LIBERA UNIVERSITÀ INTERNAZIONALE DEGLI STUDI SOCIALI

Factors affecting the accuracy of mean analysts' forecasts

With a focus on dispersion, and the implementation of a forecast improving model

Antonio Orsini
ID 176561

Supervisor: Giuseppe Ragusa

Co-Supervisor: Siria Angino

A thesis presented for the degree of
Economics and Business

Department of Economics and Finance
Academic year 2014/2015
Italy

Contents

Chapter 1

Introduction

Nowadays still, many private investors, banks and institutions rely on analysts forecasting in order to make decisive investment choices, but at the same time the sector concerning quantitative forecasting is increasingly stimulating the interest of the whole financial environment.
One of the most indicative variable, which attracts the interest of forecasters and investors is without any doubt the future net income of a company.

Net income is a good approximate measure of how a firm will perform in the future years, and part of it is still used in many price evaluation methods. When looking at net income forecasts by analysts, often the actual figure used to represent the whole sample of predictions, is simply the mean of the various analysts' forecasts, because it saves time on evaluating every single specialist opinion, and often includes both optimistic and worse scenarios in its core.
This paper will focus on studying the reliability of the mean analysts' forecasts for net income, across a wide range of companies over few years.
Literature on this subject is not present at the time being or supplies unsatisfactory explanation.
In particular, this study tries to analyse whether some variable affect the error made by the mean analysts forecast on the real value of future net income, and ultimately if this forecasts could be improved by the results obtained.

The first and basic model used for explanatory purposes, is an econometric linear model, which focusses on just one variable as particularly affecting or explaining this error made by analysts, the distribution of forecasts, commonly associated to the variance of the mean of analysts' forecasts.
The primary assumption is that as variance increases, this indicates an increase in uncertainty among forecasters, then signalling that specialists do not significantly agree on the future estimations of net incomes. On the other end,

if the variance is lower, there is a signal of cohesion among analysts' ideas and speculations, and this will probably lead to a smaller error of the overall measure, simply because agreement may be linked to more solid information available in the market.
The model is later enlarged, and other variables are added, which have already been found to have diverse correlations with the accuracy of forecasters, so to improve the explanatory capabilities of the econometric regressions. These other factors are proxies for geographical institutions, relative size of the company, industry types, companies governance disclosure level and speed and constancy of revenues growth.

Results from these models show that all the variables studied, and especially dispersion, are correlated with the magnitude of the error made by mean analysts' forecasts, in different ways.
With these findings, a model to improve the forecasts is built and compared to a well regarded econometric model, the Random Walk Forecasting, to test the possible forecasting improvements stemming from this paper. The model seems to be working and improving the actual analysts' predictions despite some technical issues which may be solved by further research.

Chapter 2

Literature Review

2.1. Previous researches

Net income forecasting has long been the center of numerous studies given its importance on the financial markets.
Slack et al.(2004) noted that even though forecasts can include various degrees of errors, they are necessary to support the management decision making process in any business. In fact, nowadays, most of the businesses listed in the market are under scrutiny by analysts and forecast specialists who try to give their best prediction on the future income or losses of the company regardless of the type of business, in fact according to Wallace and Stahl (2008) every business can forecast, even if the level of forecast accuracy might vary significantly between business environments.
Managers use forecasts to adjust operations and better evaluate the viability of diverse options.

At the same time not only managers keep a close look on net income forecasts, but also investors, private and institutional, or even creditors, like banks, who need to ascertain the solvency of a debtor, heed the forecasts.
Givoly, Dan, and Josef Lakonishok (1984) recognize the importance of forecasts, and the fact that in particular analysts' forecasts are often used by investors in taking their decisions, they found analysts opinions to be superior to time-series model and to be a good proxy for market expectations in general. Prior to that also Fried, Dov, and Dan Givoly (1982) suggested the supremacy of analysts' forecasting over time-series models as providing better surrogates for market expectations than the models, and investigated on some of the factor contributing to the performance of analysts.

