economic Review, 45*(3), 65-90.

Wang, Y.-J., & Lee, H.-S. (2008). A clustering method to identify representative financial ratios. *Information Sciences, 178*(4), 1087-1097.

Yu, S., & Kim, S.-H. (2009). Analysis of Business Week hot-growth stocks: Momentum and fundamental investment approaches. *Journal of Asset Management, 10*(3), 192-204.

Zhang, G. (2000). Accounting information, capital investment decisions, and equity valuation: Theory and empirical implications. *Journal of Accounting Research, 38*(2), 271-295.

Appendices

Appendix A

This table represents the most commonly used indicators in the literature of FA. Of course, this research is limitative and other sources could highlight other indicators. However, we think that this table is representative for our study.

Indicator	**Reference(s)**
Acceleration of turnover growth	Beneish et al. (2001).
Assets to turnover ratio	O'Connor (1973); Ou and Penman (1989); Fairfield and Whisenant (2001); Ohlson and Mossman (2003); Lam (2004); Edirisinghe and Zhang (2007); Wang and Lee (2008); Samaras et al. (2008)
Assets to turnover ratio (Δ)	Ou and Penman (1989); Ohlson and Mossman (2003)
Average profit growth (3y)	Yu and Kim (2009)
Average turnover growth (3y)	Yu and Kim (2009)
Beta	Ou and Penman (1989); Mironiuc and Robu (2013)
Book value (Δ)	Abad et al. (2004); Chen and Zang (2007); Jiang and Penman (2013); Callen (2013)
Cash ratio	Ohlson and Mossman (2003)
Current ratio	Ou and Penman (1989); Lipe (1998); Ohlson and Mossman (2003); Lam (2004); Edirisinghe and Zhang (2007); Wang and Lee (2008); Samaras et al. (2008)
Current ratio (Δ)	Ou and Penman (1989); Ohlson and Mossman (2003)
Dividend (Δ)	Ou and Penman (1989); Ohlson and Mossman (2003); Jiang and Penman (2013); Mironiuc and Robu (2013)
Dividend yield	Clark-Murphy and Soutar (2004); Lam (2004); Mironiuc and Robu (2013)
Earnings per share (Δ) (EPS)	Estep (1987); Fairfield and Whisenant (2001); Lam (2004); Chen and Zang (2007); Edirisinghe and Zhang (2007); Yu and Kim (2009); Jiang and Penman (2013); Mironiuc and Robu (2013); Callen (2013)
EBITDA margin	Ou and Penman (1989); Fairfield and Whisenant (2001); Ohlson and Mossman (2003); Chen and Zang (2007); Wang and Lee (2008); Samaras et al. (2008)
EBITDA margin (Δ)	Ou and Penman (1989); Fairfield and Whisenant (2001); Ohlson and Mossman (2003); Chen and Zang (2007); Wang and Lee (2008); Samaras et al. (2008)
Market capitalization	Ou and Penman (1989); Abad et al. (2004); Lam (2004); Yu and Kim (2009) ; Chen et al. (2010)
Momentum	Jegadeesh and Titman (1993); Clark-Murphy and Soutar (2004); Yu and Kim (2009)
Net profit margin	Ou and Penman (1989); Ohlson and Mossman (2003); Lam (2004); Chen and Zang (2007); Edirisinghe and Zhang (2007); Wang and Lee (2008); Halkos and Tzeremes (2012)
Net profit margin (Δ)	Ou and Penman (1989); Ohlson and Mossman (2003); Lam (2004); Chen and Zang (2007); Edirisinghe and Zhang (2007); Wang and Lee (2008); Halkos and Tzeremes (2012)
Operational margin	Ou and Penman (1989); Lev and Thiagarajan (1993); Abarbanell and Bushee (1997, 1998); Al-Debie and Walker (1999); Ohlson and Mossman (2003); Chen and Zang (2007); Wang and Lee (2008); Samaras et al. (2008)
Price to book ratio (P/B)	Basu (1977); Reinganum (1981); Rosenberg, Reid and Lanstein (1985); Estep (1987); Jaffe, Keim and Westerfield (1989); Chan, Hamao and Lakonishok (1991); Fama and French (1992); Kallunki et al. (1998); Ohlson and Mossman (2003); Bodie et al. (2007); Edirisinghe and Zhang

