Master's thesis Economics & Informatics

An Aggregation Method for Stock Market Recommendations

Author:
Laurens Geleedst
289602

Supervisor:
Dr. Flavius Frăsincar

Co-reader:
Viorel Milea

Erasmus School of Economics, Erasmus University Rotterdam
October 1, 2009

Abstract

In this thesis a research is conducted to find a good aggregation method for stock market advices. This aggregation method is an improvement of the original aggregation method developed in de Market Advices Aggregation System by (Stibbe, 2007). The improvement lays in the fact that an underwriter is assigned an adjusted weight, which is based on the experience of the underwriter (amount of advices issued) and the smoothed performance figure. In addition, the method is extended with a bias correction method, which tries to correct the aggregation method by correcting for affiliated underwriters. The correction method gives different weights to underwriters who have done the Initial Public Offering for the stock which advices are aggregated.

Two types of experiments are conducted: experiments in which the bias of underwriters is tested and experiments in which the improved aggregation method is tested against the original method. Experiments are executed and results are collected over a period of 48 months, between January 2003 and December 2006. The returns of the aggregated advices are calculated for a period of one day, one week, one month and half a year. According to the results, the improved aggregation method performs significantly better on the short run, while on the large run, there is no significant improvement observed anymore.

In regards to the correction method for bias, the results are adjusted for the market capitalization of the stocks (small-cap, mid-cap, large-cap). In all cases, underwriters who have done the Initial public Offering add informational value to the aggregation method and should not be given less weight. However, at large cap companies on the short run, some extreme positive differences showed up when excluding the underwriters who have done the IPO for the stock.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

The nature of the stock markets can be described as highly volatile, which at the same time allows for high gains and also high losses. Because of the complex behavior of the stock exchange market, investment banks and independent analysts are issuing advices in regards to particular stocks. The analyst's advices are usually in the form of "buy", "hold" and "sell", which is the summary of the report the analyst wrote about the company. These financial analysts mainly work at banks, insurance companies, pension funds, and other business, and are trying to help their companies or other clients to make investment decisions. The financial analysts get their information about companies by studying public records of the companies and participating in public meetings of the companies in which they can ask direct questions to the management. In the past, financial analysts had the opportunity to have private conversations with the top-management, but stricter regulations prohibited it in later stadia.

In order to maximize profits and minimizing the risk of losing, one could follow these recommendations issued by underwriters/analysts. These advices contain information on whether one should buy, sell or hold a particular stocks. The problem arises when at a particular point in time more of these advices are issued by different underwriters from the same stock. When there is in addition no consensus among the underwriters regarding the stock, investors cannot make any decisions based on these recommendations. The most extreme cases would be a situation in which three advices are issued with all a different advice type (buy, hold or sell). Solely based on this information, investors cannot make a decisions what to do with the stock. Therefor an application, which aggregates the individual advices into one advice, could help out potential investors.

The previous described dilemma has been addressed by (Stibbe, 2007), which developed an application who collects market advices to be able to aggregate them instantly. These market advices are fetched from online sources and are saved in a database for later use. Once the user wants to have an aggregated advice for a particular stock, the application fetch the relative advices from the database and performs some calculation which yields and aggregated advice. Based on this aggregated advice, the investor can make a decision what to do with the stock. The described system is called Market Advices Aggregation System (MAAS) and was developed by (Stibbe, 2007).

1.2 Goal

As discussed earlier, good aggregation of advices could improve the decision making process of an investor and indirectly increase the profit of the investor. In the original market advices aggregating system (MAAS), aggregating is done by calculating the daily returns for all the advices given by analysts for the particular stock. Based on these daily returns, an average performance per analyst can be calculated. Once these values are calculated, the aggregated advice for the stock can be determined by averaging the analyst performance per advice type. The advice type with the maximum average performance is the aggregated advice. This is one way of having the aggregated advices for a particular stock, but as discussed in (Stibbe, 2007) it is not performing very well. In the case of a buy advice, the aggregation method gives a return between 1 and 3 percent, while in the case of a sell advice the returns are negative. Therefore in our research we are going to attempt to improve the aggregation method of advices and compare it with the original aggregation method. The research question is defined as follows:

1. *What is a good method for aggregating market advices?*

1.3 Methodology

To be able to answer the research questions, a number of tasks must be completed. These tasks are described in this section.

- **Understand the working of the MAAS application and the TOWL language.**
  In order to get a better understanding of the application we analyse the code. Furthermore we study papers such as (Milea, Frasincar, and Kaymak, 2008) to get a better understanding of the TOWL language.

- **Bug fixing and improving the working of the MAAS application.**

The MAAS application was written back in 2007. Since it was completed the application has not been updated. This means the application is wrapping the advice from the sources (Section 1.4), still by DOM parsing the sources, as explained in (Stibbe, 2007). There is a possibility the sources have changed their layout, this could lead to the case in which the advice retrieval part of the application does not retrieve the proper information anymore. Furthermore we could look at the availability of RSS with the sources and use the RSS feed of the sources instead.

- **Literature study of possible methods that can be used to aggregate the market advices.**
  As described in Section 1.2, we are going to attempt to improve the aggregation of advices. Before we propose and design a new method, we first need to study the literature, which is done in Section 2. Based on the literature study, we design and develop our own method.

- **Implement the developed aggregation method into the application.**
  Once the aggregation method has been proposed and developed it needs to be implemented in the existing application. Since the application is object-oriented, a new module can be easily added to the application.

- **Determine statistical method to measure the performance per aggregating method.**
  To compare the performance of the new proposed method with the existing aggregating method, a statistical method is needed.

When executing the last two tasks of the list above, a couple of tools are needed. When implementing a new method, knowledge of the Jena framework and the POS Tagger is necessary. Furthermore, a numerical computing environment will be necessary to compute the performance between the different aggregation methods. These tools are described in detail below:

- Jena: Jena is a framework, which is used to build semantic web applications (see `http://jena.sourceforge.net/documentation.html`). It provides a programmatic environment for RDF/RDFS, OWL, and SPARQL. Furthermore, it includes a rule-based inference engine. The framework is necessary because we are going to extend the original MAAS application, which based on Jena.

- Stanford POS Tagger: The Part-Of-Speech tagger used is the Stanford tagger (see `http://nlp.stanford.edu/index.shtml`). This tagger is highly reliable (97%) and free. Since the recommendations on websites such as Analist.nl are in human readable text, we need to transform the text into machine readable text. This way the recommendations can be added to the ontology.

- Matlab: Numerical computing environment and programming language. This program will be used to measure the performance of the designed method and of the original method.

1.4 Sources

Since the Market Advice Aggregation System is based on advices issued by brokers, the system needs to retrieve these advices. Therefore the following data sources are used by the system.

- **Analist**
  On the Analist website, recommendations for different stock exchanges are issued by brokers. Furthermore the advices of different brokers are aggregated and a consensus about the stock is displayed. From this data source advices are extracted for stocks indexed at EuroNext Amsterdam, EuroNext Brussel, EuroNext Paris, Dow Jones, NASDAQ 100, and FTSE 100.

- **GuruWatch**
  On the GuruWatch recommendations are issued by brokers for the dutch market only, namely the EuroNext Amsterdam.

- **Yahoo!Finance**
  Yahoo!Finance is used to retrieve current and historical prices of stocks to measure the performance of the brokers. Furthermore Yahoo!Finance is used to measure the performance of the aggregation method in the advice application.

1.5 Outline

The thesis start with a comprehensive literature survey about market recommendations, as described in Chapter 2. It is followed with a description of the original market advice aggregation system, the way it works and how the different parts of the application interact with each other. The way the application extracts, stores and computes advices is explained in detail. In Chapter 4, the aggregation method as implemented in the original market advice application is discussed. Furthermore different methods for improving the original method are discussed. Finally the new aggregation method is outlined and discussed in detail. Next in Chapter 4, a comparison of the the new and original method is given and in addition to the implementation of the new method in the original application is described.

Chapter 5 deals with the experiments and results. It starts by describing the experimental setup, which parameters must be determined through experiments, how the methods are compared and in which way the experiment conducted. Finally the chapter concludes with the results.

The last chapter, Chapter 6, gives answer to the research question stated in Section 1.2. In addition, future work is discussed in order to improve the existing work.

Chapter 2

Literature Survey

As described in Section 1.3, we start by studying the working of the MAAS application and the TOWL language. For this purpose we study papers such as (Stibbe, 2007), (Welty, 2006) and (Milea et al., 2007). As described in Section 1.1, since the advices issued by analysts hold for a certain period, the ontology in which the advices are stored must support temporal parts. In (Welty, 2006), the authors choose OWL to represent their ontologies. The problem with OWL as representing language arises when using fluents. Fluents are relations which hold in a certain time interval and not in others. Since ternary predicates are not allowed in OWL, the authors discuss a common approach to work around the fixed length, which is 2, of predicates, called reification. Reification knows a couple of downsides, the most important for the authors is the limited use of OWL reasoning when using reification. Therefore the authors suggest the "Four-Dimensional Approach", which is used in (Stibbe, 2007) and explained in more detail in (Welty, 2006) and (Milea et al., 2007).

Next, we investigate the effects of recommendations on investors and returns. In (Bjerring, Lakonishok, and Vermaelen, 1983), it is described that it is possible to get abnormal returns (Glossary Abnormal Returns) by following the advice of brokers. Abnormal return is the difference between the actual return and the market average return, thus it is the return on top of the market average return. This study is based on the advices issued by Canadian brokers for stocks on the TSE (Toronto Stock Exchange) and the NYSE (New York Stock Exchange) and the abnormal returns are adjusted for transaction costs. Furthermore in (Morgan and Stocken, 1998), a research done by SRI International in 1987 is cited. This research discovered that 66,7% of individual investors and professional investors rely heavily on the advices generated by brokers. Although investors rely heavily on advices, they are skeptical about the motives of the brokers and therefore apply some strategic filter over the advices. In (Morgan and Stocken, 1998), the authors develop a model of the investment decision of an investor who obtains a recommendation from a broker/analyst. The model focuses on the effects of uncertainty about the incentives of the brokers on the information content of the recommendations. As soon as the investors have the smallest uncertainty about the incentives of the brokers, for example being baised, it leads to a significant loss of the information content in recommendations. Due to the fact investors may not rely the particular advice and further issued advices anymore. Furthermore in (Barber, 2001) it is stated that following advices issued by brokers, investors can generate high returns in the case where the consensus among the brokers is the highest.

On the contrary, (Barber, 2001) it is stated that advices with low consensus perform very poor to advices with high consensus. Although the investor should rebalance his portfolio on daily basis to take advantage of this phenomena. The explanation given by the authors for the abnormal returns is: chance, poor assets pricing or most likely a market which is semistrong inefficient in which investors can yield abnormal returns with absence of transaction costs. They investigated more then 360.000 recommendations issued by brokers and analysts in the United States, extracted from the "Factset-JCF" data source (see `http://www.factset.com/`). Next the firms are placed in 5 different portfolios, from being the most favorable to the least favorable. The degree in which the firm is favorable is based on the consensus, which is based on average analysts ratings. Although the authors correct their result for market risk and price-momentum effects (see Glossary Market Risk and Momentum), they leave out the transaction costs in their results and remark that by including transaction costs the possible abnormal returns will possibly vanish.

Now that we have investigated the influence of advices issued by brokers on investors, we study whether the advices issued by brokers are biased. Since research departments, who issue advices, do not have enough money in most case to be independent, they are part of an investment bank. Because of the other activities of the investmentbanks there can be a conflict of interest between the research department and the investment bank, which can lead to advices that are biased.

In (Morgan and Stocken, 1998) and (Bjerring, Lakonishok, and Vermaelen, 1983), it is stated that advices of brokers are prone to be biased. In addition (Michaely and Womack, 1999) shows that stocks that investmentbanks analysts recommend, perform more poorly than recommendations by unaffiliated analysts. Furthermore (Michaely and Womack, 1999) state that investmentbank analysts shows significant evidence of bias and the market does not recognize the full extent of this bias. The authors came to this conclusion by testing two hypothesis over data from the "Investment Dealer Digist". From this data source, firms are identified who conducted IPO's (Initial Public Offering, see Glossary IPO), which are the first sale of stocks by a company. Next recommendations for firms during the preset time window are gathered from "First Call" database. The first hypothesis state that investmentbank analysts should have superior information about issuing advices for IPO firms, because of the due diligence (Glossary Due Diligence) process and therefore their advices should be of greater quality. The authors do not find any evidence to support this hypothesis. The second hypothesis states that investmentbank analysts have a strong incentive to recommend IPO firms that their firm recently took public regardless the quality of the advice. The author's evidence corresponds with this hypothesis. The downside of the research done in (Michaely and Womack, 1999) is the narrow time window of only 2 years. By having a bigger time window in which the recommendations performance are measured the outcome would be more justifiable.

Possible example of this bias in the past, is the bankruptcy of Enron. Months before the bankruptcy, analysts who followed the company gave strong buy advices in most of the cases (15 out of the 17 analysts), despite the fact of massive reporting of account losses and market value decreases in the media. This particular occurrence and similar examples lead to the Global Settlement agreement. This enforcement agreement, signed on April 23 of 2003, was between SEC (Security and Exchange Commission), NASD (National Association of Securities Dealers), NYSE (New York Stock Exchange) and the major investment banks. The agreement was about addressing potential conflicts of interest within investment banks, between analysts and the investment bank itself.

Despite this stricter regulations the results of (Balboa, Gmez-Sala, and Lpez-Espinosa, 2009) still confirm the existence of bias in recommendations. Furthermore the results show that bias is dependent on national level. In the UK and US the bias is still the highest among the other investigated countries. These outcomes are based on recommendation information, such as monthly returns and consensus among the recommendations, stored in the "Factset-JCF" dataset. In (Morgan and Stocken, 1998) and (Balboa, Gmez-Sala, and Lpez-Espinosa, 2009) it is stated that individual and institutional investors sense over opportunism, which corresponds to biased brokers, and correct for this bias in their investment decision. In addition, (Barber, Lehavy, and Trueman, 2007) empirically test that buy recommendations by independent research firms perform better then investment banks. On the contrary the results of the test show that hold and sell recommendations perform better when issued by investments banks. In (Barber, Lehavy, and Trueman, 2007), the authors use the "First Call" database, which contains more then 300.000 recommendations, issued by different analysts for the U.S. stock market. The recommendations are divided into 4 groups, namely the investment banks which got sanctioned by the enforcement agreement, non-sectioned investment banks, syndicate members and independent research firms. They compare the performance of the independent research firms with the 3 other categories, the performance is measured in daily abnormal return.

Contradicting the previous cited sources which states that unaffiliated recommendations perform better than unaffiliated recommendations, (Mcnichols, O'Brien, and Pamukcu, 2006) shows that unaffiliated recommendations do not perform better than affiliated recommendations. Investors discount affiliated recommendations and unaffiliated recommendations arrive in most of the time to late to provide useful trading advices. In this light affiliated recommendations are recommendations issued by analysts that have a connection with the recommended firm, while unaffiliated analysts do not have any "known" connection. The authors extend the work of (Michaely and Womack, 1999), by doing their research over a bigger time window, while using the same data source, namely the "First Call" data source. The authors conclude that unaffiliated analyst recommendations do not earn higher abnormal buy-and-hold returns than recommendations from affiliated analysts at intervals of three, six or twelve months after the recommendation. Instead, their results change per observing year, while there is no significant difference in abnormal returns between affiliated and unaffiliated recommendations. This result is in conflict with (Michaely and Womack, 1999), who discovered a significant bias in affiliated recommendations, since affiliated recommendations performed worse than unaffiliated recommendations. Furthermore they remark that buy recommendations issued by affiliated brokers are discounted by investors, which is in line with the results of (Michaely and Womack, 1999).

Although there is no consensus among the authors of different papers whether their is bias or not, (Balboa, Gmez-Sala, and Lpez-Espinosa, 2009) developed a correction method. This correction method is based on the ratio between buy and sell recommendations and the theoretical balanced situation in which buy and sell recommendations are equal. The further the ratio of a country differs from the theoretical balanced situation the more the recommendations issued in that particular country must be corrected for bias. In other words the proposed correction method, corrects for bias introduced by optimism of financial analysts. Furthermore (Balboa, Gmez-Sala, and Lpez-Espinosa, 2009) results show that recommendations corrected for bias deliver significant higher returns than raw recommendations. In addition the results show that countries with minor bias in their recommendations have higher returns on those recommendations. Although the proposed correction method by (Balboa, Gmez-Sala, and Lpez-Espinosa, 2009) is unique, the method itself is not very strong, since it is based one a balanced situation in which the buy and sell recommendations are evenly distributed. Which is nothing near reality, since the market is never really consistent and a lot of external factors determine whether the price of a stock is rising or decreasing and thus whether a buy or sell advice is issued.