These findings, although being still very relevant in the field, should be observed considering that since the 80's forecasting models have considerably

improved, and researches on the matter showed a slightly different picture. Bradshaw, Mark T., et al. (2011) go over the undisputed supremacy of analysts' forecasts with a different perspective, finding that on the long term forecasting models tend to be more precise then analysts, and should then be taken into account when analysing different predictions.
Nonetheless, analysts forecasts remain one of the more accurate if not the most accurate forecasts on net incomes, especially in the short term. Therefore it can be stated that net income forecasts and in particular the one made by analysts do have an effect on the market behaviour and investment decisions, deserving a particular and in detail attention.

Often investors interested in analysts' forecasts rely on different analysts' predictions, basing their decisions on the analysts' past history of successes, but still a great part of the financial markets looks at the mean of those different analysts' forecast, as a consensus and aggregate of their predictions. Kim, Oliver, Steve C. Lim, and Kenneth W. Shaw.(1998) forward a critique on the mean as a good predictor and aggregation of forecasts, because it gives too much weight on the common informations available in the market and it does not give the right importance to private information held by each analysts, problem which is tried to solve by Bloomberg new adjusted measure for income forecasts, which takes into account the past result and relevance of each single analyst. Nonetheless they still agree that the mean of forecasts is a starting point to improve, and it is widely used not only by investors but even for accounting purposes.

The problem that arises is how much can we rely on these mean forecasts', if there is some variable which statistically affect these figures and if there is some quantitative method to better understand the functioning of these aggregate predictions.
The system proposed in this paper, relies on a simple linear model which tests the influence of some particular variable on the accuracy of mean forecasts' across a wide range of companies and over a timespan of six years. This model is based on some fundamental assumption which follow logic reasoning and previous quantitative studies' results here presented.

Orie E. Barron, and Pamela S. Stuerke (1998) research examines whether dispersion in analysts' earnings forecasts reflects uncertainty about firms' future economic performance. They base their own results on previous papers as Daley et al. (1988), and Imhoff and Lobo (1992). Abarbanell et al. (1995), argue that, if forecast dispersion after (i.e., conditional on) an earnings announcement reflects uncertainty about firms' future cash flows and this uncertainty causes investors to desire additional information, then dispersion will be positively associated with both (a) the level of demand for more information and (b) the magnitude of price reactions around the subsequent earnings release. In their study they find positive and significant correlation to both factors thus not rejecting their initial hypothesis.

Hope, O.-K.,(2003) instead finds that forecast accuracy is positively associated with disclosure level and the enforcement of accounting standards, suggesting that managers are more likely to follow prescribed accounting rules when they know that their accounting choices are better monitored.

Also macroeconomic characteristics seem to be relevant on the matter. Gauri Bha, Ole-Kristian Hope, Tony Kang (2006) outline the importance of governance transparency on the accuracy of forecasts, adding the proposition of an increasingly significant effect of such transparency levels in an environment of weak legal enforcement, while Thiess Buettne and Bjoeran Kauder (2010) outline that there are significant differences in the conditions and practies of revenue forecasting in the OECD countries other than tax structures and macroeconomic fundamental.

But as difference in macroeconomic variables may be fundamental to understand and capture the error made by mean analyst' forecasts, also differences among industries and sector may reveal to be a key factor.
Kwon, Sung S. (2002) provide evidence for lower unsigned error and dispersion in high-tech firms vis-a-vis low-tech firms. The higher forecast accuracy and forecast convergence for high-tech firms relative to low-tech firms in financial analysts' forecasts of earnings can be attributed to the information effect prevailing over the noise effect.
Das, Somnath, Carolyn B. Levine, and Ki Sivaramakrishnan(1998) present and elaborate with empirical data the assumption that analysts tend to make higher prediction, and result in a wider positive error, for firms which operate in less predictable industries, so that they may have better access to information from the direct management of the firm.