	(2007); Chen, Kim, Yao and Yu (2010); Jiang and Penman (2013); Mironiuc and Robu (2013)
Price to earnings ratio (P/E)	Basu (1977); Reinganum (1981); Jaffe et al. (1989); Chan et al. (1991); Ohlson and Mossman (2003); Clark-Murphy and Soutar (2004); Edirisinghe and Zhang (2007); Yu and Kim (2009); Chen et al. (2010); Mironiuc and Robu (2013)
Quick ratio	Ou and Penman (1989); Ohlson and Mossman (2003); Edirisinghe and Zhang (2007); Samaras et al. (2008)
Quick ratio (Δ)	Ou and Penman (1989); Ohlson and Mossman (2003)
Return on assets (ROA)	Ou and Penman (1989); Fairfield and Whisenant (2001); Ohlson and Mossman (2003); Edirisinghe and Zhang (2007); Wang and Lee (2008); Samaras et al. (2008); Halkos and Tzeremes (2012); Mironiuc and Robu (2013)
Return on equity (ROE)	Ou and Penman (1989); Lipe (1998); Ohlson and Mossman (2003); Chen and Zang (2007); Edirisinghe and Zhang (2007); Wang and Lee (2008); Halkos and Tzeremes (2012); Mironiuc and Robu (2013)
ROE (Δ)	Estep (1987); Ou and Penman (1989); Ohlson and Mossman (2003); Samaras et al. (2008)
Solvability	Ou and Penman (1989); Ohlson and Mossman (2003); Lam (2004); Edirisinghe and Zhang (2007); Wang and Lee (2008)
Solvability (Δ)	Ou and Penman (1989); Lipe (1998); Ohlson and Mossman (2003); Samaras et al. (2008)
Turnover (Δ)	Estep (1987); Ohlson and Mossman (2003); Abad et al. (2004); Chen and Zang (2007); Edirisinghe and Zhang (2007); Yu and Kim (2009)
Turnover to current assets ratio	Wang and Lee (2008)
Turnover to equity ratio	Wang and Lee (2008); Samaras et al. (2008)
Variance dividend	Clark-Murphy and Soutar (2004); Graham and Dodd (1934) cited in Grimm (2012)
Variance profit	Lipe (1998), Mohanram (2005); Graham and Dodd (1934) cited in Grimm (2012)

Appendix B

This table provides an overview of all variables incorporated in the CI, accompanied by their calculation. At the bottom of the table, some complementary notes can be found.

Indicator	Group	Calculation
Net margin	A1	$= \frac{\text{Net profit}_{t-2}}{\text{Turnover}_{t-2}}$
EBITDA margin	A2	$= \frac{\text{EBITDA}_{t-2}}{\text{Turnover}_{t-2}}$
Operational margin	A3	$= \frac{\text{Operational profit}_{t-2}}{\text{Turnover}_{t-2}}$
ROE	A4	$= \frac{\text{Net profit}_{t-2}}{\text{Equity}_{t-2}}$
ROE (Δ)	A5	$= \left[\left(\frac{\text{Net profit}_{t-2}}{\text{Equity}_{t-2}}\right) - \left(\frac{\text{Net profit}_{t-3}}{\text{Equity}_{t-3}}\right)\right] / \left\lvert\frac{\text{Net profit}_{t-3}}{\text{Equity}_{t-3}}\right\rvert$
ROA	A6	$= \frac{\text{Net profit}_{t-2}}{\text{Assets}_{t-2}}$
ROA (Δ)	A7	$= \left[\left(\frac{\text{Net profit}_{t-2}}{\text{Assets}_{t-2}}\right) - \left(\frac{\text{Net profit}_{t-3}}{\text{Assets}_{t-3}}\right)\right] / \left\lvert\frac{\text{Net profit}_{t-3}}{\text{Assets}_{t-3}}\right\rvert$
Solvability	B1	$= \frac{\text{Equity}_{t-2}}{\text{Assets}_{t-2}}$
Solvability (Δ)	B2	$= \left[\left(\frac{\text{Equity}_{t-2}}{\text{Assets}_{t-2}}\right) / \left(\frac{\text{Equity}_{t-3}}{\text{Assets}_{t-3}}\right)\right] - 1$
Turnover (Δ)	C1	$= \left(\frac{\text{Expected turnover}_{t-1}}{\text{Turnover}_{t-2}}\right) - 1$
Acceleration turnover growth	C2	$= \left[\left(\frac{\text{Expected turnover}_{t-1}}{\text{Turnover}_{t-2}}\right) - 1\right] - \frac{1}{3}\sum_{i=t-4}^{t-2}\left[\left(\frac{\text{Turnover}_{i}}{\text{Turnover}_{i-1}}\right) - 1\right]$
Earnings per share (Δ) (EPS)	C3	$= \frac{(\text{Expected EPS}_{t-1} - \text{EPS}_{t-2})}{\text{Stock price}_{t}}$
Acceleration EPS growth	C4	$= \frac{(\text{Expected EPS}_{t-1} - \text{EPS}_{t-2})}{\text{Stock price}_{t}} - \frac{1}{3}\sum_{i=t-4}^{t-2}\left(\frac{\text{EPS}_{i} - \text{EPS}_{i-1}}{\text{Stock price}_{i+1}}\right)$
EBITDA (Δ)	C5	$= \frac{(\text{Expected EBITDA}_{t-1} - \text{EBITDA}_{t-2})}{\text{Stock price}_{t}}$