As we have seen in this literature study, there is not really a consensus among the research about investors relying on advices issued by brokers. In (Bjerring, Lakonishok, and Vermaelen, 1983), (Morgan and Stocken,

1998) and (Barber, 2001), it is stated that investors can yield abnormal returns when following recommendations issued, although in (Barber, 2001) investors yield these high returns only when following recommendations with high consensus among the analysts. In addition to these results, (Barber, Lehavy, and Trueman, 2007) and (Michaely and Womack, 1999) shows that underwriter recommendations and affiliated analysts recommendations perform worse than unaffiliated/unbiased recommendations. Thus differentiating between the issuer of the advice when aggregating advices is of great importance to yield higher return. In addition (Bjerring, Lakonishok, and Vermaelen, 1983), (Morgan and Stocken, 1998) and (Michaely and Womack, 1999) underwrite the assumption of potential bias in recommendations issued by investments banks. (Morgan and Stocken, 1998) concludes that the smallest uncertainty by the investor in the incentive of the analyst makes the informational content of the recommendation less valuable. Furthermore (Michaely and Womack, 1999) state the market does not recognize the full extent of bias in recommendations. Finally in (Balboa, Gmez-Sala, and Lpez-Espinosa, 2009) it is stated, despite the stricter regulations, there is still bias in recommendations issued. Furthermore (Balboa, Gmez-Sala, and Lpez-Espinosa, 2009) propose a method to correct for this bias and by applying this method, they conclude that corrected recommendations yield higher abnormal returns than uncorrected recommendations.

By taking this information into consideration when designing a new aggregation method for the MAAS application, the method should take extra notion in the case where the consensus among the analysts is the lowest and should make a distinction between affiliated and unaffiliated recommendations and thus correct the advice for biased influences.

2.1 Global Analyst Research Settlement

The Global Analyst Research Settlement, was an enforcement agreed between the U.S. Security and Exchange Commission (SEC), New York Stock Exchange, Nasdaq Stock Market and ten of the biggest investment firms in the United States. The settlement was reached on April 28th of 2003 and was agreed to address issues in conflict of interest within their business (Security and Exchange Commission, b).

The central issue, which emerged this settlement, was the conflict of interest between the investment banking department and analysis department within the ten biggest investment banks in the United States. Several cases where known, in which analysts issued fraudulent research reports in violation of various sections within the Securities Exchange Act of 1934 (Security and Exchange Commission, a). These violations occurred in banks such as: Credit Suisse First Boston Corp and UBS Warburg LLC. These fraudulent reports where mainly issued by high fees and pressure from within the investment bank department.

The Global Settlement incorporates mainly regulations to prevent abuse of investment reports by pressure of investment bankers on analysts. Examples of these regulations are: insulate banking and analyst departments from each other physically, budget allocation for analyst department must be independent of investment bank department. In addition analysts are prohibited to go with bankers to sale pitches (planned presentation for a product to close a sale for the same product) during the advertising and promotion period of IPO's. Besides these regulations, the quite period in which persons are legally prohibited to promote an upcoming IPO is increased from 25 to 40 days (Security and Exchange Commission, b).

Globally speaking this is the one and only introduced regulation which minimize the conflict of interest between the bankers and analysts and thus minimize potential bias in the recommendations.

Chapter 3

Market Advice Aggregation System

3.1 Introduction

As described in Section 1.1, at a particular point in time multiple advices will be most likely issued regarding one stock. Once the consensus among the different recommendations is low, it can be quite difficult for investors to make a decision based on these recommendations. Therefore an application which aggregates these market advices and form one advice, could possibly help investors to make a decision. In this light (Stibbe, 2007) developed an application which aggregates the market advices issued. In this chapter, an overview will be given of the application. Furthermore the different components of the application will be discussed in detail.

3.2 Overview of the Application

The MAAS application consists of two parts, namely the information extractor and the advice application part. First of all information about advices must be gathered from the sources: Analist.nl, GuruWatch.com, described in Section 3.3. The information extractor extracts the necessary information from these sources, as discussed in Section 3.4. The information extractor uses a POS-Tagger (see Section 1.3) to extract information from the sources. Next, these extracted advices are stored in an ontology. Once the ontology is filled with advices, the advice application (Section 3.5), aggregates them and generates a new advice.. The schematic view of the application is displayed in Figure 3.1.

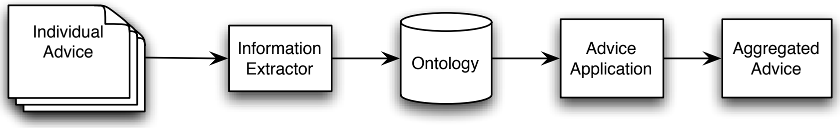

Figure 3.1: Schematic view of the working of the MAAS Application.

3.3 Input of the Application

The application is built around the sources described in Section 1.4. The application grabs the advices issued on these sources. In the original application advices issued on the Analist website are fetched through a grabber called the Market Advice Grabber in the application. This module is based on DOM parsing the necessary pages. These pages contain the information about the advice. By DOM parsing, the page is converted into an XML tree, within this XML tree one can search nodes which contains the information. At the time of the application written, the RSS feed was not updated on regular basis, therefore (Stibbe, 2007) make use of the complicated process of DOM parsing. At the time of speaking the RSS feed is update as soon as a new advice is issued (see `http://www.analist.nl/rss/analist_nl.php`).

This makes it possible to extract the advice from the website through the RSS feed. An example of an advice issued by Analist.nl through the RSS feed is displayed in Figure 3.2. The advice concerning whether one should buy/hold or sell is stated in the description tag. This way, the process of gathering advices is simplified, due to the mark up of the feeds in XML.


```
<item>
<title>RBC Capital Markets: APPLE INC. kopen.</title>
<link>http://rss.feedsportal.com/c/637/f/8254/s/1047096/story01.htm</link>
<description>(Analist.nl) Toronto - Op 6-5-2008 herhalen de analisten van RBC Capital
Markets hun koopadvies voor APPLE INC. (ISIN: US0378331005 / TICKER: AAPL).Het
12-maands koersdoel voor APPLE INC. wordt opwaarts bijgesteld van 200.00 USD naar
220.00 USD. In 2006 bedroeg.. (lees verder op: www.analist.nl)</description>
<guid isPermaLink="false">http://www.analist.nl/rss.php?id=51910</guid>
</item>
```


Figure 3.2: RSS feed of advice *51910*.

Since the application has not been updated for over a year, old advices must first be added to the ontology, before new advices can be added by RSS. Since the RSS feed only contains the latest advice, DOM parsing is still required to extract the old advices from the source.

Once the advices are extracted from the feed, the system must understand what is stated inside the advice. The machine cannot understand the words like we do. Therefore natural language processing is necessary to transform the human readable text into machine readable text. This process is explained in Section 3.4.

3.4 Information Extractor

In Section 3.3, we have discussed the sources from which the advices are extracted. Furthermore we have discussed how the advices are fetched from the source. Once the advice is grabbed from the particular source (Analist.nl), we need to transform the human-readable text into machine-readable information.

3.4.1 Natural Language Processing

The goal of Natural Language Processing (NLP) is to extract information from human-readable text and transform the text into machine-readable information. There are two main methods for this transformation process, namely the association-based approach and the pattern-based approach.

Before one can apply one of those methods, the extracted text must first be tagged, which is done with a part-of-speech tagger. The part-of-speech tagger annotates every single word. The tagger classifies each word in eight categories: verbs, nouns, pronouns, adjectives, adverbs, prepositions, conjunctions and interjections (Banko and Moore, 2004). There are different implementations of PoS-Taggers, suchs as Microsoft's HMM Tagger (Banko and Moore, 2004) and the Stanford PoS-Tagger. The Stanford PoS-Tagger is based on the Penn Treebank (Marcus, Santorini, and Marcinkiewicz, 1994). The Penn Treebank is a large annotated corpora, consisting 4.5 millions of words, which is used to tag the words. In Table 3.1 a small part of the tagset is displayed from (Marcus, Santorini, and Marcinkiewicz, 1994).

Table 3.1: Part of the Penn Treebank tagset.

1.	CC	Coordinating conjunction
2.	CD	Cardinal number
3.	DT	Determiner
4.	EX	Existential there
5.	FW	Foreign word
6.	IN	Preposition/subordinating conjunction
7.	JJ	Adjective

Due to the fact the MarketAdvices source is discontinued, the non-english source Analist.nl is used. This prohibits one to use part-of-speech tagger which uses english tag sets, such as the stanford tagger. The Analist.nl is the only source which summarizes advices for different stock exchanges, including the Amsterdam Stock Exchange (AEX), this source is therefor vital for the application. Because the source is foreign, english tag sets will not work, thus a different part-of-speech tagger is used, which includes a dutch tag set, called TreeTagger. This part-of-speech tagger implements probabilistic tagging method as explained in (Schmid, 1997). When using this method incorporated with a dutch tagging set and the following advices: "Op 6-5-2008 herhalen de analisten van RBC Capital Markets hun koopadvies voor APPLE INC. (ISIN: US0378331005 / TICKER: AAPL). Het 12-maands koersdoel voor APPLE INC. wordt opwaarts bijgesteld van 200.00 USD naar 220.00 USD.", one gets the output displayed in Table 3.2.

Table 3.2: Advice 51910 PoS-Tagged by the TreeTagger PoS-Tagger.

Word	**Pos**	**Lemma**
Op	nounsg	¡unknown¿
6-5-2008	num__card	@card@
herhalen	verbprespl	herhalen
de	det__art	de
analisten	nounpl	analist—analiste
van	prep	van
RBC	adj	¡unknown¿
Capital	nounsg	¡unknown¿
Markets	nounpl	¡unknown¿
hun	det__poss	hun
koopadvies	nounpl	¡unknown¿
voor	prep	voor
APPLE	adj	¡unknown¿
INC	nounpl	¡unknown¿
ISIN	nounsg	¡unknown¿
US0378331005	nounsg	¡unknown¿
AAPL	nounsg	¡unknown¿
wordt	verbpressg	worden
opwaarts	adj	opwaarts
bijgesteld	verbpapa	bijstellen
van	prep	van
200.00	num__card	@card@
USD	adj	¡unknown¿
naar	prep	naar
220.00	num__card	@card@
USD	adj	¡unknown¿

To extract the correct information from the advice using the pattern-based approach, we use the tagged advice as displayed in Table 3.2 as an example.

The first cardinal numeric value is the issuing date. The verb following the cardinal number indicates whether the advice is upgraded, downgraded or held constant. Next one searches for the advice type, which equals words like “kopen”, “houden” or “verkopen”. In case of an upgrade/downgrade, search for a second occurrence of these words “kopen”, “houden” or “verkopen”. The second occurrence of the advice type words, indicates to what it is upgraded or downgraded. The price target is found by searching for the first cardinal number after the word “koersdoel”. If the price target is followed with the word “niet”, there is no price target. In case of upgrading or downgrading the price target, the final price target will be preceded with the word “naar”. The ISIN identifier can be found by searching for the word ISIN and the following cardinal number is the code. The ISIN identifier (International Securities Identifying Number) is an international identification number for companies. Once one extracted this number from the advice and matches it with the numbers stored in the miniFDO ontology (see Section 3.4.2), one knows for which company the advice is issued. Finally the broker who issued the advice must be extracted. This is done by searching for the first occurrence of the word “van”. Due to the length of the name of the broker is unknown, all the words between the first occurrence of the word “van” after the word “analist” and the the next occurrence of a verb, noun or adjective is chosen as the name of the broker. By running this method when wrapping over more than hundred advices, there where no errors.

3.4.2 Ontologies Used in the MAAS Application

Quick decision making is a key ingredient of the MAAS application. To support quick decision making, the application makes use of an ontology to store the advices. Because of the temporal behavior of advices, the advices hold for a certain period and therefore the ontology must support temporal parts. As we have seen in the literature survey (Section 2), (Milea, Frasincar, and Kaymak, 2008) has developed a four-dimensional approach of the Web Ontology Language (OWL), called tOWL.

The ontology in the application is based on the tOWL language. The ontology in the application is called the “Financial tOWL ontology” and consists of three parts.

- **miniTOWL**
  The miniTOWL ontology contains the tOWL language specifications (Milea, Frasincar, and Kaymak, 2008), which is used by the other two ontology files.

- **miniFDO**
  The miniFDO ontology contains static information about companies.

For each stock listed company in the miniFDO ontology a set of properties is listed, such as the name, quoteSymbol, currency and more. These properties are used to match companies in the advices issued with stock listed companies in the ontology.

- **advices**
  Finally the advices ontology contains all the advices issued by brokers. For each advice a time slice is created. A time slice refers to exactly one time interval, while a time interval can be used by more than one time slice. A time interval consists of two time points; the starting point and the ending point. Once a time slice is created for a new advice, the time slice refers to the corresponding time slice of the broker, the company for which the advice is issued, the price target, the advice type and the corresponding time interval.

3.4.3 Updating the Ontology

Once the advice is retrieved from the source and tagged by a part-of-speech tagger, the correct information must be extracted from the tagged advice. The most well known methods according to (Gomez-Perez and Manzano-Macho, 2003) are:

- **Pattern-based extraction**
  A relation is recognized when a sequence of words in a text matches a particular pattern.

- **Association rules**
  Association rules, initially used for data mining database, are defined as frequent concepts co-occur in texts. When the correlation between concept x and concept y are high, once can imply x when y occurs.

- **Conceptual clustering**
  Concepts are clustered according to the semantic difference between the concepts.

- **Ontology pruning**
  When using the ontology pruning method, the objective is to build a domain specific ontology from different sources. This is done by building a core domain, extending the domain with a dictionary filled with domain specific words and finally prune those words which are not domain specific. This is done by analyzing the domain specific corpus and compare it with the ontology. Once the ontology contains concepts which do not occur in the domain specific corpus, these concepts are eliminated from the ontology.

- **Concept learning**
  A defined taxonomy is incrementally updated with newly acquired concepts from analysed text.

In the MAAS application, the pattern-based approach is used to extract information from advices. Due to advices have a constant structure, the pattern-based approach is more suitable. As described in Section 3.4.1 the correct information is first annotated by the part-of-speech tagger. Furthermore, the correct information is extracted by using a pattern-based method. Once the information is extracted, the ontology must be updated by incorporating the new information. As we have seen in Section 3.4.2, there are three ontologies, in which two ontologies contain the information about the companies and the advices, namely miniFDO and advices. Once the correct information is extracted from the advice, this information must be added to the ontologies:

- **Updating miniFDO**
  In case that there is an advice issued by a broker, which is not known in the miniFDO ontology, the broker must be added to the ontology. In addition, when the company is not known (e.g., no matching record with the ISIN identifier), the company must be added.

- **Updating advices**
  To be able to store an advice into the advice ontology an instance of the analyst who issued the advice and the company for which the advice is issued must be available in the miniFDO ontology.

3.5 Advice Application

As discussed in Section 3.2, the application consist of two parts. The advice application part, which is the second part of the application provides the user with an advice, as requested by the user. A schematic view of the advice application is displayed in Figure 3.3.

As one can see in Figure 3.3, the advice application consist of 4 modules. These modules are based on a model-view-controller framework. At first the GUI module, which is used to make the application user friendly, is initiated. The GUI handles the graphical user interface part of the application. Through the graphical user interface, users can specify for what company they want to get an advice, by specifying the company and the time window.

Once the request has been made, the GUI modules passes it over to the NewAdvice Module. The NewAdvice module is the application logic (controller). It evaluates the request passed on by the user interface and it sends requests to the models:

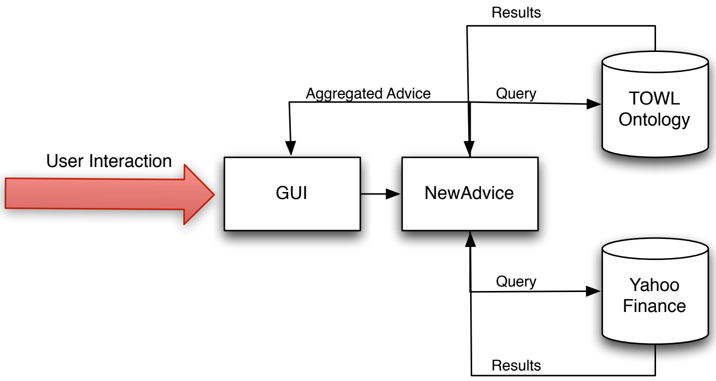

Figure 3.3: Schematic view of the advice application part.