Another important forecasting practice has long been to look at past performances and values, and the speed rate at which fundamental characteristics of the studied variables change.
On this argument, Lundholm, Russell, Sarah McVay, and Taylor Randall (2010) can surely shed some light, outlining the importance of previous years revenues and sales, to next year outcomes. The study finds that speed of growth of a company's revenues may affect the predictability of its incomes,

given that companies that grow at a faster pace may be in a development face and may be easier to predict their success, instead more stable companies may follow rather unpredictable patterns of high and low.

At last, Lang, M., and R. Lundholm (1996) finds that forecast accuracy is positively related to corporate disclosure policies and firm size. As a company grows larger it may attract the interest of more investors and consecutively larger amount of moneys may depend on well accurate forecasts, implying that probably better quality and experienced analysts will evaluate and predict the future incomes of the firm.

2.2. Assumptions

All of the paper and researches previously listed, generate perfectly fertile ground for the supposition of some relevant assumption, which will lead to the creation of few fundamental factors that are tested for significance on affecting the accuracy of mean analysts' net income forecasts.

The first element tested is the dispersion and distribution of analysts' forecasts which make up the aggregate mean forecasts analysed in this paper, and it will also be relevant and included in all the different variation of the model here presented. Is possible to assume that this dispersion in analysts forecasts could be used as a good proxy for market uncertainty about a firm's future net income.[1] This assumption then, is relevant in the model presented, since uncertainty about future incomes might be correlated with the precision of the aggregate mean of analysts' forecasts, again implying that more uncertainty would mean less accuracy of the measure.
Following this main assumption and the principal explanatory factor here analysed, there are also other variables of which influence on mean forecast accuracy is studied.

Analysts dispose of a variety of information both private and public, is therefore essential to introduce some form of information disclosure in the model.
Since forecast accuracy is found to be positively associated with disclosure level[2] a measure of corporate and general disclosure could well explain some of the error made by mean analysts' forecast.

[1] Orie E. Barron, and Pamela S. Stuerke (1998).
Daley et al. (1988), and Imhoff and Lobo (1992). Abarbanell et al. (1995).
[2] Hope, O.-K.,(2003).

Another important observation, is that different countries have different general structures, both fiscal and legal[3], is than reasonable to assume that these factors, together with the different accounting standards followed by distinct companies, will interfere with the precision of cluster of analysts' forecasts among countries.

Macroeconomic characteristics may also have to do with the growth and stability of a country, probably in region which are experiencing a constant and sustained growth, companies and industries will likely converge to a similar rate of growth, or at least will grow in similar ways. Instead in regions of the world with a more stable growth, different situation could be observed, as companies that are failing or new industries which are quickly climbing to the top. These will then affect the precision of the analysts working in so different situations, as it could be easier to predict revenues in a country where all companies converge to a similar growth instead of predicting figures in a region where things are going in different direction for each player on the market.
For such reason a proxy for macroeconomics difference is needed, and the most simple way to obtain it is just to control for the different domiciles of firms.

A similar effect to the differences stemming from diverse macroeconomic factors, could be observed across diverse types of industries. Empirical results[4], show how significant and important may be the difference and bias present among industry sectors, therefore belonging to a different sector instead of another should be taken into account in the model presented.
Considering instead variation among firms in relative largeness and growth rate[5], it could be assumed that both speed of growth of net income and previous years revenues have some relation with the error made by mean analysts' forecasts.

Summing up, the models later introduced rely on the above mentioned empirical literature and the relative assumptions that directly stem from correlated studies findings, and different proxies for the measure cited are implemented and explained.
While some of these factors have already been proven to have some effect on analysts' performances the assumption here made take a step forward

[3] Thiess Buettne and Bjoeran Kauder (2010).
[4]. Kwon, Sung S. (2002).
Das, Somnath, Carolyn B. Levine, and Ki Sivaramakrishnan(1998).
[5] Lundholm, Russell, Sarah McVay, and Taylor Randall (2010).

introducing similar arguments to the clustered measure of the mean analysts' forecast, and focussing on a variable, the forecasts dispersion, on which there is not much empirical research done.