Book value (Δ)	C6	$= \left(\frac{\text{Expected book value}_{t-1}}{\text{Book value}_{t-2}}\right) - 1$
Market capitalization	C7	$= \ln(\text{stock price}_t * \text{amount of outstanding shares})$
Net profit margin (Δ)	C8	$= \left[\left(\frac{\text{Expected net profit}_{t-1}}{\text{Expected turnover}_{t-1}}\right) - \left(\frac{\text{Net profit}_{t-2}}{\text{Turnover}_{t-2}}\right)\right] \Big/ \left\lvert\frac{\text{Net profit}_{t-2}}{\text{Turnover}_{t-2}}\right\rvert$
EBITDA margin (Δ)	C9	$= \left[\left(\frac{\text{Expected EBITDA}_{t-1}}{\text{Expected turnover}_{t-1}}\right) - \left(\frac{\text{EBITDA}_{t-2}}{\text{Turnover}_{t-2}}\right)\right] \Big/ \left\lvert\frac{\text{EBITDA}_{t-2}}{\text{Turnover}_{t-2}}\right\rvert$
Dividend (Δ)	C10	$= \frac{(\text{Expected dividend}_{t-1} - \text{Dividend}_{t-2})}{\text{Stock price}_t}$
Momentum	D1	$= \left(\frac{\text{Stock price}_t}{\text{Stock price}_{t-1}}\right) - 1$
Earnings to price ratio	E1	$= \frac{\text{Net profit}_{t-2}}{\text{Stock price}_t}$
Book to price ratio	E2	$= \frac{\text{Book value}_{t-2}}{\text{Stock price}_t}$
Payback time	E3	$= \frac{(\text{Stock price}_{t-1} * \text{\# of outstanding shares}) + \text{net financial debt}_{t-2}}{\text{EBITDA}_{t-2}}$
Expected payback time	E4	$= \frac{(\text{Stock price}_{t-1} * \text{\# of outstanding shares}) + \text{expected net financial debt}_{t-1}}{\text{Expected EBITDA}_{t-1}}$
Current ratio	F1	$= \frac{\text{Current assets}_{t-2}}{\text{Current liabilities}_{t-2}}$
Current ratio (Δ)	F2	$= \left[\left(\frac{\text{Current assets}_{t-2}}{\text{Current liabilities}_{t-2}}\right) \Big/ \left(\frac{\text{Current assets}_{t-3}}{\text{Current liabilities}_{t-3}}\right)\right] - 1$
Quick ratio	F3	$= \frac{\text{Current assets}_{t-2} - \text{Inventory}_{t-2}}{\text{Current liabilities}_{t-2}}$
Quick ratio (Δ)	F4	$= \left[\left(\frac{\text{Current assets}_{t-2} - \text{Inv}_{t-2}}{\text{Current liabilities}_{t-2}}\right) \Big/ \left(\frac{\text{Current assets}_{t-3} - \text{Inv}_{t-2}}{\text{Current liabilities}_{t-3}}\right)\right] - 1$
Cash ratio	F5	$= \frac{\text{Cash and cash equivalents}_{t-2}}{\text{Current liabilities}_{t-2}}$
Cash per share	F6	$= \frac{\text{Net cash per share}_{t-2}}{\text{Stock price}_t}$
Dividend yield	G1	$= \frac{\text{Expected dividend}_{t-1}}{\text{Stock price}_t}$