- YahooFinance
  The YahooFinance module is used to retrieve current and historical price of the requested company. This information is needed to compute a new aggregated stock recommendation.

- TOWLOntology
  The TOWLOntology module is used to retrieve information from the two ontologies (miniFDO and advices). This information can be advices stored in the advices ontology or company information stored in the miniFDO ontology.

Next, it retrieves data from the models and manipulates the data in order to fulfill the request made by the end user. Once the proper action is performed and thus leading to the requested information by the user, this data is send back to the user interface and presented to the end user.

In short, once the user requests an advice about a particular company within a particular time window, the user interface retrieves the request and send it down to the NewAdvice module. The NewAdvice module handles the request, in which the model is addressed and the newly formed advice is passed to the user interface.

3.6 User Interface

In order to make the application user friendly a graphical user interface is developed by (Stibbe, 2007). This graphical interface is based on the graphical interface guidelines outlined by (KDE Techbase). All the interfaces within the application are built up through a wizard, which guides the user through the information gathering process. By using wizards, less buttons can be used and more overview is given to the user.

The application consist of 4 tabs, in which different aspects can be calculated.

- Standard: In the standard tab screen, a simple method is used to calculate an aggregated advice for a selected company.
- Portfolio: In the portfolio tab, aggregated advices are calculated for all the companies added to the portfolio.
- Advanced: In the advanced tab, aggregated advices are calculated using a more advanced method.
- GICS: In the Global Industry Classication Standard (GICS) tab, the performance per advisor can be viewed by looking up the GICS code.

The most used screens are the standard and advanced tabs. These are schematically displayed in Figures 3.4 and 3.5.

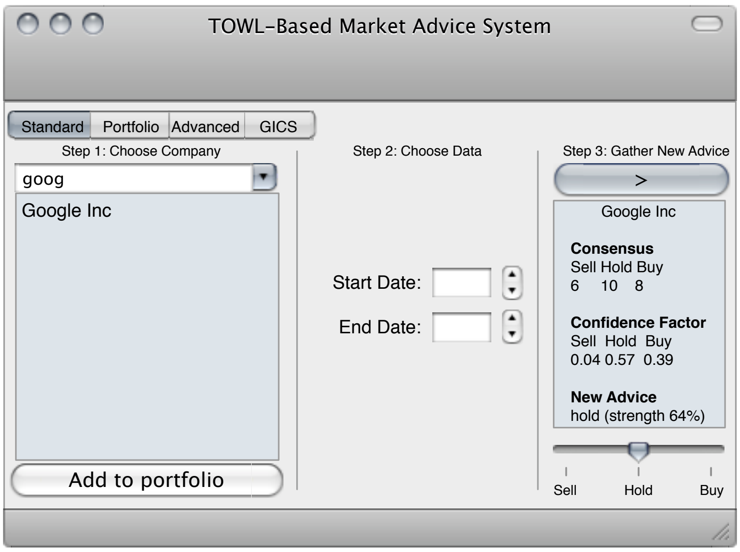

Figure 3.4: Schematic view of the advice application part.

In Figure 3.4, the user starts by selecting the company for which it wants an aggregated advice. Next, it selects the start and end date for which it wants an aggregated advice. Finally, by pushing the start > button, the aggregated advice is calculated. According to (KDE Techbase), users give the most attention to the first and last row when information is presented, therefore the generated advice is displayed in the last row.

In Figure 3.5, the advanced tab wizard is displayed. The advanced tab, calculates a new advice for the selected company based on the performance of the underwriters within the same industry. It looks very familiar to the

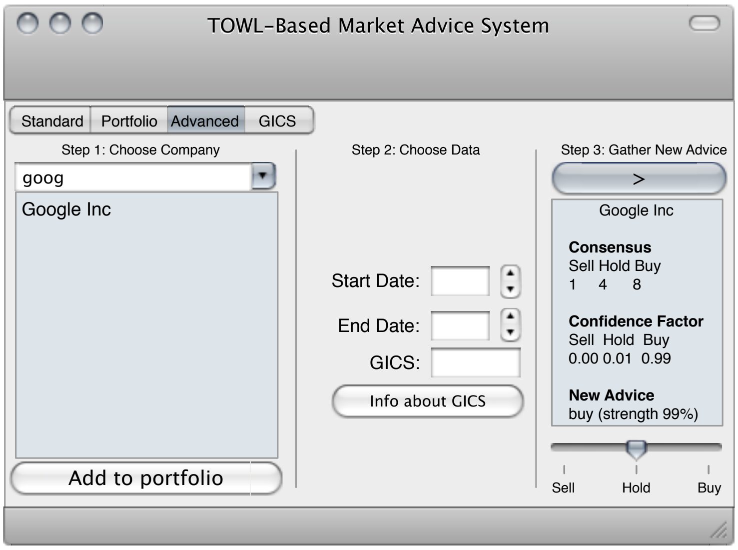

Figure 3.5: Schematic view of the advice application part.

standard tab displayed in Figure 3.4. The only addition in the GUI, is the GICS text field and the "more info" button. Once the user selects a company in the first step, the GICS code is automatically filled into the text field. Next, the user can specify the time period for which it wants the aggregated advice and by clicking the start button, this advice is calculated using the advanced method.

3.7 Conlusion

The Market Advice Aggregation System (Stibbe, 2007), aggregates market advices into one aggregated advice. These market advices are fetched from publicly available sources. Once the advices are extracted, they are processed by a natural language processor to extract the relevant information. This information is added to the ontology by using a pattern-based extraction method. Once the ontology is filled with advices, the advice application part can aggregate advices at a particular point in time and generate one aggregated advice to the end user.

Chapter 4

Aggregation Method

4.1 Introduction

As described in Section 1.1, due to the complexity of the stock market, following advices issued by brokers can increase the return on investment. This assumption is confirmed by different papers, as we have seen in the literature (Chapter 2). The authors of (Bjerring, Lakonishok, and Vermaelen, 1983), (Morgan and Stocken, 1998) and (Barber, 2001) have concluded that investors can obtain abnormal returns by following advices. In addition to this result, some authors suggest that these abnormal returns can only be yield by following advices where the consensus among the brokers is high. One can infer from these results that aggregation of advices by analysts is important to yield possible abnormal returns. In this light (Stibbe, 2007) developed an application which aggregates the market advices issued. In this chapter the original aggregation method is discussed in detail in Section 4.2. It is followed by an overview of methods to correct potential affiliation in recommendations. At the end of the chapter the improved aggregated method is outlined (Section 4.3).

4.2 Advice Aggregation in the Original Application

Before explaining the original aggregation method, a schematic view of the advices, advisors and companies is given in order to illustrate the way the aggregation method works. In Figure 4.1, this schematic view is displayed. In this example there is a timeline, there are 7 advisors and 7 companies. At a particular time point t, one could want an aggregated advice for a particular company $C3$. At that particular time point t, all the advices are taken, which still hold, meaning the advices that are issued with a maximum of six month before time point t. For each of the advisors the performance is measured. This is done by calculating the returns of previously issued advices.

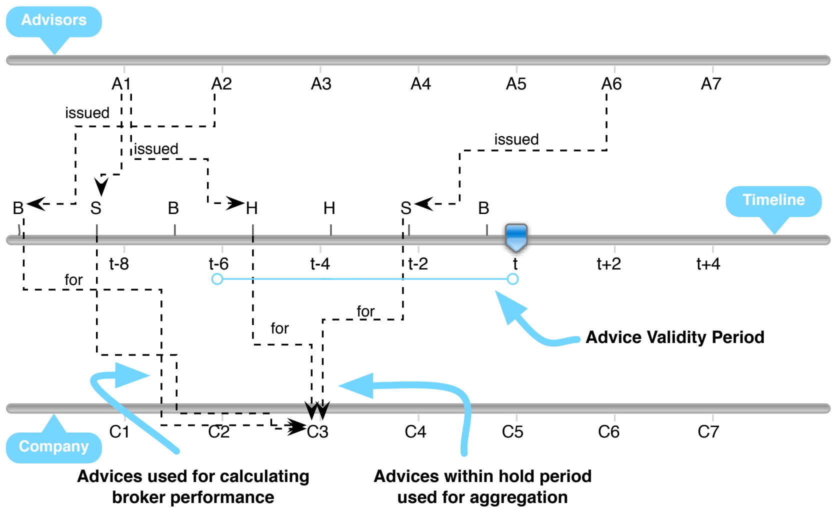

Figure 4.1: Schematic view of the relation between advices, advisors and companies.

Due to the advice requested spans a time window up to six months before, the advices gathered in this six month time window are from different time periods. In order to compare and aggregate the advices, the return per day per advice must be calculated, as explained by (Stibbe, 2007). For a buy and hold advice the return is calculated as displayed in Equation 4.1 and for a sell advice as displayed in Equation 4.2. These Equations are asymmetrical, because the performance differs per advice type. In case of a hold and a buy, the performance is positive in case of an increase in price, while the performance is negative for a sell advice in case of a price increase.

$$dailyReturn = \frac{price_{t+1} - price_t}{price_t} * 100 \tag{4.1}$$

$$dailyReturn = \frac{price_t - price_{t+1}}{price_{t+1}} * 100 \tag{4.2}$$

In Equation 4.1 and 4.2, *price* is the price of the stock and t defines the day and *t+1* is the following day. The calculation of the daily return is necessary to calculate the average return per day for each advice.

$$averageReturnPerDay = \frac{\sum_{i=1}^{x} dailyReturn_i}{x} \quad (4.3)$$

In Equation 4.3, the average return per day (for days 1,..,x) is calculated by summing the *dailyReturn*, as calculated in Equation 4.1 and 4.2. Next, it is divided by the amount of the days the advice is valid, denoted by x.

Once the average return per day is calculated per advice, one can measure the performance of the broker, as displayed in Equation 4.4.

$$brokerPI = \frac{\sum_{i=0}^{k} averageReturnPerDay_i}{k} \quad (4.4)$$

In Equation 4.4, the performance indicator of the broker is calculated. This is done by summing the *averagereturnperday* for all the advices k, issued by the broker (starting from 2003 and onwards). This performance indicator is of importance when delivering new advices to the end user, since in case of a negative performance indicator the broker is excluded from the advice aggregation.

Once the performance indicator of the broker is calculated, the confidence factor can be determined. This confidence factor is used when generating a new advice by the advice application. This advice is based on which advice (sell, buy or hold) has the greatest confidence factor. In Equation 4.5 the equation to calculate the confidence factor is displayed.

$$CF_{[buy,hold,sell]} = \frac{\sum_{i=0}^{l} brokerPI^{i}_{[buy,hold,sell]}}{\sum_{b=0}^{m} brokerPI^{b}_{buy} + \sum_{c=0}^{o} brokerPI^{c}_{hold} + \sum_{d=0}^{p} brokerPI^{d}_{sell}} \quad (4.5)$$

In Equation 4.5, the formula is displayed to calculate the confidence factor. This confidence factor is calculated for each advice type (buy, hold and sell). In the case of calculating the confidence factor for the buy advice type, the sum is taken over all the performance indexes of advisors who issued a buy advice for the particular stock, during the valid period as displayed in Figure 4.1. Next this value is divided by the sum of all the performance indexes of brokers who have issued an advice for the particular stock (brokers issued buy advice are denoted by m, brokers issued hold advice are denoted by c etc). Once Equation 4.5 is executed for every advice type, the advice type with the biggest outcome and thus the biggest confidence factor is chosen as newly aggregated advice.

Once the confidence factor is calculated and the advice is determined from this outcome, one can determine the strength of the advice. In the Market Advice Aggregation Application (Stibbe, 2007), there are two methods for determining the strength of the advice.

- Strength R
  The R strength method has output between -1 and 1 and is calculated as follows:
  $$Strength_r = (CF_{buy} * 1) + (CF_{hold} * 0) + (CF_{sell} * -1) \tag{4.6}$$
- Strength V
  Method V calculates the strength in contrast to the other two types of advices:

$$Strength_v = \frac{\begin{array}{c}(max(CF_{buy}, CF_{hold}, CF_{sell}) - min(CF_{buy}, CF_{hold}, CF_{sell}) + \\ max(CF_{buy}, CF_{hold}, CF_{sell}) - median(CF_{buy}, CF_{hold}, CF_{sell}))\end{array}}{2} \tag{4.7}$$

In some cases it can occur that there is no possibility of calculating the confidence factor. This can occur when there are no brokers who issued advices for a particular stock in a particular time window. In this case the advice application will always generate a hold advice. Both methods present drawbacks, which will be discussed in Section 4.3.

4.3 Improved Advice Aggregation Method

In Section 4.2, we have described the aggregation method currently used in the MAAS application (Stibbe, 2007). This method is based on confidence factors, which are based on the performance indicator of a broker. The performance of a broker is based on the average return per day per issued advice.

4.3.1 Methods for Discovering Potential Bias

As we have discussed in Section 2, advices issued are prone to be biased. Despite stricter regulations in the United Stated, (Balboa, Gmez-Sala, and Lpez-Espinosa, 2009) state that advices are still possibly biased. A possible explication of bias lays in the fact investment banks have clients, which they want to keep and therefore do not want to issue advices which may be bad for their client. In this light there can be made a disjunction between the regional and national investment banks, since national firms have more reputation capital (see Glossary Reputation Capital), which lead to better performing advices in contrast to regional counterparts.

Possible bias can occur in cases when advices are issued for companies who are in some way connected with the firm who issued the advice. In order to correct for the bias in advices, one must discover relationships between the firms. Possible bias can be detected in different settings.

- Search for relationship between investment banks and issuing company. This relationship could possibly indicate bias.
- Investigate whether the investment bank has done the Initial Public Offering (IPO, see Glossary IPO) for the issuing company. In case the investment bank has done the IPO, the quality of the advices in regards to the issuing company could contain bias.
- Investigate whether the investment bank has done buy recommendations just before the filing date of a bankrupt firm. According to the research of (Allen and Isreal, 2003), quite a few investment banks issued buy recommendations months before the filling date of a bankrupt firm, while the media states possible bankruptcy of the particular firm.

Search for relationship between investment banks and issuing company

In order to discover a public relationship between the stock and the issuing firm, one needs to search for news articles in which such a relationship can be discovered. Since the information on the web is overwhelming, information extractors are designed as a service to gather relevant information from multiple sources, as described by (Hongwei Zhu and Madnick, 2001). Newspaper and internet portals are example of information extractors. These extractors are mostly general, while we want specific generators, one can obtain specific information from different sources.

We are looking for information extractors to search for news item in which the relationship between the stock and the issuer can be determined.

Table 4.1: Advice 53873; Bank of America issues advice for Invitrogen.

Date	Bank	Stock	Advice
17 June	Bank of America Securities	INVITROGEN	=Hold

Searching the internet delivers us the following news item:
"Invitrogen to buy Applied Biosystems for $6.7 bln - .....To fund the deal, Invitrogen plans to use cash on hand and proceeds from a fully underwritten **debt financing** from **Bank of America**, UBS Investment Bank and Morgan Stanley. The total debt of the combined company would be $3.5 billion, including $2 billion in new debt....." (see `http://news.yahoo.com/`

s/nm/20080612/bs_nm/appliedbio_invitrogen_dc_5). This could indicate a possible conflict of interest between different departments within the Bank of America and thus lead to biased recommendations for the Invitrogen stock.

To gather such relationships, information extractors are necessary. This information extractor should search for news headlines which contain the name of the stock and, next, it should search through the complete message in order to discover the name of the issuer, as displayed in Figure 4.2.

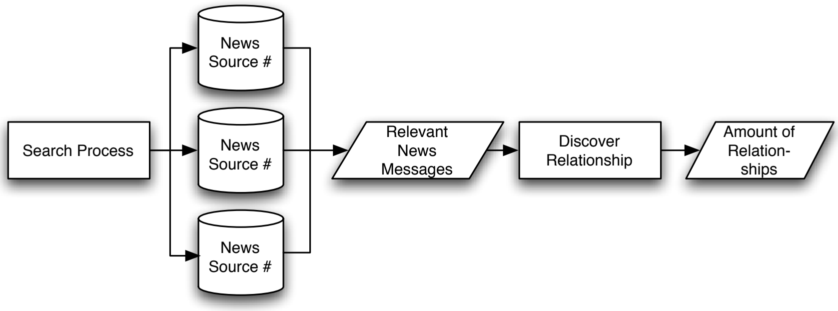

Figure 4.2: Schematic view of the extraction of information from internet sources.

As displayed in Figure 4.2, the information extractor process starts by a search process. In this search process the stocks are defined and also the brokerage firm (Glossary Brokerage Firm), these values combined is the search string. Next the search string is passed onto different news sources (i.e. yahoo finance, bloomberg et al). The information aggregation fetches the relevant results from the news sources. Next the messages are investigated on potential relationships between the analyst and the stock, which might lead to conflict of interests.