Chapter 3

The Data

In this section there will be a brief review of the data used in the model later introduced, as well as a clarification and explanation on some of the proxies used to represent the variable assumed to be relevant in the previous chapter.

3.1. Bloomberg

Due to the financial nature of the study here proposed, there is a need of a good, accurate, modern and professional source of data, which could also be reliable and relatively large in scale.
For these reasons the data used for this research are taken from Bloomberg databases. Bloomberg L.P. is a privately held financial software, data and media company headquartered in New York City, which provides financial software tools such as an analytics and equity trading platform, data services and news to financial companies and organizations through the Bloomberg terminal.

Being one of the most used financial software in the sector, Bloomberg databases are wide both in scale and in time, meaning that it is possible to acquire data on a vast range of companies over a time period of six year, which allowed for more observation, summing up to 28,6168, leading to an overall improved reliability of results.
Another advantage of Bloomberg databases is given by the fact that they are largely international, therefore it is possible to source data on companies which have domicile and are situated in different countries, about eighty, which are disseminated in the whole five continents, and which can be classified as developed, or developing economies.
Moreover Bloomberg data go beyond the simple figures coming from yearly disclosures as sales revenues or costs, since its database incorporate a number of information which are essential to build relevant proxies and measures of

the variables proposed in this paper by the theory and the related assumptions. In particular, Bloomberg is a useful source of aggregate information about analysts' previsions. It is capable of computing by itself mean and standard deviation of analysts' forecasts, which is essential for the construction of the model here presented.

3.2. Proxies

In Chapter 1, through the study of past researches, various assumption are made, which need to be transformed and represented by material data and measurable figures. The Bloomberg database offers a wide range of applicable variables, which are able to reflect the assumptions coming from the theory into numbers, and act as proxies for the factors hereby listed:

1. Forecasts dispersion;
2. Governance disclosure and information openness;
3. Macroeconomic differences;
4. Industry and sector types;
5. Relative size;
6. Past performances and speed of growth.

In this study, forecasts dispersion, refers to a quantitative measure, of how spread out and distributed are the analysts' forecast, relative to the average, as a mean to understand how analysts' opinions differ from each other. A statistical measure to quantify dispersion is found precisely in the standard deviation, which Bloomberg can easily source and calculate for any given year and company.

On the same database is also possible to find a good proxy for general, and in particular governance, disclosure level, and that is the ESG Bloomberg score. ESG score starts from 0.1 for companies that disclose a minimum amount of environmental, social and governance data to 100 for those that disclose every data point collected by Bloomberg.
The same scale applies to the sub-measure of ESG score which takes only into account governance disclosure, for that each data point is weighted in terms of importance, with board of directors data carrying greater weight than other disclosures. The score is also tailored to different industry sectors, in this way, each company is only evaluated in terms of the data that is relevant

to its industry type.
Unfortunately such data are not available for all the companies selected to be part of the sample analysed, therefore the model including ESG level and governance disclosure score has relatively less observations (20,499).

Regarding macroeconomic differences instead, an easy method to integrate fundamental macroeconomic structure, different tax practices and legislation systems, is simply to control for the different countries. From the database of reference, denomination of the country of domicile is extracted, for each firm, creating a sample containing companies from almost 80 different countries.
To account for differences across industries, there is need of a comparable denomination for industries and sectors. the Industry Classification Benchmark (ICB) is an industry classification taxonomy launched by Dow Jones and FTSE in 2005 and now owned solely by FTSE International. It is used to segregate markets into sectors within the macroeconomy. The ICB uses a system of 10 industries, partitioned into 19 supersectors, which are further divided into 41 sectors, which then contain 114 subsectors.
The ICB is used globally to divide the market into increasingly specific categories, allowing investors to compare industry trends between well-defined subsectors.
In this study, industry classification is included in the model to capture consistent differences among analysts practices and accuracy in different industries.