Average turnover growth (3y)	G2	$= \frac{1}{3}\sum_{i=t-4}^{t-2}\left[\left(\frac{\text{Turnover}_i}{\text{Turnover}_{i-1}}\right)-1\right]$
Average profit growth (3y)	G3	$= \frac{1}{3}\sum_{i=t-4}^{t-2}\left(\frac{\text{EPS}_i-\text{EPS}_{i-1}}{\text{Stock price}_{i+1}}\right)$
Variance dividend (5y)	G4	$= \text{variance}(\text{dividend}_{t-6},\dots,\text{dividend}_{t-2})$
Variance profit (5y)	G5	$= \ln[\text{variance}(\text{EPS}_{t-6},\dots,\text{EPS}_{t-2})]$
Turnover to assets ratio	G6	$= \frac{\text{Turnover}_{t-2}}{\text{Assets}_{t-2}}$
Turnover to assets ratio (Δ)	G7	$= \left[\left(\frac{\text{Turnover}_{t-2}}{\text{Assets}_{t-2}}\right)\Big/\left(\frac{\text{Turnover}_{t-3}}{\text{Assets}_{t-3}}\right)\right]-1$
Turnover to current assets ratio	G8	$= \frac{\text{Turnover}_{t-2}}{\text{Current assets}_{t-2}}$
Turnover to equity ratio	G9	$= \left[\left(\frac{\text{Turnover}_{t-2}}{\text{Current assets}_{t-2}}\right)\Big/\left(\frac{\text{Turnover}_{t-3}}{\text{Current assets}_{t-3}}\right)\right]-1$
Beta	/	$= \frac{\text{Covariance(market return, stock return)}}{\text{Variance (market return)}}$

Note 1: The time suffix t can be 2011, 2012 or 2013 in our analysis, depending on the year we want to analyse. If we want to analyse t=2012, the most recent data available in annual accounts is for period t-2 or 2010. Hence, when using t-1 in our equations, this means that this number is based on predictions from the management of that company and financial analysts as the annual report for this year has not been published yet.

Note 2: Some equations have the stock price at time t in the denominator. We did this to deal with some extreme outliers. If we would try to calculate the growth rate of the EPS with EPS in the first period = 0,01 and EPS = 1 in the second period, the relative growth in percentages would become unmanageably large. However, when calculating growth as the difference between EPS in both periods (i.e. 0,99) and divide this by the stock price at the beginning of year t (e.g. 10euro), this value will become much more realistic. Additionally, the interpretation remains very straightforward: per euro of stock price, the EPS increased by 0,099euro. This makes the comparison across stocks very easy.

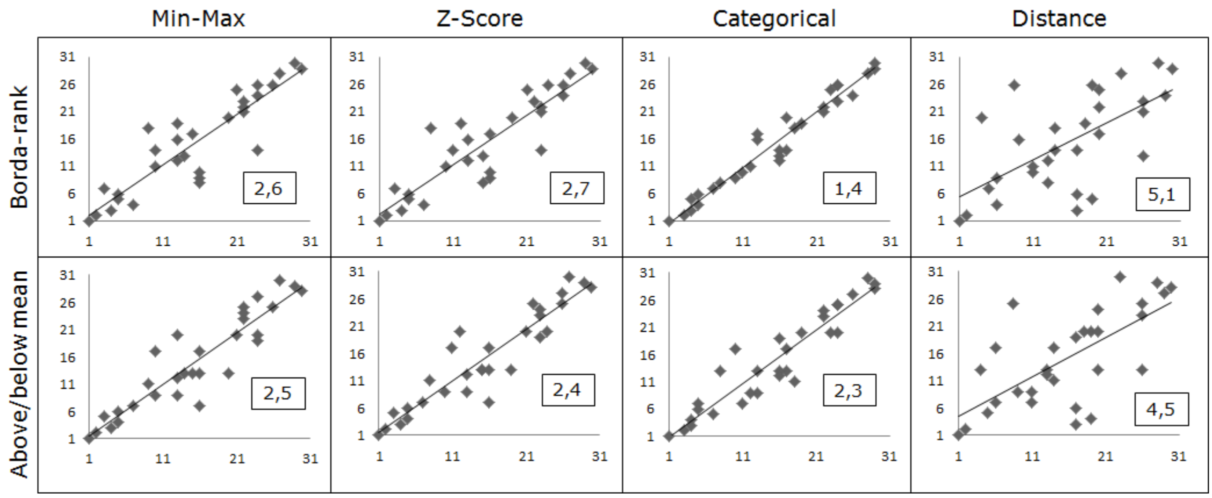
Appendix C

Total return of each stock in 2011, 2012 and 2013.