However, this method can only detect potential relationships by searching publicly available information, one can imagine that most of the relationships between companies are not publicly known and therefore this method is not sufficient.

Investigate whether the investment bank has done the IPO for the issuing company

Another method to detect possible bias, is by searching for Initial Public Offering processes. As described in (Michaely and Womack, 1999), investment banks who have done the due diligence process for a company should have superior information and therefore their advices for a particular company should perform better. However the contrary is true, investment banks having done the due diligence for a particular company have the incentive to publish positive recommendations for companies for whom due diligence is done by the investment bank, regardless the quality of this recommendation (Michaely and Womack, 1999).

Although this method is very promising, it has one major drawback. The method cannot be applied to every stock listed company, especially older stock listed companies.

Investigate bias of investment banks by looking at advices issued for bankrupt firms

In (Allen and Isreal, 2003), the extent of bias in the financial world is investigated. They observe the recommendations for companies prior to their filling date. The data shows that the amount of sell recommendations a couple of months prior to the filling date of the bankrupt firm is equal to the amount of sell advices in the total universe of the publicly traded stocks. In addition (Allen and Isreal, 2003) remarks that despite the description (to each advice a small description is added by the underwriter) added to the recommendation making notice of possible bankruptcy, the advice is in most cases still positive (buy).

Although this method looks promising in detecting bias investment banks, the process of detection bias can only be done while examining historical data and gives no guarantee to detect the company bias.

Conclusion

From these three possibilities, detecting possible bias by examining IPO data is the easiest method to quantify and to execute. However it does only detect a small portion of potential bias. Detecting bias by searching for relationships or by looking into recommendations done just before the filling date of bankrupt firms is more comprehensive, due to the fact of quantifying the relationship. Furthermore basing bias on advices issued for companies who went bankrupt gives no guarantee that the investment bank is biased for other companies.

Therefore in Subsection 4.3.3, an indirect method is proposed for correcting bias. This method gives penalties for advisors whose advices are performing worse. In addition, for every stock, the underwriters who did the IPO are looked up and saved. Since IPO information is not available in most of the cases, this addition is therefore only applicable for companies for which IPO information can be looked up.

4.3.2 Direct Method to Correct Bias

As discussed in Section 4.3.1, there are a couple of methods for discovering potential bias in predictions. The most promising method of these three, is discovering bias by looking up which underwriters have done the Initial Public Offering (IPO) for the particular stock. This is because:

- IPO information is publicly available. It can be extracted from publicly available sources (Hoovers)

- Information, once obtained, is easy to quantify. One can experiment by including or excluding affiliated underwriters

- According to (Michaely and Womack, 1999), underwriters that have done the IPO for a particular stock, have the incentive to publish positive recommendations for that particular stock.

- The information obtained about potential biased underwriters in regards to particular stock is still valid. This information is not obsolete and can be used in the aggregation process.

One minor downside of this direct method for correcting bias is the scalability issue. The method can only be used for stocks on which the IPO information can be obtained.

There are several publicly available sources in which information, such as IPO information, is represented for a particular stock. One of these sources is Hoovers (Hoovers). This source is very comprehensive and in addition it provides free information for US stock listed firms from 1996 and onwards.

Since we want to test whether Initial Public Offering has influence on the performance of an advice issued, we use this source and therefore accept the compromise of not being able to apply this correction method for every stock listed in the ontology.

The bias correction method is implemented on top of the improved aggregation method discussed in Section 4.3.3. In this correction method, advisors who have done the IPO of a company are assigned with different weights, the Equation 4.21, must be adjusted in order to apply to this constraint. In Equation 4.8, the new Equation is displayed for calculating the Adjusted Weight per advisor.

$$AW_{at} = x * EF_{at} * SP_{at} \tag{4.8}$$

In Equation 4.8, the adjusted weight is calculated by taking the sum of the Experience factor of advisor a at time t and the Smoothed performance of advisor a at time t. The difference with the initial Equation (4.21), is the addition of the weight x. This weight is multiplied with the smoothed performance and the experience factor. The explanation of Equation 4.8 is given in Section 4.3.3.

4.3.3 Improved Aggregation Method

As discussed, for correction bias an additional method is developed on top of the improved aggregation method. This improved aggregation method has been built from scratch, but some parts of the original method are kept, although transformed or updated.

First of all, issuers publish advices, an advice holds for a period of six months or until a new advice is issued. As soon as one of these conditions is met, the performance of the advice over that period is calculated. This is done by calculating the *dailyReturn* for every day, as displayed in Equation 4.9 for buy advices and in Equation 4.10 for sell advices.

$$dailyReturn_{buy} = \frac{price_{t+1} - price_t}{price_t} \tag{4.9}$$

$$dailyReturn_{sell} = \frac{price_{t-1} - price_t}{price_t} \tag{4.10}$$

In Equation 4.9 and 4.10, $price_t$ stands for the price of the stock at a particular day in the period the advice holds. The Equation for calculating the hold advice performance is different in comparison to the others. In the original method (Stibbe, 2007) a hold advice is treated the same way as a buy advice. Which is incorrect, because they substantially differ, in case the price of a stock rises, the performance of a buy advise for that particular stock should also be rising, but the performance of a hold advice should not rising in the same way as the performance of a buy advice. Otherwise the underwriter should have issued a buy advice. Hold advices are issued when advisors do not predict any ups or downs in the particular stock. Therefore the performance of a hold advice should be positive when the future prices ranges within the range of the initial price. In order to accomplish this, the hold advice is an aggregate of the performance of a buy advice and the performance of a sell advice, as displayed in Equation 4.11 and 4.12.

$$dailyReturn_{hold} = dailyReturn_{sell} + dailyReturn_{buy} \tag{4.11}$$

$$dailyReturn_{hold} = \frac{price_{t+1} - price_t + price_{t-1} - price_t}{price_t} \tag{4.12}$$

In Equation 4.12, the performance of a hold advice is calculated by taking the sum of the calculation of a sell performance and a buy performance.

Once the performance over the period per day is calculated, the arithmetic mean $\mu_{dailyReturn}$ is taken over that period, as displayed in Equation 4.13.

$$AvgDailyReturn_a = \frac{\sum_{i=0}^{n} dailyReturn_i}{n} \tag{4.13}$$

In Equation 4.13, the *AvgDailyReturn* for advisor a is calculated by taking the sum over each *dailyreturn* and then it is divided by the amount of days in the period n. Next, the daily return of the advice is compared with the average performance of the industry d in the same period, displayed in Equation 4.14.

$$AvgIndustry_d = \frac{\sum_{j=0}^{d} \frac{price^j_{TIend} - price^j_{TIstart}}{price^j_{TIend}}}{d} \tag{4.14}$$

In Equation 4.14, the average performance of a particular industry d is calculated over the same period as the advice for which one wants to measure the performance. This is done by summing the percentage change in price over the period for every company j in the same industry d and taking the arithmetic mean over this value. Companies from the same industries are retrieved by looking up the GICS (Global Industry Classification Standard, see Glossary GICS). Next the distance between the the two performance measurements is calculated, as displayed in Equation 4.15.

$$x_t = \mu_{dailyReturn_a} - \mu_{industry_d} \tag{4.15}$$

In Equation 4.15, the distance is calculated by subtracting the performance of the industry d off the *dailyReturn* of Advisor a. Once the distance between the performance of the individual company in regards to the industry is calculated, an index is made of this value, as displayed in Equation 4.16.

$$x_{t_{index}} = x_t + 1; \tag{4.16}$$

As displayed in Equation 4.16, the index of the return is made by adding 1. If an advisors advice performs better than the average industry performance, the index has an value greater than 1. In case of equal performance the index is 1 and in case of worse performance the index is smaller than 1. This is because the performance x_t is always an value between 0 and 1 and thus the index performance $x_{t_{index}}$ is always an value between 0 and 2. The value is indexed to prevent negative values, thus leading to exclusion of the analyst. Next the index performance value $x_{t_{index}}$, is passed onto Brown's exponential smoothing function (single exponential smoothing). Since the time series data possibly does not contains any trend, single exponential smoothing is chosen over double exponential smoothing. The data is smoothed to give more weight to recent observations and have declining weights for older observations. This way, one gets smoothed data of daily returns of advisors.

Before the index value is smoothed, the weight of every advisor W_a is set to be 1. Which means one assumes at the beginning, every advisors performs equal to the industry of the advice issued, as given in Equation 4.17.

$$W_a^{t=0} = 1 \tag{4.17}$$

Once the weight per advisor is determined, the weight can be used to smooth the performance for each advisor as displayed in Equations 4.18 and 4.19.

$$W_a^t = \alpha x_{t_{index}}^a + (1 - \alpha)W_a^{t-1} \tag{4.18}$$

$$SP_a^{t+1} = \begin{cases} W_a^t & W_a^t > 0 \\ 0 \leq 0 & W_a^t \leq 0 \end{cases} \tag{4.19}$$

In Equation 4.18 and 4.19, the weight for an advisor a is adjusted by the latest observation of the performance of his advice $x_{t_{index}}$. The α variable needs to be determined, higher α values give more weight to recent observations, which is the purpose of exponential smoothing. The optimal value of α must be determined during experiments. The smoothed performance SP of advisor a at time t, is always between 0 and 2, and never exceed these limits.

By incorporating weights for the individual advisors, one advisor can get penalties in cases when his advice is performing badly. This way advisors, who perform over the last couple of periods worse than the industry, get less weight W_a^t in the eventual aggregation of advices, because their weight has dropped below the index value of 1. In the most extreme case, the weight of an advisor can drop below zero when the advisors only issue advices which perform badly (below the industry performance for a long period of time).

However, once the weight of the advisor is below 0, the weight is set to 0 and the advisor is not accounted for anymore. Because of the behavior of exponential smoothing, which gives more weight to recent observations, an advisor with a very low weight can quickly increase in weight once his advices are performing better than the industry performance.

This smoothed performance SP_a^t is updated with every new advice of the advisor. However smoothed performance does not say everything. There could be situations in which some advisors have a very high smoothed performance SP_a^t, solely because this value is based on just one advice. Therefore extending this model with the amount of advices issued is necessary. In this light, the *experience factor* of an advisor is introduced. The experience factor of an advisor expresses the amount of issued advices in the past.

As displayed in Figure 4.3, the slope of the experience factor of an advisor is decreasing over the amount of advices. In the initial stage the difference between an advisor with 10 advices issued and an advisor with only 2 advices issued is much bigger than an advisor with 50 advices in contrast to one with 40 advices. The experience factor formula is as displayed in Equation 4.20.

$$EF_a^t = ln(z + 1) \tag{4.20}$$

In Equation 4.20, z are all the advices issued by advisor a from the start of the experiment until the point in time the advice aggregation is requested.

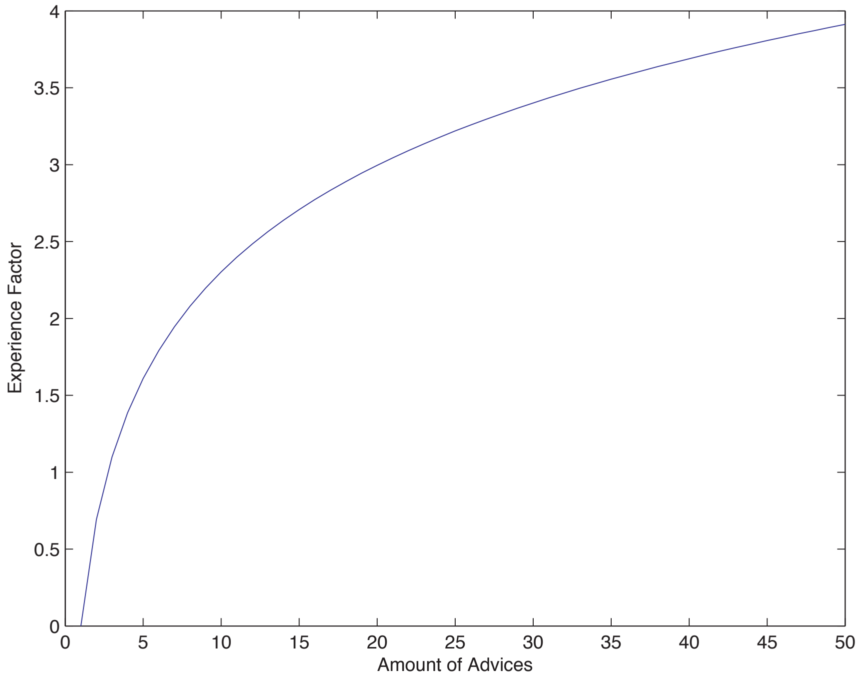

Figure 4.3: The graph of the experience factor equation.

The natural logarithmic function is used, because of the behavior of the logarithmic function, as displayed in 4.3, which is a very steep slope in the beginning and flattening when the amount of advices increases. Due to the fact the natural logarithm of one, $ln(1)$ is zero, every amount of advices is increased by one. In this case, the experience factor of an advisor without any advices is zero and thus not accounted for in the model. Once the experience factor and the smoothed performance of the advisor is calculated, the *adjusted weight* of the advisor can be calculated.

$$AW_a^t = EF_a^t * SP_a^t \tag{4.21}$$

In Equation 4.21, the adjusted weight AW for an advisor a at time t, is the product of the experience factor (Equation 4.20) and the smoothed performance (Equation 4.19) of the advisor. With the *adjusted weight* per advisor the average weight per advice type can be calculated.

$$AWAT^t_{[buy,hold,sell]} = \frac{\sum_{k=0}^{q} AW_{k_{[buy,hold,sell]}}}{q} \tag{4.22}$$

The *average weight per advice type*, Equation 4.22, is calculated per advice type. This is done by taking the sum over the all the weights per advice type and divide it by the amount of advices per advice type q. Next, the *confidence factor* per advice type is calculated.

$$CF^t_{[buy,hold,sell]} = \frac{AWAT^t_{[buy,hold,sell]}}{AWAT^t_{buy} + AWAT^t_{hold} + AWAT^t_{sell}} \tag{4.23}$$

In Equation 4.23, the confidence factor per advice type is based on the average weight per advice type. The biggest confidence factor is taken as the aggregated advice type.

For the advice application the strength of the given aggregated advice much be calculated. Although, as stated in (Stibbe, 2007), strength measurements do not hold any prediction power, it still gives some extra information about the strength of the newly generated advices for the end user.

$$Strength_{s_l} = max(CF_B, CF_H, CF_S) - median(CF_B, CF_H, CF_S) \tag{4.24}$$

In Equation 4.24, the strength of an aggregated advice is calculated by measuring the distance between the biggest confidence factor and the median confidence factor. This method is chosen over the original method outlined in Equation 4.7, due to the behavior of the later. Strength method 4.7, always incorporates three advices types, while it is very likely only two, or maybe only one, advice type is issued for a particular time window.

In Table 4.2, a comparison is made between the different strength methods (Equation 4.6, 4.7, 4.24) for different situations. In Table 4.2, B stands for buy, H stands for hold and S stands for sell. In the column B,H,S different confidence factors for buy, hold and sell are given. In the different strength columns the strength value and the chosen advice type is given. The chosen situations are mostly extreme cases.

Table 4.2: Comparison of different strength methods.

B,H,S	s_r	s_v	s_l
0.33,0.33,0.33	0.00H	0.00H	0.00H
0.49,0.51,0.00	0.49B	0.26H	0.02H
0.33,0.66,0.00	0.33B	0.49H	0.33H
0.20,0.80,0.00	0.20H	0.50H	0.60H
0.20,0.00,0.80	-0.60S	0.5S	0.60S
0.33,0.00,0.66	-0.33S	0.49S	0.33S
0.05,0.25,0.70	-0.65S	0.55S	0.45S

Based on the Table 4.2, one can state that strength method s_l is the best option. Besides giving more realistic strength figures (see second row), it is in addition also the method with the least drawbacks. Furthermore the method will only be used to give some extra information about the aggregated advice for the end user and cannot be used in the aggregation process, since strength values do not give any extra informational value to the aggregation process, as discussed in (Stibbe, 2007).

4.3.4 Comparison Between Aggregation Methods

In Section 4.2, the original method of the Market Advice Aggregation System is described. This method is schematically displayed in Figure 4.4. In Figure 4.5, the new designed aggregation method is displayed, as described and explained in Section 4.3.3. White blocks correspond to similar parts in the method, while colored blocks correspond to adjustments made in the original method.

Figure 4.4: Schematic view of the old aggregation method.