Enterprise value is another important and acknowledged factor which is directly related to company size. Enterprise value is composed by market value plus enterprise value components (Preferred Equity + Minority Interest + Total Debt - Cash and Equivalents – Adjustments) and it can be thought as the theoretical takeover price if the company were to bought.
Larger and smaller companies surely have distinct enterprise value, therefore making EV a good proxy for relative size of a firm. But often firm with slightly different enterprise value may not really have a considerably different size. For such reason, the firms sample is split in four different quartiles according to EV largeness, which better capture the magnitude and difference between large and small firms.

Chapter 4

Error measures

Different size of the firms making up the sample considered in this study, is surely an advantage in terms of a higher reliability and comprehensiveness of the sample, but may also lead to some practical problem in the functioning of the model and the comparability of the errors made by mean analysts' forecasts across companies.
The theoretical idea behind a possible drawback, is quite simple and linear: The error made by analysts' could be systematically correlated to the size of net incomes, and in turn with the main variable studied, the dispersion of analysts' forecasts.

For example, let's take into consideration two different firms, the first makes one billion of profit each year, and the second makes one million of profit each year. Now, looking from a realistic prospective, is easy to speculate that forecasters will make larger errors in terms of volume in forecasting the income of the firm which makes one billion of profit each year. Even if in both of the companies forecasters make an average error of 10% on each firm, the error made on the first one would be relatively greater, therefore the dispersion among forecasts could be systematically greater in volume or the other factor may not result as being so significant as they should if firms of the same size would be compared.

At this point when studying the correlation between the error made by mean analysts' forecasts and the dispersion, the standard deviation of forecasts, a systematic correlation may arise which may not be due to the fact that dispersion, and therefore uncertainty among forecasters, is relevant in studying the error made, but just be due to the relative size of different firms. While this problem may arise and should therefore be taken into account, it could also be the case that it doesn't, since the emerging of such hurdle comes only by supposition.

For such reasons there is a need of finding also different measurement for both errors and relative dispersion, in order to standardize companies of different sizes and tackle the possible problems arising from their dissimilarities.

4.1. Theoretical measures

The first proposed theoretical measure of error, is just a simple transformation of the mean squared error, the root mean squared error (RMSE)[1]. Let us denominate F_t as the forecast at time t, and Y_t=Net Income at time t. Then define the forecast error as $e_t = Y_t - F_t$.

$$MSE = mean(e_t^2)$$

$$RMSE = \sqrt{MSE}$$

This measure is simple and relatively reliable, mostly because does not undergo particular transformations and it will be useful when trying to calculate a precise estimate for the error. However, as Chatfield (1988) and Armstrond and Collopy (1992) pointed out, the MSE is not appropriate for comparisons between series as it is scale dependant.

If we then consider the previous speculation about a possible systematic correlation which would hamper the explanatory power of the model, it is clearly needed an alternative measure for the error studied, in order to standardize different sizes of revenues.
The second measure of error here introduced, was proposed by by Makridakis (2000) and is a transformation of another measure used by the same Makridakis during the original M-competition 1982.
The original error measure was the mean absolute percentage error (MAPE):

$$MAPE = mean\left(\left|\frac{100e_t}{Y_t}\right|\right)$$

Unfortunately this measure was noticed to put a heavier penalty on positive errors than on negative errors, and most relevant for this specific case, was very sensible to outliers, therefore if applied to the study proposed would not value equally enough all the firms' different forecasts.