	2011		2012		2013
AB INBEV	11,75%	AB INBEV	41,49%	AB INBEV	18,80%
AGFA-GEVAERT	-62,27%	AGFA GEVAERT	3,91%	ARSEUS	81,29%
ALFACAM GROUP	-41,76%	ARSEUS	45,59%	BARCO	6,79%
ATENOR	-22,82%	ATENOR	40,47%	BEKAERT	20,75%
BEKAERT	-69,75%	BARCO	43,78%	BELGACOM	3,11%
BELGACOM	2,80%	BEKAERT	-6,98%	CFE	48,61%
CFE	-27,73%	BELGACOM	0,58%	CMB	57,35%
CMB	-22,44%	CMB	-6,12%	COLRUYT	10,67%
COLRUYT	-20,71%	COLRUYT	31,45%	DECEUNINCK	48,28%
DECEUNINCK	-60,11%	DELHAIZE GROEP	-26,24%	DELHAIZE GROUP	41,52%
DELHAIZE GROEP	-20,65%	D'IETEREN	-10,22%	D'IETEREN	22,30%
DUVEL MOORTGAT	3,36%	EURONAV	22,73%	ECKERT&ZIEGLER BEBIG	-6,02%
ELIA	7,92%	EVS BROADCAST EQUIP.	18,41%	ELIA	3,33%
EURONAV	-70,14%	EXMAR	42,43%	EURONAV	88,45%
EVS BROADCAST EQUIP.	-12,63%	FLUXYS	20,85%	EVS BROADCAST EQUIP.	7,05%
EXMAR	14,95%	FOUNTAIN	-2,16%	EXMAR	59,43%
HAMON	-43,25%	HAMON	-19,46%	FOUNTAIN	-63,96%
IBA	-40,07%	IBA	15,93%	GDF SUEZ	14,48%
LOTUS BAKERIES	5,94%	MELEXIS	24,24%	IBA	40,14%
MELEXIS	-19,95%	MOBISTAR	-42,97%	JENSEN-GROUP	17,58%
MOBISTAR	-6,55%	MOURY CONSTRUCT	4,05%	KINEPOLIS	44,65%
NYRSTAR	-46,26%	NYRSTAR	-23,77%	LOTUS BAKERIES	30,38%
OMEGA PHARMA	1,23%	PICANOL	81,78%	MIKO	31,70%
PINGUINLUTOSA	-22,40%	RTL GROUP	7,23%	MOBISTAR	-19,75%
REALDOLMEN	12,38%	SIPEF	3,26%	MOURY CONSTRUCT	-14,93%
RESILUX	-7,63%	SOLVAY	76,59%	NYRSTAR	-49,00%
SAPEC	-13,96%	SPADEL	31,44%	PICANOL	34,16%
SIPEF	-19,43%	UCB	36,02%	REALDOLMEN	20,18%
SOLVAY	-16,48%	UMICORE	33,95%	RECTICEL	13,09%
TELENET GROUP	14,56%	ZETES INDUSTRIES	-2,63%	RESILUX	76,39%
TESSENDERLO	-19,41%			ROULARTA MEDIA	14,90%
TRANSICS	3,16%			SIOEN INDUSTRIES	32,88%
UCB	24,41%			SOLVAY	6,17%
UMICORE	-18,51%			TER BEKE	23,56%
VPK PACKAGING	9,03%			TESSENDERLO	-20,08%
				UCB	28,30%
				UMICORE	-19,79%
				VAN DE VELDE	13,36%
BEL20	-20,80%		16,91%		16,54%
RANDOM3	-17,00%		16,27%		19,93%

Appendix D

Impact of normalization methods on composite rankings for year 2012 with focus on omitting the absolute level of information (with median random weights and arithmetic aggregation). The value in the square indicates the average shift in ranking.



Appendix E

Impact of adding additional restrictions to BOD models on composite rankings for year 2012. The value in the square indicates the average shift in ranking.