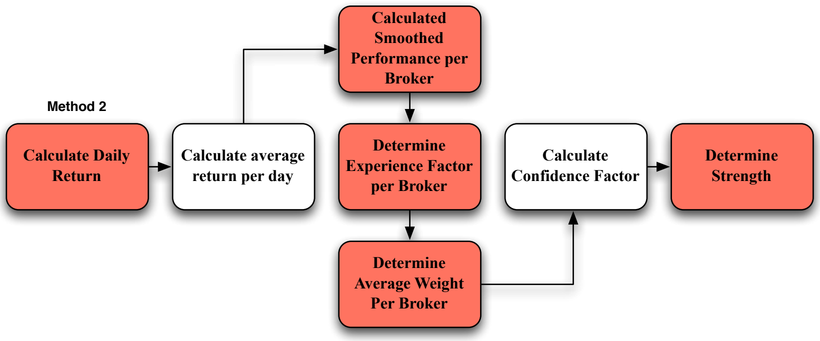

Figure 4.5: Schematic view of the new aggregation method.

As one can see in Figure 4.4 and 4.5, the main adjustments are made between calculating the average return per day and determining the confidence factor. In the original method, the broker performance is calculated by taking the average over all the average returns of advices issued by the broker. Furthermore only positive performance indexes are taking to measure the confidence factor. By incorporating only positive indexes, a lot of information is lost, which otherwise could add information value to the aggregation process (i.e. bad performing underwriters all state sell, while good performing underwriters mostly state hold, this could strengthen the new generated advice). Therefore in the new designed model an average weight per advisor is determined. This weight is based on the smoothed performance of the advisor and the experience of the advisor. This approach is chosen since smoothed performance on its own does not say enough about the advisor. Therefore by adding information about the experience of the advisor, one gets a balanced weight per advisor.

Furthermore the calculation of daily return has been adjusted. In the original method, there was no distinction made between a hold and a buy advice, and therefore were both treated the same. In the new method, both advice types are treated different. Buy advices performs well in case of an increase in future prices, hold advices do not rise in the same fashion as a buy advice in case of a price increase.

Calculating the confidence factor is in essence the same, only with different variables. However calculating the eventual strength of the new generated advice differs from the original method. While the original method always accounted for all three advice types issued, the new strength method only calculates the difference between the chosen advice type (advice type with the biggest confidence factor) and the median advice type. This way, one knows the distance between the advice type chosen and the median advice type.

4.4 Implementation of New Aggregation Method

The original MAAS application by (Stibbe, 2007) is designed in Java. Due to the fact the application is completely Object Oriented, new methods can be integrated quite easily.

As discussed in Section 3.2, the application consists of three parts, namely the advice application, extra gear module and the information extractor.

- Advice Application: The main part of the application. A GUI is used by users to specify the company for which it wants an aggregated advice. By specifying the date, the user gets an aggregated advice.
- Information Extractor: Used to populate the ontology with new advices from the Analist source (Section 1.4). This part of the application has been modified in order to correctly extract information from the sources. Due to the fact the advices are from a different source and foreign, this module had to be reconstructed.
- Extra Gear: Mainly used to run experiments on large amount of data, where there is no GUI available in order to save system resources.

4.4.1 Implementation of Bias Correction Method

In Section 4.3.2, the method for correcting bias by incorporating IPO information is described. In order to get this particular information, one must first extract the relevant information from the Hoovers source on the internet (Hoovers).

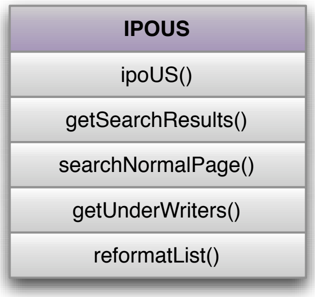

Figure 4.6: UML class diagram of the IPO class.

In order to do so, an Extract class is added to the Extra Gear Module, as displayed in Figure 4.6. The class contains two search methods, the first method searches the database by Quote Symbol code, while the second method search the whole database by using the normal name. This way exceptions are handled, once the quote symbol cannot be looked up or when there are multiple search results for a quote symbol. Next the underwriters are extracted, once the proper entry is found. Eventually the retrieved underwriters are stripped from long names (i.e. Merill Lynch is transformed to Merill), so they are easy comparable with advisors stored in the ontology.

4.4.2 Implementation of Improved Aggregation Method

As discussed in the introduction of this Section 4.4, the application consists of three modules. The improved aggregation method must be first of all implemented in the advice module, so it is usable through the user interface. In addition, the extra gear module is used to run experiments and therefore this module must also be adjusted in order to work with the improved aggregation method. Due to the fact the original application is completely object oriented, implementing the new method is done by adding two new classes. The original classes which are affected by the new method are displayed in Figure 4.7. The *NewAdvice* class is depended on the *Advisor* class.

In Figure 4.8, the newly designed classes are displayed, in addition the changes are displayed in bold. The Advisor class contains some new set and get methods, in which the logic is done, as described in Section 4.3.3. Furthermore, the NewAdvice class is adjusted to properly use the new set and get methods in the Advisor class. In addition, a strength method and a method to get the biggest confidence factor is added. In the original application, calculating the strength of an advice and determining the biggest confidence factor, was done in the GUI class. This is against the MVC

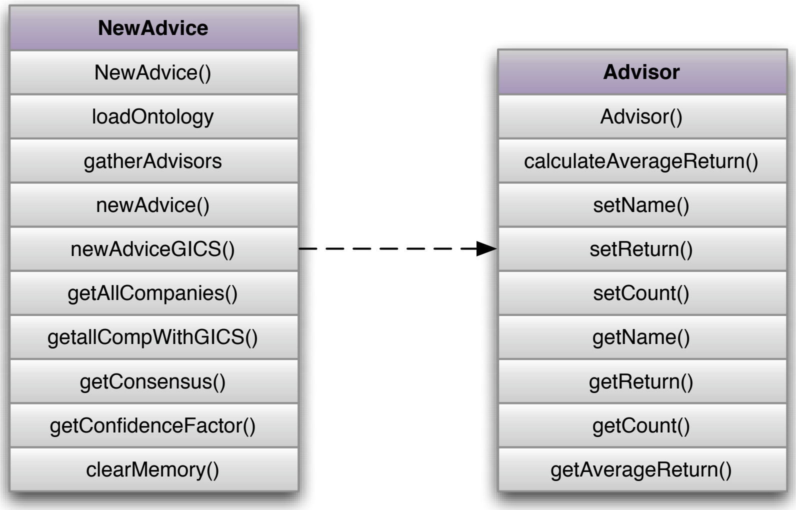

Figure 4.7: UML diagram of the original classes, whom are affected by the new method.

(model, view, controller) approach and therefore this logic is transferred to the NewAdvice class.

In order to run experiments with the new classes, the LargeRun class in the Extra Gear module, must only import these new classes. The output of experiments is saved into CSV (comma separated values) files, in order to process the data in statistical programs.

4.5 Conclusion

In Chapter 3, the methodology used in the Market Advice Aggregation System is explained in detail. In this chapter the updated method is proposed. The original method aggregates the advices to one advice, by first determining the daily return per advice issued. Next, per broker the performance of this broker is determined by taking the average of all the daily returns of advices issued by this particular broker. Finally the confidence factor is calculated for all the brokers who have issued an advice for the requested stock in the requested period. The biggest confidence factor is chosen as the aggregated advice.

In order to improve this method, different approaches are discussed, such as; searching for relationships between stock listed companies and analysts, looking up IPO information and analyze historic data about bankruptcy and the behavior of advisors just before the bankruptcy of stock listed companies. Although each approaches knows some disadvantages, the new improved

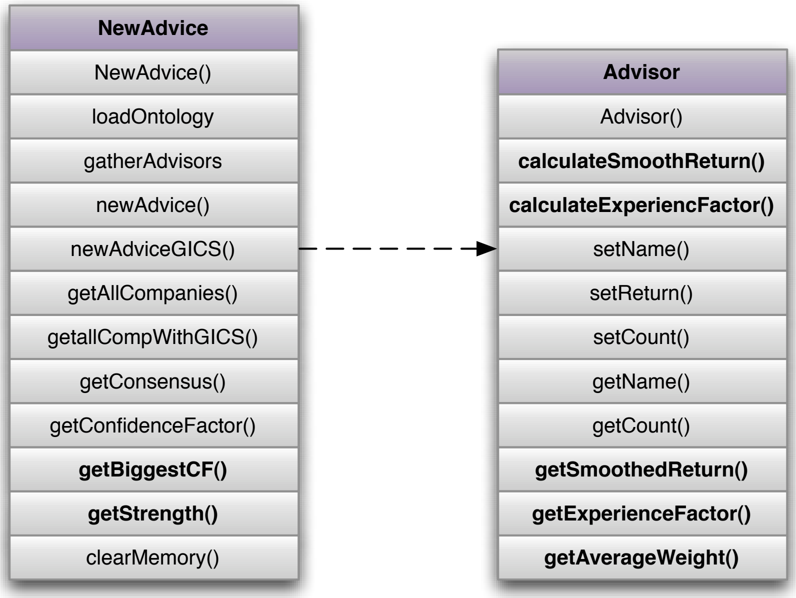

Figure 4.8: UML diagram of the newly coded classes.

aggregation method is extended with Initial Public Offering data in order to correct for bias. In the experimental phase one can give different weights to the potential affiliated analysts (0 till 1) and eventual omit them from aggregation to analyze the impact of possible affiliation.

For each analyst the daily return per advice is calculated, although a distinction is made between hold and buy advices, while the original method treated them the same. These average daily returns are added to a smoothing performance function, which smoothes the performance of the analyst every time a new observation (daily return of a new advice) is added. In addition, the experience factor is calculated per advisor, which is based on the amount of advices. The product of smoothed performance and experience factor gives a balanced weight to each advisor.

Next the product of the smoothed performance and the experience factor is taken, which is the eventual weight for the advisor. In the aggregation process, each advice type is multiplied with the average weight of the advisors who issued the particular advice type. The advice type with the biggest weight is chosen as the aggregated advice.

Chapter 5

Experiments and Results

5.1 Introduction

Once the newly designed method has been integrated, as described in Section 4.4, experiments must be executed in order to determine the performance of the new method. The experiments can be divided into two groups, namely: old aggregation method versus improved aggregation method and correction for affiliated underwriters versus no correction for affiliated underwriters. In regards to these experiments the average return per company per period (one day, one week, one month and half year) is calculated and compared.

The experiments are executed over a period of 48 months for 42 selected companies. These companies are selected based on the amount of information can be found in regards to the Initial Public Offering.

5.2 Experimental Setup

The original experimental setup relies on advices from the I/B/E/S data set, since Analist.nl (formerly known as marketadvices) only covers a too narrow time window. Due to the original experiment is executed over a period of 48 months between 2003 and 2006, most of the advices must be extracted from the I/B/E/S data set, as Analist.nl does not cover that period to the full extend, as displayed in Figure 5.1.

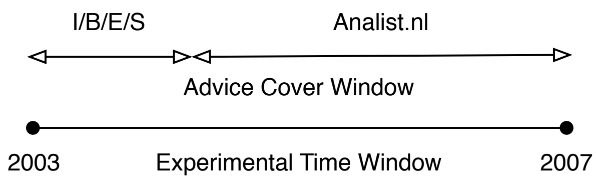

Figure 5.1: Sources coverage of experimental time window.

Information from this data set is extracted and added to the ontology. In order to make the best comparison between the two methods, the same data is used. Therefore extracting new advices is not necessary.

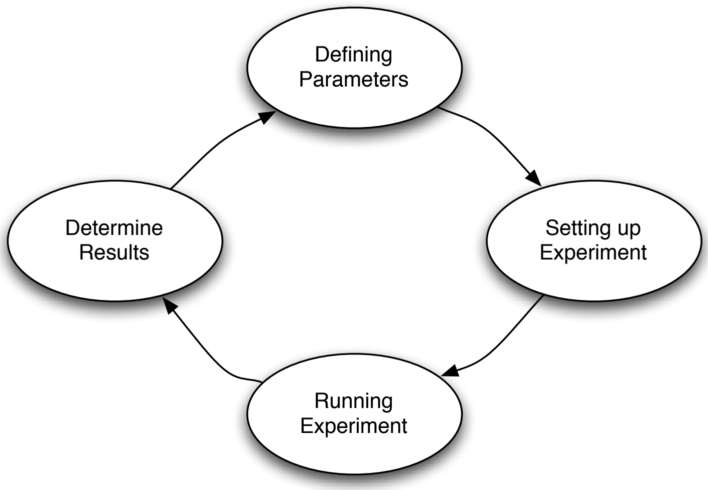

Figure 5.2: Experimental setup.

In Figure 5.2, the experimental setup is schematically displayed. It is a loop in which parameters are adjusted after every experiment in order to achieve optimal results and solution. For the first experiment (old aggregation method versus improved aggregation method) the experimental setup is as follows:

- Setting parameters: In the newly developed method, some parameters can be adjusted. Most important parameter is the weight given to latest observations.

- Setting up experiments: The experimental setup must correspond with the original setup. This make comparison more accurate.

- Running experiments: Experiments will be executed using the Extra Gear module of the application. In this module a special command line class is developed for running large experiments

- Interpretation of results: Studying the outcome of the experiment and possibly define new values for the parameters. In addition, tests of significance are conducted to be able to answer the research question.

For the second experiment, correction for affiliation versus no correct for affiliation, the experimental setup is as follows:

- Setting parameters: The weight given to affiliated underwriters can differ. The weight per underwriters will be tested between zero weight and normal (1) weight.
- Setting up experiments: The experimental setup must correspond with the first setup in regards to the time window. Secondly the companies must be selected based on the amount of information available.
- Running experiments: Experiments will be executed using the Extra Gear module of the application. In this module a special command line class is developed for running large experiments
- Interpretation of results: Studying the outcome of the experiment and possibly define new values for the parameters. In addition, tests of significance are conducted to be able to answer the research question.

The prices of the stocks at particular time intervals are extracted from the Yahoo Finance website (prices are always closing day prices). Due to this fact a lot of requests are made on the Yahoo Finance server, therefore (Stibbe, 2007) decided to persistently store all the relevant price information, as soon as it is requested. This way there is no need to send out multiple request to the Yahoo server for getting price data.

5.3 Determine Parameter Values

Multiple experiments will be executed to determine optimal values for both the experiments.

5.3.1 Overview of Parameters

In the newly designed method (Section 4.3.3), a couple of parameter values must be chosen. The most important parameters for the method are:

- Weight to observations: The new method contains a smoothing function, which smoothes the average daily returns of advisors. With this method one can give variable weight to recent observations in contrast to the accumulated older recommendations. This weight is denoted by a in Equation 4.18.
- Weight given to affiliated underwriters: In order to measure the impact of bias in recommendations, potential affiliated underwriters are assigned variable weights, while unaffiliated underwriters are given normal weight. The affiliation of an underwriter is based on whether or not the underwriter had done the Initial Public offering for the stock listed company. This weight is denoted by x in Equation 4.8.

The best value for these two parameters must first be found in order to run the experiments.

5.3.2 Weight to Observations

In order to find the optimal value for the α value, an experiment is executed 10 times, for 10 different α values. For each experiment the α value is incremented with 0.1. This gives the following output, as displayed in Figure 5.3. This figure is the return per α value for Zions Bancorporation.

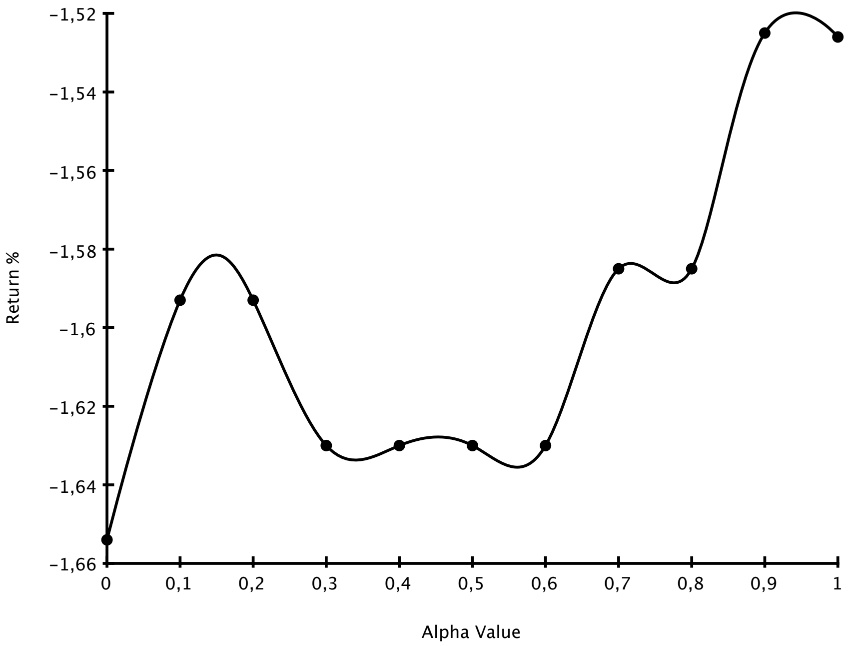

Figure 5.3: Average return per week/month/half year for different α values.