[1]References from Gooijer, Hyndman (2006)

Makridakis (2000) proposed a symmetric version of the same measure (sMAPE), which was bounded to an upper value of 2, sensibly reducing the effect of outliers, standardizing better the whole error and also being more balanced in weighting negative and positive values.

$$sMAPE = mean\left(\frac{2 * |Y_t - F_t|}{(Y_t + F_t)}\right)$$

4.2. Adapting to the model

The error measures previously presented are essential in the construction of the model later introduced, nonetheless there is a need of performing some transformation in order to make them more fit to the data, and efficient in their estimation.
Let us denominate now the variables with the real data which will be implemented in the subsequent model, mF_t as the mean analysts' forecasts at year t, while Y_t remains the real net income at year t[2]. This first error will be called "positive gross error" as it represents the basic error made by the mean analysts' forecasts at year t, taken always in its positive form.

$$PosGroErr_t = \sqrt{(mF_t - Y_t)^2}$$

The second measure of error used take its original roots from the sMAPE. In this particular case more positive forms of the single values are needed in order to get a more balanced error, therefore the denominator of this measure is split in two components which are both made positive by taking the square root of their respective square values.
If such transformations where not implemented, then given the fact that forecasts may be both negative or positive in value, and their arithmetic sum may end up being either positive or negative, or very near zero, the measure would loose its peculiar ability to reduce the error boundaries and eliminate most of the outliers bias. This measure is here called "symmetric error" since its theoretical fundamentals steam directly from the symmetric MAPE.

$$SymErr_t = \left(\frac{2 * \sqrt{(mF_t - Y_t)^2}}{\sqrt{mF_t^2} + \sqrt{Y_t^2}}\right)$$

or

[2] Data for Net Income are also taken from Bloomberg database, and refer to the actual net income of a company at year t, as by public internal disclosure.

$$SymErr_t = \left(\frac{2 * PosGroErr_t}{\sqrt{mF_t^2} + \sqrt{Y_t^2}} \right)$$

But for the standardization to have a complete and relevant effect, not only the measure of error should be standardized for the size of different observation, but also the variables which may be size dependant should be standardized. In particular in this study, as previously stated, the dispersion might be related to size, as if there is a larger value for a company's net income, probably the analysts' forecasts will diverge in larger magnitude even if their relative percentage precision to forecast for a smaller industry may be similar.
The easiest method to standardize the dispersion in this case, is to take the percentage standard deviation in respect to the different real net incomes of the companies.

$$\sigma_t^{Rel} = \left(\frac{\sigma_t}{Y_t} \right)$$

In this way, not only the relative dispersion will be comparable across different industries, but also it will be not systematically related to the error measure.

If the error measure would have been standardised with the same transformation applied to the dispersion, there would have been a false correlation between the error and the dispersion, due to the net income changing its magnitude and affecting both factors.

$$\left(\frac{PosGroErr_t}{Y_t \uparrow} \right) \uparrow \;\sim\; \left(\frac{\sigma_t}{Y_t \uparrow} \right) \uparrow$$

Using instead a combination of different transformations, the SymmErr for the error, and a standard percentage for the dispersion, the two measure will not have a systematic bias.

Thanks to the SymErr and the related transformations the result of the model will be more solid and comparable since all observation will have about the same weight relative to each other. The PosGroErr will be instead fundamental to have an unmodified view of the result of the model, trying to see whether natural correlation applies and most of all, the measure will be important for prediction purposes.

Chapter 5

Explanatory Models

This chapter tries to empirically verify and explain the theoretical hypotheses which were presented in the previous sections, through the use of linear econometric models.

5.1. Testing dispersion

The first group of regressions testing the main hypothesis, which is the focus of this study, refers to the relevance of dispersion of analysts in affecting the accuracy of the mean analysts' forecasts, again the hypothesis under this connection is that dispersion mirrors the uncertainty and information discrepancy in the market.
The first two linear models try in fact to verify this relation empirically and quantitatively, but using the two different errors and dispersion measure transformations. In the first model the dependent variable, the error, is the $PosGroErr_t$ previously introduced, while the dispersion is measured simply by the unmodified standard deviation, here called $BEstNISTDDEV_t$. In the second regression instead the dependent variable is represented by the $SymErr_t$, while the dispersion is measured as the σ_t^{Rel}, which is here referred to as $RelBEstNISTDDEV_t$.
This two different models present two view of the same object, the first one is unmodified, while the second one is transformed to account for firms' size inequality

$$PosGroErr_t = \beta_0 + \beta_1 BEstNISTDDEV_t + \varepsilon_t$$

$$SymErr_t = \beta_0 + \beta_1 RelBEstNISTDDEV_t + \varepsilon_t$$

The results[1] reveal a significant and positive effect of the dispersion in both of the models. Therefore, the hypothesis of dispersion as a measure for