As one can see in Figure 5.3, the optimal value for the α value for Zions Bancorporation is about 0.9. Since this value eventually gives an average return of -1.525. However, this process of determining the optimal α value per company is quite time consuming, since per company the experiment must be executed 10 times in order to determine for 10 α values the return.

In order to overcome this problem, the optimal α value is determined individually per company. This is done by calculating for each company the optimal α value for the first month of the 48 month period. The optimal alpha value is chosen and used for the rest of the experiments for the selected company. This brings down the total time of running experiments.

5.3.3 Weight to IPO Underwriters

Finally the weight for IPO advisors must be determined. As described in Section 2, underwriters who have done the due diligence process for a particular stock, should have superior information regarding the stock and thus their recommendations should potentially perform better than average. According to (Michaely and Womack, 1999), the contrary is true and underwriters have an incentive to issue positive advices, regardless whether or not this advice is in best interest for investors.

To investigate whether this conclusion is valid and thus whether there is bias in recommendations both the experiments will be executed. Once with all the IPO underwriters included and once with all the IPO underwriters excluded (for the particular stock). In order to exclude and include IPO underwriters, Equation 4.8 is used. The value for x is adjusted in order to include and exclude the IPO underwriters (x takes the value 0 and 1).

5.4 Results

In this section the results will be presented and described. The parameters are set as described in Section 5.3.

5.4.1 Performance of IPO Underwriter Recommendations

As described in Section 5.3.1, experiments are executed twice for stock listed companies for which IPO information is available. The experiment is executed twice, once with all the underwriters included in the aggregation process and once without the underwriters who have done the IPO for the particular stock. According to (Michaely and Womack, 1999), leaving out the underwriters who are potential biased (underwriters who have done the IPO for the particular stock), should lead to better aggregation, since the underwriters having done the due diligence process are prone to give more positive recommendations regardless of the quality of the recommendation.

To test this finding, the experiment will be executed with and without the affiliated underwriters, as displayed in Table 5.1. The results in Table 5.1 display only a small portion of the companies tested. The results are an average over the 48 months of experimental time window.

In Table 5.1, the performance per week, per month and per half a year is calculated for a sample of the 42 companies. Each experiment is executed two times, one with and one without affiliated underwriters (IPO underwriters included and IPO underwriters excluded). These potential affiliated analysts are looked up as described in Section 4.4.1. In Table 5.1, the Excl row is without the potential affiliated analysts, while the Incl row is with all the underwriters included.

In Figures (5.4, 5.5, 5.6 and 5.7), the one day, one week, one month and half year return for 42 tested companies is displayed. The blue bar chart (lighter color) is the aggregated performance without the IPO underwriters, while the red bar chart (darker color) is with all the advisors included. As one can see, on the one week return, there are some peeks in performance when excluding IPO underwriters, while on the longer run (one month and

Table 5.1: Comparison of performance of IPO underwriters.

Company	IPO	Day Rtrn	Week Rtrn	Month Rtrn	$\frac{1}{2}$ Year Rtrn
Yahoo	Excl	0,375	-1.3786	-2.2857	-10.8562
	Incl	0,4166	-1.5812	-1.7822	-9.8186
WellPoint	Excl	0,1666	-0.4158	-1.8460	-10.2263
	Incl	0,2708	-0.4692	-1.6895	-9.8564
UPS	Excl	0,7501	0.2079	-0.6228	-1.7381
	Incl	0,625	0.1468	-0,1692	-1.6767
Polo RL	Excl	0,8291	-0.3655	-2.0596	-15.5313
	Incl	0,0	-0.5130	-2.1665	-10.9786
NOV	Excl	0,4375	-2,5069	-0,3560	-11,6970
	Incl	0,4583	-2,5522	-0,7370	-11,6280
Google	Excl	0,1916	-0.6481	-1.7163	-10.7317
	Incl	0,2291	-0.7859	-2.0003	-11.9759
Amazon	Excl	0,1041	-0.3276	-1.4045	-4.6551
	Incl	0,1666	-0.3276	-1.4045	-4.6551

half a year), there is no significant difference anymore, when included or excluded. As one can see the lines on the longer run are overlapping each other, which means the performance of both methods does not differ significantly.

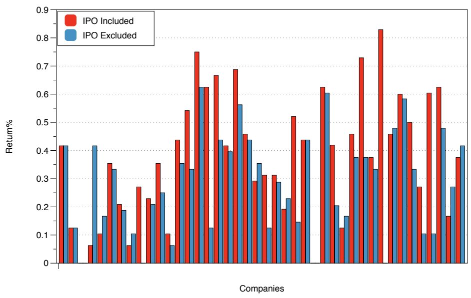

Figure 5.4: One day return for the 42 selected companies

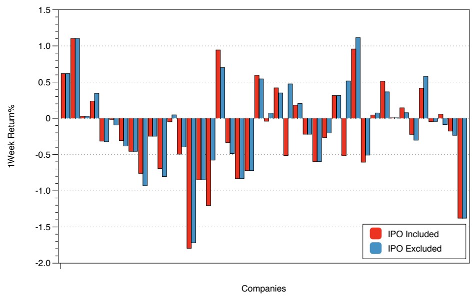

Figure 5.5: One week return for the 42 selected companies

Since no clear conclusion can be deducted from the results displayed in Figures 5.4, 5.5, 5.6 and 5.7, the distribution of the returns is displayed in Figures 5.8, 5.9, 5.10 and 5.11. In these figures the blue chart (lighter color) represents the aggregation with all the underwriters included, the gray chart (darker color) represents the aggregation without the IPO underwriters included. Next the significance of the returns must calculated.

In order to test whether or not the returns of advice aggregation without IPO underwriters yield higher returns in contrast to advice aggregation with all underwriters, a test of significance is conducted.

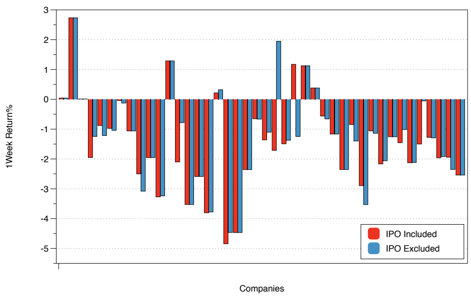

Figure 5.6: One month return for the 42 selected companies

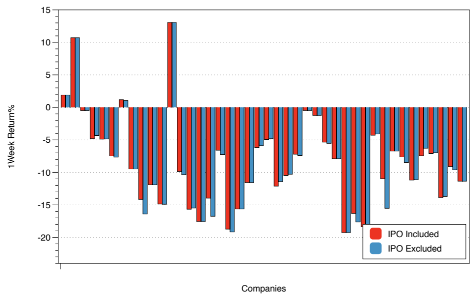

Figure 5.7: Half year return for the 42 selected companies

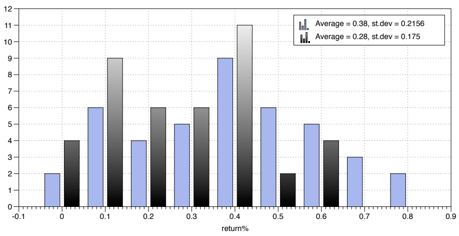

Figure 5.8: One day return distribution for the 42 selected companies

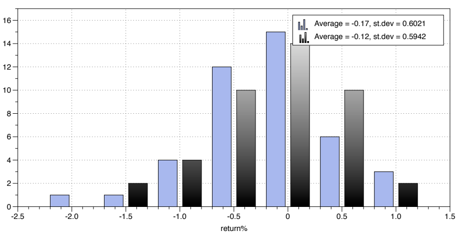

Figure 5.9: One week return distribution for the 42 selected companies

The Wilcoxon signed-rank test is used (Aczel and Sounderpandian, 2002), since it is useful in comparing two populations in which one has paired observation. In addition this test is useful in cases whether it is not clear if the differences between the paired observations are normally distributed.

Since the results per company are paired observations, nonparametric tests are the most useful to use (Aczel and Sounderpandian, 2002). Except the Wilcoxon signed-rank test, once can use the normal sign test. However the sign test does not account for the magnitude of differences between the paired observation, therefore the Wilcoxon test is more applicable.

First the null hypothesis and alternative hypothesis are defined, as displayed below:

- $H_0 : \mu = \mu_0$
- $H_a : \mu > \mu_0$

The null hypothesis states that there is no difference in mean return between including all underwriters μ_0 and excluding the affiliated underwriters μ. The alternative hypothesis states that the mean return of excluding affiliated underwriter μ is greater than including all underwriters μ_0. Once the null hypothesis can be rejected, the alternative is accepted, which lead

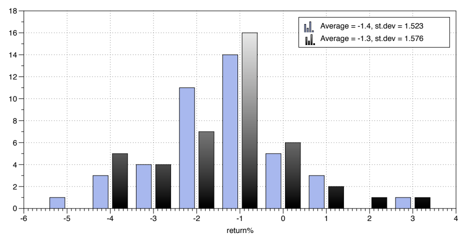

Figure 5.10: One month return distribution for the 42 selected companies

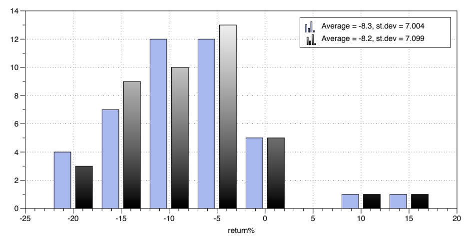

Figure 5.11: Half year return distribution for the 42 selected companies

to the conclusion that affiliated underwriters do not add information value to the aggregation process.

For each pair of observation i out of the total amount of observations $0, .., n$, the difference is calculated, as displayed in Equation 5.1.

$$D_i = x_i - \mu_0 \tag{5.1}$$

In Equation 5.1, x_i stands for a particular return without affiliated underwriters, μ_0 stands for average return with all underwriters included. The outcome D_i, describes the difference between both variables.

The next step is to rank the absolute values of the differences calculated in Equation 5.1, which is displayed in Equation 5.2

$$R_i = |D_i| \tag{5.2}$$

In Equation 5.2, the absolute value is taken per difference D_i and is assigned a rank R_i relative to the other differences.

Next, sums are are taken of the ranks of the positive and of the negative differences, which is called the Wilcoxon T statistics. This process is displayed in Equations 5.3 and 5.4.

$$T^+ = \Sigma_i(positiveR_i) \tag{5.3}$$

$$T^- = \Sigma_i(negativeR_i) \tag{5.4}$$

In Equation 5.3, the positive signed ranks R_i are summed up (ranks R_i in which D_i was already positive). This lead to the variable T^+. In Equation 5.4, the negative signed ranks R_i are summed up (ranks R_i in which D_i was negative).

If the amount of pairs n are below 50, the critical point w_a of the distribution of the test statistic T for level of significance a can be easily looked up in tables. Since the hypotheses stated earlier are right tailed tests, the critical value must first be converted to a right tail value, as displayed in Equation 5.5

$$w_{1-a} = \frac{n(n+1)}{2} - w_a \tag{5.5}$$

In Equation 5.5, the critical value w_{1-a} for significance level a for the right tail is calculated, according to (Aczel and Sounderpandian, 2002). This is done by multiplying the amount of observation with its increment $n(n+1)$ and divide it by 2, next the critical value for n observations with a level of significance a is subtracted. Once the positive Wilcoxon T^+ statistic is greater than the critical value of the right tail, the null hypothesis is rejected as displayed in Equation 5.6.

$$T^+ > w_{1-a} \tag{5.6}$$

Without looking up critical values, one can convert the Wilcoxon T statistic to a standarized z statistic as described in (Aczel and Sounderpandian, 2002). This standardization process is shown in Equations 5.7, 5.8, 5.9, 5.10.

$$T = min(T_+, T_-) \tag{5.7}$$

To convert to a standardized z statistic, the smallest signed rank is chosen as displayed in Equation 5.7.

$$E(T) = \frac{n(n+1)}{4} \tag{5.8}$$

In Equation 5.8, the estimated mean value for the T statistic is calculated, in which n is the amount of paired observations.

$$\sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{24}} \tag{5.9}$$

Next, the standard deviation for the T statistic is calculated as displayed in Equation 5.9.

$$z = \frac{T - E(T)}{\sigma_T} \tag{5.10}$$

Once the estimated mean and standard deviation are calculated for the T statistic, the z value can be calculated as displayed in Equation 5.10.

The level of significance α is set to 5%, which means the probability of making an Type I Error, rejecting the null hypothesis while the null hypothesis is true, is set to 5 percent. In Table 5.2, the null hypothesis is tested for the one day, one week, one month and half year return.

Finally the p value is calculated. The p value is a probability, calculated using the z test statistic, that measures the support provided by the sample for the null hypothesis (Aczel and Sounderpandian, 2002). In case the p value is below the level of significance of 5% the null hypothesis is rejected.

Table 5.2: Significance test of IPO advisors.

	1 Day Rtrn	1 Week Rtrn	1 Month Rtrn	$\frac{1}{2}$ Year Rtrn
μ_0	0,38	-0,17	-1,4	-8,3
$E(T)$	451,5	451,5	451,5	451,5
σ_T	79,97	79,97	79,97	79,97
z	2,7445	-0,7439	0,3188	-0,6938
p	0,9969	0,2284	0,3749	0,2438

As displayed in Table 5.2, all the p values are higher then the significance value of 0.05. Therefore in none of the cases the null hypothesis can be rejected.

Globally speaking, leaving out the IPO underwriters yields significantly no improved results. Therefore the companies are divided into groups accordingly to their market capitalization (Glossary Market Capitalization), to see if there can be improvements discovered per category. This leads to three subgroups, namely companies with small market capital (Glossary Small-cap), companies with medium capital (Glossary Mid-cap) and big companies with large capital (Glossary Large-cap). The division of used companies is displayed in Table 5.3.

Table 5.3: Division of companies in capital.

Small-cap	Mid-cap	Large-cap
8	21	11

Performance of IPO Underwriters at Small-Cap Companies

As described in the Glossary (entries: Small-cap, Mid-cap and Large-cap), companies market capital determines the group which they belong to. The market capital information is extracted of the Yahoo! Finance website.

In Figures 5.12, 5.13, 5.14, 5.15, the performance of the 8 small-cap companies is displayed. The blue bar chart (lighter color) represents the aggregated performance without the IPO underwriters and the red bar chart (darker color) with every underwriter

As one can see in the Figures (5.12, 5.13, 5.14 and 5.15), the monthly performance is slightly increased when the IPO underwriters are left out. However, on average base, including all the underwriters yield slightly better results for small-cap companies.

In Figures 5.16, 5.17, 5.18 and 5.19, the distribution of returns for small-cap sector is displayed. The blue bar represents the aggregation with all the underwriters included, while the gray one is without the IPO underwriters.

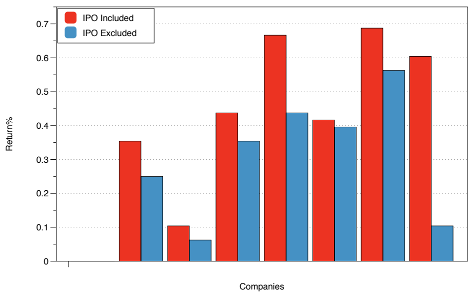

Figure 5.12: One day return for the Small-cap companies

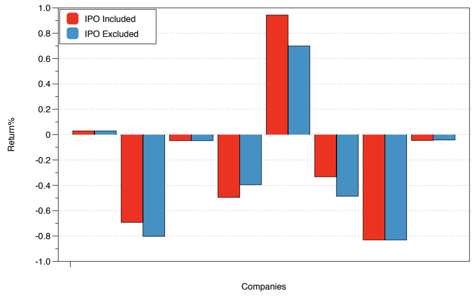

Figure 5.13: One week return for the Small-cap companies

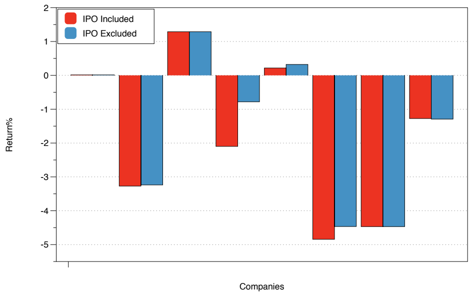

Figure 5.14: One month return for the Small-cap companies

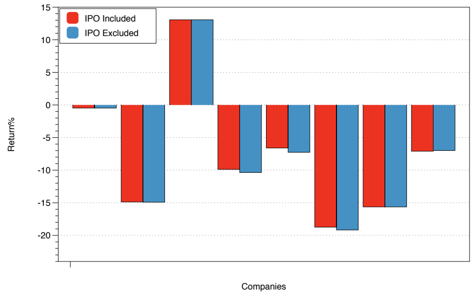

Figure 5.15: Half year return for the Small-cap companies

In order to test whether or not leaving out IPO underwriters yield better returns, the same null and alternative hypothesis are defined as before. This lead to the results displayed in Table 5.4.

Table 5.4: Significance test of IPO advisors for the small-cap sector.

	1 Day Rtrn	1 Week Rtrn	1 Month Rtrn	$\frac{1}{2}$ Year Rtrn
μ_0	0,41	-0,18	-1,8	-7,5
$E(T)$	18	18	18	18
σ_T	7,14	7,14	7,14	7,14
z	-1,6803	-0,5601	-0,28	-0,42
p	0,9464	0,2877	0,3897	0,3372

As one can see in Table 5.4, all the p values are bigger then the chosen level of significance of 0.05. This means the null hypothesis can not be rejected.

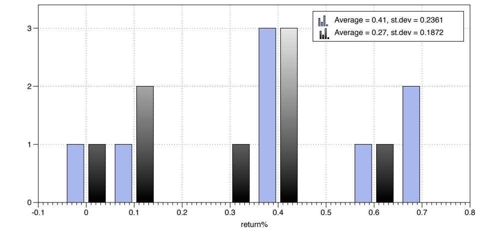

Figure 5.16: One day return distribution for the Small-cap companies

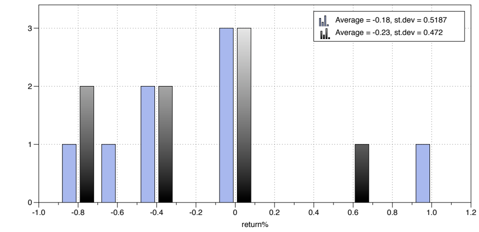

Figure 5.17: One week return distribution for the Small-cap companies

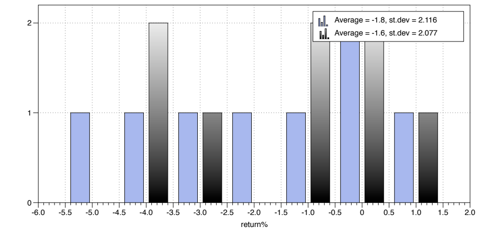

Figure 5.18: One month return distribution for the Small-cap companies

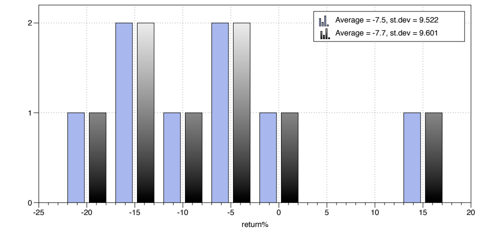

Figure 5.19: Half year return distribution for the Small-cap companies

Performance of IPO Underwriters at Mid-Cap Companies

In Figures 5.20, 5.21, 5.22 and 5.23, the aggregated performance of one day, one week, one month and half a year of mid-cap companies is displayed. The blue bar chart represents the aggregated performance without the IPO underwriters, while the red bar chart represents the aggregated performance of all the advisors.

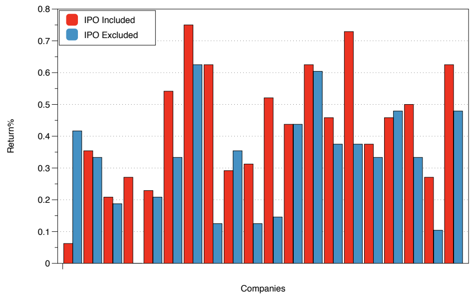

Figure 5.20: One day return for the Mid-cap companies

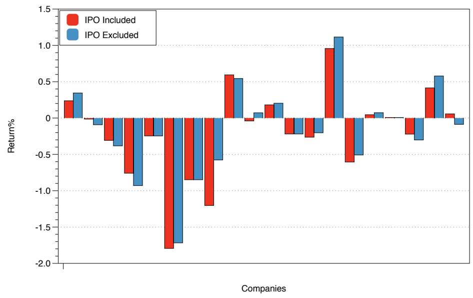

Figure 5.21: One week return for the Mid-cap companies

As one can see in the Figures (5.20, 5.21, 5.22 and 5.23), including the IPO underwriters yields in all cases better returns. Thus, in respect to mid-cap companies, IPO underwriters add additional value and information in

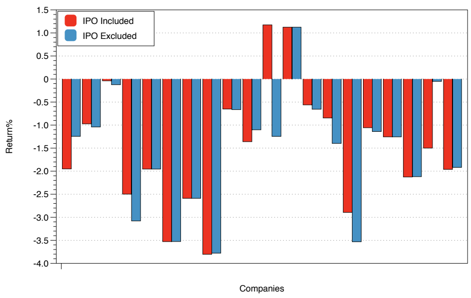

Figure 5.22: One month return for the Mid-cap companies

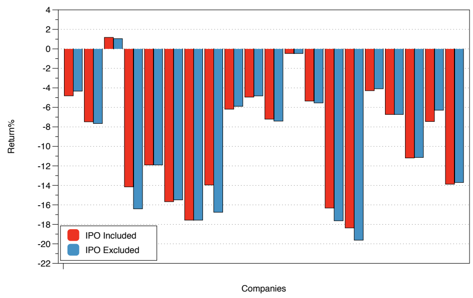

Figure 5.23: Half year return for the Mid-cap companies

the aggregation process. To test whether this statement is true, the same test of significance is conducted. The distributions of returns are displayed in Figures 5.24, 5.25, 5.26 and 5.25

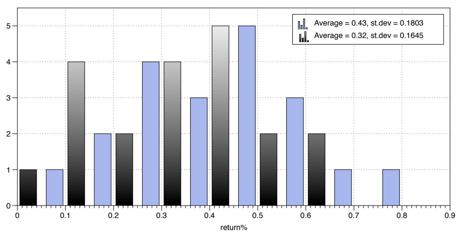

Figure 5.24: One day return distribution for the Mid-cap companies

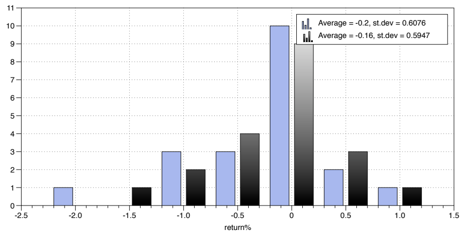

Figure 5.25: One week return distribution for the Mid-cap companies

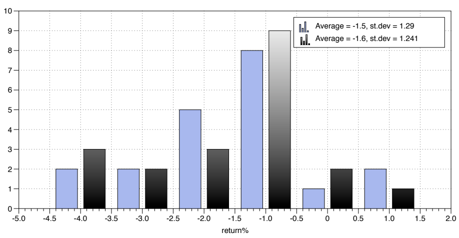

Figure 5.26: One month return distribution for the Mid-cap companies

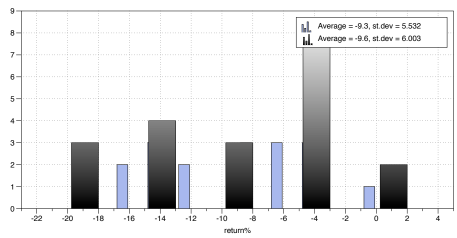

Figure 5.27: Half year return distribution for the Mid-cap companies

In order to test whether or not leaving out IPO underwriters yield better returns in respect to mid-cap companies, the same null and alternative hypothesis are defined as before. This lead to the results displayed in Table 5.4.

As one can see in Table 5.5, all the p values are bigger then the chosen level of significance of 0.05. This means the null hypothesis can not be

Table 5.5: Significance test of IPO advisors for the mid-cap sector.

	1 Day Rtrn	1 Week Rtrn	1 Month Rtrn	$\frac{1}{2}$ Year Rtrn
μ_0	0,43	-0,2	-1,5	-9,3
$E(T)$	105	105	105	105
σ_T	26,78	26,78	26,78	26,78
z	-2,7252	-0,5226	-0,1119	-0,1866
p	0,993	0,3006	0,4554	0,4259

rejected.

Performance of IPO Underwriters at Large-Cap Companies

In Figures 5.28, 5.29, 5.30 and 5.31, the aggregated performance of one week, one month and half a year of large-cap companies is displayed. The blue bar chart (lighter color) represents the aggregated performance without the IPO underwriters, while the red bar chart (darker color) represents all the underwriters.

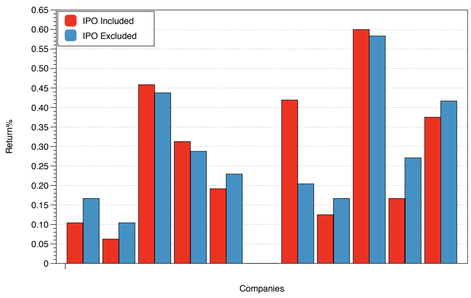

Figure 5.28: One day return for the Large-cap companies

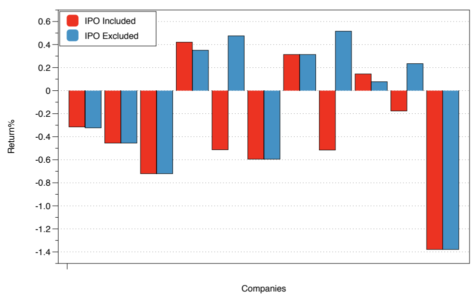

Figure 5.29: One week return for the Large-cap companies

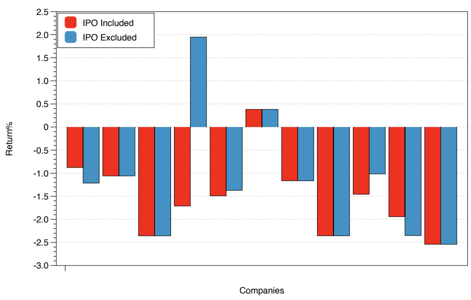

Figure 5.30: One month return for the Large-cap companies

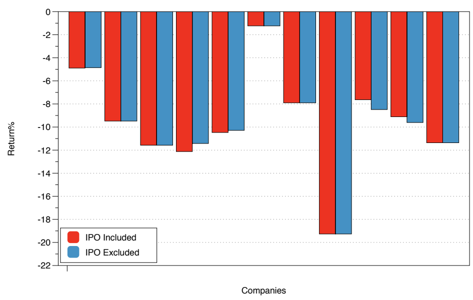

Figure 5.31: Half year return for the Large-cap companies

As one can see in the Figures 5.28, 5.29, 5.30 and 5.31, excluding IPO underwriters yields in the majority of cases higher return, in some cases the difference is very big. Especially on the short to medium run, the returns are higher when excluding the IPO underwriters. On the very short run (one day performance), for more than the half of the companies, excluding the IPO underwriters yield higher returns. To test this assumption, a test of significance is conducted. In Figures 5.32, 5.33, 5.34 and 5.35, the distribution of returns is displayed.

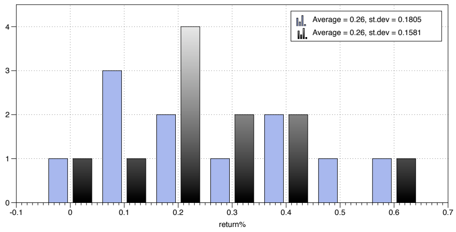

Figure 5.32: One day return distribution for the Large-cap companies

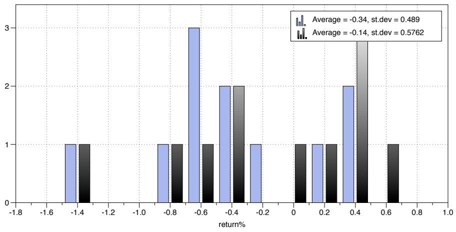

Figure 5.33: One week return distribution for the Large-cap companies

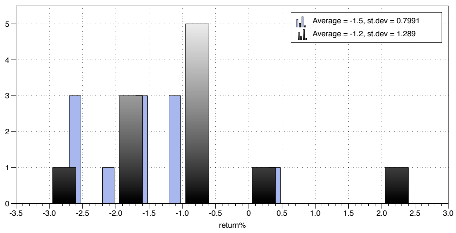

Figure 5.34: One month return distribution for the Large-cap companies

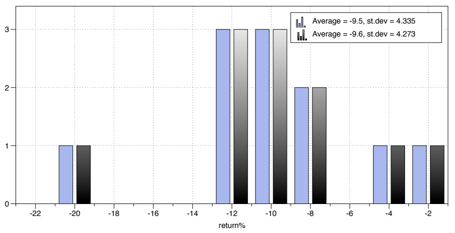

Figure 5.35: Half year return distribution for the Large-cap companies

In order to test whether or not leaving out IPO underwriters yields better returns in respect to large-cap companies, the same null and alternative hypothesis are defined as before. This leads to the results displayed in Table 5.6.

Table 5.6: Significance test of IPO advisors for the large-cap sector.

	1 Day Rtrn	1 Week Rtrn	1 Month Rtrn	$\frac{1}{2}$ Year Rtrn
μ_0	0,26	-0,34	-1,5	-9,5
$E(T)$	33	33	33	33
σ_T	11,24	11,24	11,24	11,24
z	-0,1778	-1,1558	-0,2667	-0,3556
p	0,4294	0,1238	0,3948	0,3610

As one can see in Table 5.5, all the p values are bigger then the chosen level of significance of 0.05. This means the null hypothesis can not be rejected.

Conclusion

In the experiment conducted above, the performance of the IPO underwriters is measured. This is done by including the underwriters and excluding the underwriters from the total aggregation process. As displayed in Tables 5.2, 5.4, 5.5 and 5.6, for none of the experiments their is a significance positive difference for excluding IPO underwriters. This lead to the conclusion that including them yield better results and thus should be included in the total aggregation process.

5.4.2 Performance of Improved Aggregation Method

This section describes the performance of the improved aggregation method, as described in Section 4.3.3. The results are compared with the performance of the original method as described in Section 4.2. For consistency the methods are tested against the same selected companies in Section 5.4.1.

Improved Aggregation Method on the 42 IPO Companies

As discussed in Section 4.2, the original aggregation methods has some drawbacks. Therefore a new improved aggregation method has been proposed in Section 4.3.3. In this Section, the performance of both aggregation methods are compared on the 42 companies selected in Section 5.4.1. In Table 5.8, a small portion of the selected companies is displayed and for each company the one day, one week, one month and half year performance is displayed. As one can see, the improved version performs significantly better then the original method. Only for Google Inc., the original method performs better.

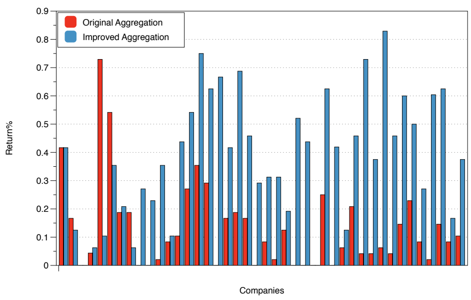

Figure 5.36: One day return for the 42 selected companies

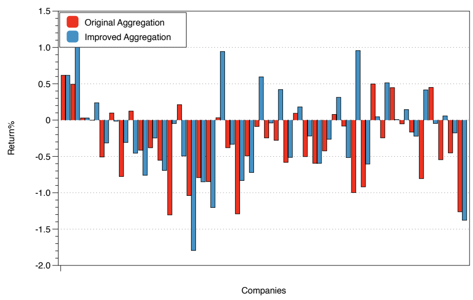

Figure 5.37: One week return for the 42 selected companies

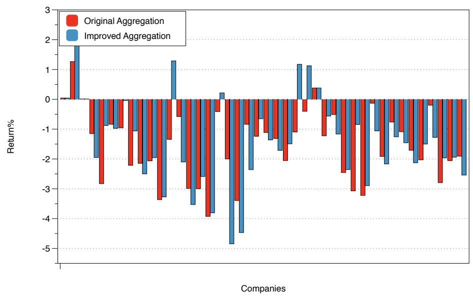

Figure 5.38: One month return for the 42 selected companies

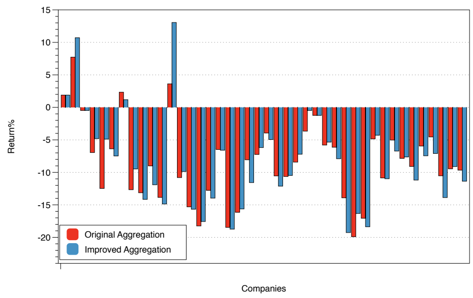

Figure 5.39: Half year return for the 42 selected companies

In Figures 5.36 5.37, 5.38 and 5.39, the one day, one week, one month and half year aggregated performance of the original and improved method are displayed. The blue bar (lighter color) represents the new improved aggregation method, while the red bar (darker color) represents the original aggregation method. On the very short run there is a lot difference between both methods, as displayed in Figure 5.36. As one can see, the improved method outperforms the original method in most of the cases.