[1]The results of this models are presented under table A.1 in Appendix A.

uncertainty among analysts cannot be here rejected. The positive sign of the independent variable only strengthen the suppositions, since a high level of uncertainty would surely hamper the precision of the forecasters as a group.

5.2. Testing auxiliary variables

After having tested the basic significance of the main variable in this study, the next models focus on gradually adding the explanatory variables of which hypothesis were previously discussed. The regression are carried out using both the modified and the unmodified measure as done in the two previous models.
In the first linear model, other than the dispersion, are introduced variables referring to speed of growth, and size of the firm, which is measured observing the relative enterprise value of each company, and its relative belonging to one of the four enterprise values size quartiles of the entire market, introduced as dummy variables[2].

$$PosGroErr_t = \beta_0 + \beta_1 BEstNISTDDEV_t + \beta_2 NI_g + \beta_3 Quart + \varepsilon_t$$

$$SymErr_t = \beta_0 + \beta_1 RelBEstNISTDDEV_t + \beta_2 NI_g + \beta_3 Quart + \varepsilon_t$$

Then, are added dummy variables linking every firm to its country of domicile, and to the relative industry in which they operate, according to the ICB industry denomination.

$$\begin{aligned} PosGroErr_t = {} & \beta_0 + \beta_1 BEstNISTDDEV_t + \beta_2 NI_g + \beta_3 Quart \\ & + \beta_4 CNTRYofDMC + \beta_5 IndName + \varepsilon_t \end{aligned}$$

$$\begin{aligned} SymErr_t = {} & \beta_0 + \beta_1 RelBEstNISTDDEV_t + \beta_2 NI_g + \beta_3 Quart \\ & + \beta_4 CNTRYofDMC + \beta_5 IndName + \varepsilon_t \end{aligned}$$

At last, proxies for governance disclosure and general level of disclosure are added. Unfortunately this latter model has less observation on which to base its result, since information disclosure proxies provided by Bloomberg are not available for every firm in this database, the companies missing this particular value were excluded.

[2]The results of this models are presented under table A.2 - A.3 in appendix A.

$$PosGroErr_t = \beta_0 + \beta 1BEstNISTDDEV_t + \beta_2 NI_g + \beta_3 Quart$$
$$+\beta_4 CNTRYofDMC + \beta_5 IndName + \beta_6 GovDisc + \beta_7 Esglvl + \varepsilon_t$$

$$SymErr_t = \beta_0 + \beta 1RelBEstNISTDDEV_t + \beta_2 NI_g + \beta_3 Quart$$
$$+\beta_4 CNTRYofDMC + \beta_5 IndName + \beta_6 GovDisc + \beta_7 Esglvl + \varepsilon_t$$

5.3. Results Summary

The different models have mixed outcomes but do point out few important effects. Analysing the result obtained from the models using the unprocessed measure of error [3], it can be noticed that dispersion is still relevant, and its coefficient is not much altered, after addition of different variables. [4]. This result subsist also in the model using the processed (SymErr) and relative measure of error and dispersion.

Speed of revenue growth seem to have an effect on the error made by mean analysts' forecasts in both the version of the model. Nonetheless its sign is inverted in respect to the one hypothesised previously, showing a positive correlation. This fact may be explained by the empirically driven counter-hypothesis that companies growing slower are easier to predict due to the fact that their revenues are more stable over time, and forecasters can draw their information also from past performances.
Size, in particular quartile of enterprise value, seem to be positively relevant only for bigger companies, meaning that forecasters tend to be relatively less accurate on the biggest companies in the market. Although this may be explained by the difference in size of revenues and the possible bias arising from it, it's possible to notice that the same effect persist also in the model with the standardized measure of error, in this case also being in the third quartile seem to be relevant for forecast accuracy, but still not as much as for the relatively biggest companies.