On the medium run, one week, the improved aggregation method performs significantly better than the original method for most of the selected companies. This can be explained by the fact that the original aggregation method yields in more cases a buy advice, which performs more poorly in contrast to the improved method for which aggregations yield a hold advice. In some cases the difference between the original method and the improved method is quite extreme.

On the longer run, the one month performance, the improved method is still performing slightly better although the difference is less clear in contrast to the one week performance. On the long run, half year, the effect is minimized and there is no significant difference anymore.

In Figures 5.40, 5.41, 5.42 and 5.43, the distribution of returns is displayed. The blue/purple (lighter color) chart represents the original aggregation method, while the gray chart (darker color) represents the improved aggregation method.

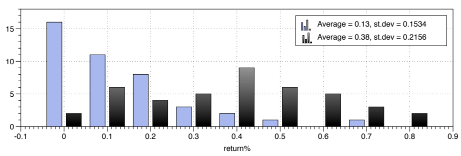

Figure 5.40: One day return distribution for the 42 selected companies

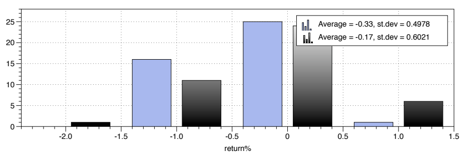

Figure 5.41: One week return distribution for the 42 selected companies

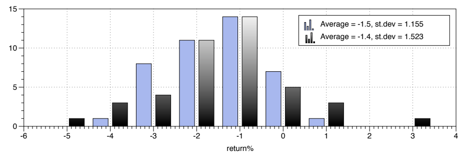

Figure 5.42: One month return distribution for the 42 selected companies

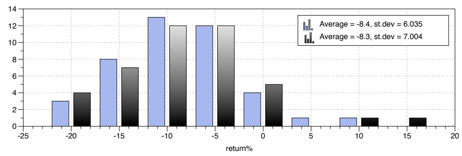

Figure 5.43: Half year return distribution for the 42 selected companies

As one can see in Figures 5.40, 5.41, 5.42 and 5.43, the distribution of returns for both methods is plotted. Next, we investigate whether the improved aggregation method returns significantly differ from the original method.

In order to test whether or not the returns of the improved aggregation method significantly perform better then the original method, a test of significance is conducted for each period.

As discussed in Section 5.4.1, the same Wilcoxon signed-rank test is conducted to determine whether or not the method significantly differs from the original method.

First the null hypothesis and alternative hypothesis are defined, as displayed below:

- $H_0 : \mu = \mu_0$
- $H_a : \mu > \mu_0$

The null hypothesis H_0 states that there is no difference between the mean return of the original aggregation method μ_0 and the improved aggregation method μ. The alternative hypothesis H_a states that the mean return of the improved aggregation method μ is greater than the original method μ_0. In case of rejection of the null hypothesis, the alternative hypothesis is accepted, which leads to the conclusion that the improved aggregation method performs better.

The null hypothesis is not rejected when the mean of the improved aggregation method is below the mean of the original method. However, if the mean of the improved aggregation method is above the mean of the original method, the null hypothesis is rejected and the alternative hypothesis is accepted.

The level of significance α is set to 5%, which means the probability of making a Type I Error, rejecting the null hypothesis while the null hypothesis is true, is set to 5 percent. In Table 5.7, the null hypothesis is tested for the one day, one week, one month and half year return.

As one can see in Table 5.7, the one day and one week return is smaller than the significance level of 0.05, which means for the one day and one week return the null hypothesis is rejected and thus the improved aggregation

Table 5.7: Significance test of improved aggregation method.

	1 Day Rtrn	1 Week Rtrn	1 Month Rtrn	$\frac{1}{2}$ Year Rtrn
μ_0	0,13	-0,33	-1,5	-8,4
$E(T)$	451,5	451,5	451,5	451,5
σ_T	79,97	79,97	79,97	79,97
z	-5,0952	-1,7692	-0,0937	-0,4313
p	0,0001	0,0384	0,4626	0,3330

method performs significantly better. In contrast, for the one month and half year returns, one cannot reject the null hypothesis and therefore the null hypothesis is not rejected for these periods. This means their is no evidence to state that the improved aggregation method performs better for the one month and half year return.

Table 5.8: Comparison of performance of original and improved aggregation method.

Company	IPO	Day Rtrn	Week Rtrn	Month Rtrn	$\frac{1}{2}$ Year Rtrn
Yahoo	Old	0,1041	-1.5051	-2.7336	-10.9719
	New	0,375	-1.5812	-1.7822	-9.8186
WellPoint	Old	0,0833	-0.7324	-2.0538	-9.4523
	New	0,1666	-0.4692	-1.6895	-9.8564
UPS	Old	0,3541	0.2983	0.5920	-3.0593
	New	0,7501	0.1468	-0,1692	-1.6767
Polo RL	Old	0,0625	-0.0041	-2.4099	-12.3690
	New	0,8291	-0.5130	-2.1665	-10.9786
NOV	Old	0,1666	-2,5069	-0,3560	-11,6970
	New	0,4583	-2,5522	-0,7370	-11,6280
Google	Old	0,125	-0.7268	-2.0394	-11.3690
	New	0,1916	-0.7859	-2.0003	-11.9759
Amazon	Old	0,7291	-1.0616	-0.9248	-8.3809
	New	0,1041	-0.3276	-1.4045	-4.6551

Conclusion

In the above conducted experiment, the improved aggregation method is tested against the original aggregation method. On the short run (one day and one week), the improved method is significantly performing better. On the long run, there is no significant difference anymore. This can be explained by the fact on the long run prices fluctuate more than on the short run, which lead to equal performance. In addition, the short run is the most important period, yielding higher returns on a short notice is more interesting then on the long run.

5.5 Discussion

In the previous sections (Sections 5.4.2 and 5.4.1), the performance of the improved aggregation method and the performance of the IPO underwriters has been tested. At first, tests where conducted to measure the impact of underwriters who have done the Initial Public Offering for a certain stock. The results are divided into market capitalization of the stocks on which it is tested. For small, mid, and large-cap companies, IPO underwriters add additional value to the aggregation process. However for large-cap companies on the short run, excluding IPO underwriters yields in most cases higher returns, but not enough to significantly differ and thus rejecting the null hypothesis (which states: IPO underwriters yield the same or more returns).

Next, the improved aggregation method is tested against the original aggregation (with all underwriters included). At a 5% significance level, the improved aggregation method significantly performs better on the short run (one day, one week), while on the longer run, the null hypothesis cannot be rejected. Although this is logical, as on the long run prices more fluctuate and thus leading to equal performance for both methods. Therefore the short run is the more important period.

Chapter 6

Conclusion and Future Work

In Section 5.2, the improved aggregation method is tested against the original method. In addition it is tested whether or not IPO underwriters add additional value to the aggregation process. In this section, answers will be given to the research questions stated in Section 1.2. In addition, future research will be discussed.

6.1 Conclusion

As stated in Section 1.2, in this thesis a method is proposed to aggregate market advices. Based on the literature survey, this improved aggregation method is extended with a method to correct bias. According to the literature survey, bias still occurs in market recommendations, and correction for this bias could improve the informational value of the recommendations (see Section 2). Therefore in this thesis a correction method is developed on top of the improved aggregation method. This bias correction method corrects bias by giving different weights to underwriters who have done the Initial Public Offering for the company whose advices are aggregated.

In the experiment chapter (Chapter 5), first the correction method is tested. This is done by giving giving IPO underwriters different weights (from 0 to 1) and examination of the outcome. At a significance level of 5%, one cannot state that giving less weight to underwriters who have done the IPO process for the company in subject yields higher returns.

Next, the improved aggregation method, in which advisors are given adjusted weight (a combination of a experience factor and a smoothed performance figure), is tested against the original aggregation method developed by (Stibbe, 2007). At a significance level of 5%, one can say the improved aggregation method yields higher returns on the short to medium term, while on the longer run the effect is diminishing and does not significantly differ anymore. This results is very important, since the short period is the most important period in respect to returns (investors like it more to get higher returns on short periods than over long periods). In addition, over longer periods prices more fluctuate, which cannot be handled by the methods. This leads to equal performance on the long period.

Therefore one can state that an aggregation method, which gives advisors an experience factor and a smoothed return figure, yields higher returns on the short run and therefore is a good aggregation method. However one cannot state that correcting bias by giving less weight to advisors who have done the Initial Public Offering for the company whom advices are aggregated, yields higher returns.

6.2 Future Work

In the original application by (Stibbe, 2007), there are only three advice types, this could be extended by adding more advice types (adding e.g. strong buy and strong sell) and thus have more distinction between different aggregated advices. Furthermore there could be more research done into the calculation of a hold advice type. Although it is discussed in this paper that it should not be handled equally as a buy advice, more research should be done in designing a good method for calculating the performance of a hold advice.

Besides focusing on the advice types, one could extend this method with other bias correction methods, such as an method which automate investigate the past of an underwriter in respect to bankrupt firms and based on that information performs some correction method in regards to the underwriter.

Finally, it would be very interesting to design an method which can interpret news messages, which can add additional useful information to the aggregation of market advices and thus leading to possible better aggregation of advices. As discussed in the literature survey, solely recommendations does not include all information. Accordingly there where numerous examples in which buy advices where issued for companies, while news messages states heavily loss for those particular companies. Later on these companies went bankrupt. Incorporating relevant news messages and understanding the meaning of these messages will possible add informational value to the aggregation process.

Glossary

Abnormal Returns The abnormal return, is the return for one stock or a portfolio in contrast to the average market return for a particular broad index. This lead to the following equation $AbnormalReturn_i = Return_i - MarketAverageReturn$, in which i is a particular stock in a portfolio, p. 11.

Agency Costs Agency costs are internal costs for the organization, which occurs because agents (employers) acting on behalf of the organization. The costs consist of two main sources. First of all risks associated with agents using organization resources for their own benefit. Second, stock options for managers to keep their interest and let them perform.

Brokerage Firm A brokerage firm buys and sells securities for their clients and for themselves. In addition these brokerage firms have research departments for their client benefit, p. 32.

Due Diligence Due Diligence is the process before the initial public offering. The due diligence process audits whether the statements made in the prospectus hold. More credibility increases the chance of a successful IPO, p. 13.

Excess Returns Excess returns are the returns over the risk-free returns. Risk-free returns are those returns obtained by investing in risk-free portfolio's.

GICS Global Industry Classification Standard. Each stock listed company is assigned to a sub-industry, to an industry, industry-group and sector based on the principal business activity of the company, p. 36.

IPO IPO stands for Initial Public Offering and refers to the event in which a particular private company goes public by offering stocks on a official stock exchange, p. 12.

Large-cap Large-capital companies are companies with a market capital which exceeds the $10 billion, p. 54.

Market Capitalization Market Capitalization is the measurement of the corporate size of a company, which is based on the product of the stock price and the amount of outstanding stocks, p. 54.

Market Risk The chance that a stock decreases in price due to external factors, such as general decline in financial markets, p. 12.

Mid-cap Medium-capital companies are companies with a market capital between $2 billion dollar and $10 billion dollar, p. 54.

Momentum The momentum is the empirically observed tendency of stocks who performed well in the past to still outperform stocks who performed poor in the past, p. 12.

Reputation Capital Reputation capital is a quantitive measure of a company his reputational value. The Intangible Asset Finance Society defines reputation capital as the sum of all the intangible assets of a company, p. 30.

Small-cap Small-capital companies are companies with a market capital between $200 million dollar and $2 billion dollar, p. 54.

Bibliography

Aczel, A., and Sounderpandian, J. 2002. *Complete Business Statistics*. McGraw-Hill/Irwin: The McGraw-Hill Companies, 5. ed. edition.

Allen, M., and Isreal, S. 2003. An analysis of the relevance and bias of analyst recommendations: The case of bankrupt companies. *The Financier* 10(1-4):25–45.

Balboa, M.; Gmez-Sala, J. C.; and Lpez-Espinosa, G. 2009. The value of adjusting the bias in recommendations: International evidence. *European Financial Management* 15(1):208–230.

Banko, M., and Moore, R. C. 2004. Part of speech tagging in context. In *In Proceedings, 20th International Conference on Computational Linguistics (Coling 2004)*, 556–561. International Conference on Computational Linguistics.

Barber, B. M.; Lehavy, R.; and Trueman, B. 2007. Comparing the stock recommendation performance of investment banks and independent research firms. *Journal of Financial Economics* 85(2):490–517.

Barber, B. 2001. Can Investors Profit from the Prophets? Security Analyst Recommendations and Stock Returns. *Journal of Finance* 56(2):531–563.

Bjerring, J.; Lakonishok, J.; and Vermaelen, T. 1983. Stock Prices and Financial Analysts' Recommendations. *Journal of Finance* 38(1):187–204.

Carleton, W. T.; Chen, C. R.; and Steiner, T. I. 1998. Optimism biases among brokerage and non-brokerage firms' equity recommendations: Agency costs in the investment industry. *Financial Management* 27(1):17–30.

Conrad J., Cornell B., L. W. R., and Rountree, B. R. 2006. How do analyst recommendations respond to major news? *Journal Of Financial and Quantitative Analysis* 41(1):25–60.

Gomez-Perez, A., and Manzano-Macho, D. 2003. Deliverable 1.5: A survey of ontology learning methods and techniques. Technical report, OntoWeb.

Hongwei Zhu, M. D. S., and Madnick, S. E. 2001. Information Aggregation A Value-added E-Service. In *Proceedings of the International Conference on Technology, Policy, and Innovation: Critical Infrastructures*, 26–29. International Conference on Technology, Policy, and Innovation.

Hoovers. Hoovers IPO Database. See http://www-2.hoovers.com/global/msn/index.xhtml?pageid=1949; retrieved June 29, 2009.

KDE Techbase. KDE Usability Guidelines. See http://techbase.kde.org/Projects/Usability; retrieved September 30, 2009.

Marcus, M. P.; Santorini, B.; and Marcinkiewicz, M. A. 1994. Building a large annotated corpus of english: The penn treebank. *Computational Linguistics* 19(2):313–330.

Mcnichols, M. F.; O'Brien, P. C.; and Pamukcu, O. M. 2006. That Ship has Sailed: Unaffiliated Analysts' Recommendation Performance for IPO Firms. *SSRN eLibrary*.

Michaely, R., and Womack, K. L. 1999. Conflict of Interest and the Credibility of Underwriter Analyst Recommendations. *Review of Financial Studies* 12(4):653–86.

Milea, V.; Frasincar, F.; Kaymak, U.; and di Noia, T. 2007. An OWL-Based Approach Towards Representing Time in Web Information Systems. In *The 4th International Workshop of Web Information Systems Modeling Workshop (WISM 2007)*, 791–802. Tapir Academic Press.

Milea, V.; Frasincar, F.; and Kaymak, U. 2008. The TOWL Web Ontology Language. In *The 20th Belgian-Dutch Conference on Artificial Intelligence (BNAIC 2008)*, 343–344.

Morgan, J., and Stocken, P. 1998. An Analysis of Stock Recommendations. Technical report, Princeton, Woodrow Wilson School.

Ramnath, S.; Rock, S.; and Shane, S. 2006. A Review of Research Related to Financial Analysts Forecasts and Stock Recommendations. Georgetown University and Leeds Business School (unpublished).

Schmid, H. 1997. Probabilistic part-of-speech tagging using decision trees. In Jones, D., and Somers, H., eds., *New Methods in Language Processing*, Studies in Computational Linguistics, 154–164. London, GB: UCL Press.

Security and Exchange Commission. Securities Act of 1933. See http://uscode.house.gov/download/pls/15C2A.txt; retrieved September 30, 2009.

Security and Exchange Commission. The Global Research Analyst Settlement. See http://www.sec.gov/spotlight/globalsettlement.htm; retrieved September 30, 2009.

Stibbe, R. 2007. Towards Automated Trading On The Semantic Web. Master's thesis, Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam.

Welty, C. 2006. A Reusable Ontology for Fluents in OWL. In *Proceedings of FOIS-2006*, 226–236. Baltimore:IOS Press.