Governance disclosure does not result to be relevant, but general proxy for information disclosure does thus supporting the hypothesis that when general information on a company are more publicly disclosed the mean analysts' forecast better aggregates market public information improving in

[3]For regressions results, see Table A.1 - A.2 - A.3 in appendix A.
[4]The coefficient of dispersion varies from a value of 1.82 in the simplest model to 1.69 in the first model with auxiliary variables, down to a value of 1.61. Such differences may be highly important if this values were used for forecasting purposes, but can be considered a minimal variation if considered the principal explanatory objective of the model.

forecast accuracy.
Dummies representing differences among countries and sector show mixed results depending on the case. Although the models using the gross error do not seem to underline much the effect of industry and macroeconomic effects, results change significantly when the error is standardized for revenues size an is thus more reliable for explanatory studies. In this case, is possible to notice that almost every industry does have a different significant effect on the accuracy of mean analyst' forecasts, relative to the other industries.
Countries do not homogeneously show significance, but some of them seem to be significantly correlated with higher or lower error. These outcomes only strengthen the assumptions and hypothesis previously made, underlining that in some countries macroeconomic differences may affect forecasting precision.

In general, these result seem to be consistent with the hypothesis formerly stated. The error made by mean analysts' forecast seem to be correlated with the factor studied in this model.
The set of informations directly stemming from the results of this study could be highly relevant, and should be considered when taking into account the validity of mean analysts' forecast as an aggregated mean of forecasts.
The next step will be analysing whether these informations could be implemented in a simple model to improve analysts forecast accuracy.

Chapter 6

Applied Models

In this chapter, the models previously used to explain and verify the significance of the factor formerly assumed to be relevant in affecting the accuracy of mean analysts' forecasts, undergo slight modifications, and their new results are used to try and improve the prior forecasts.

6.1. Model Description

The fundamental idea behind the application of these explanatory models into actual improving forecasting, is that with simple regressions is possible to give an estimate of the error mean forecasts will make, and use it to adjust the already existing forecasts.

First the the error is calculated through different methods, so that later the different error estimations can be compared, and the best one can be picked and used to produce an ultimate forecast.
The measure of error used for these kind of estimation is the $PosGroErr$ which is better fit for estimation purposes since it is not complicate and it does not undergo particular mathematical transformation, making it easier to work with.
The error is estimated through the sole use of the dispersion of the forecasters, as it is the variable that among the others scales best, and is supposed to be more relevant. The other variables do not seem to be fit for these error estimation as they tend to alter too much the error estimates, due to the fact that their coefficient is calculated on a large set of observation which vary in scale, where controlling for firm size does not seem to be enough in order to limit size effect.

The different regressions used, closely mirror the previous models in the explanatory section but with the difference that the intercept is excluded

from the linear model. If added, the intercept would always take into account a base error, which is sometimes too large given the fact that since in these models the error is not standardized, different firms of different size cannot be addressed with using the same basic error.
While these problem appears with the coefficient of the intercept, since it is one and cannot be changed with respect to the different firms, the coefficient of the standard deviation does not produce the same bias, simply because standard deviation scales with firms size.

After having calculated an estimation for the error, the next step is to either add back or subtract this error to the already existing mean forecasts. Thorough this procedure, the forecast is adjusted for possible variation in the mean estimate due to the dispersion and uncertainty present among forecasters.
Currently, there is unfortunately no precise way to predict whether this error should be added or subtracted. The problem resides in the fact that *PosGroErr* is always a positive measure, while the bias on the forecast may be either positive or negative, as a prediction can be bigger or smaller than the actual value. In addition the error measure cannot be changed to shift sign or take in to account vector positivity or negativity, and the positive characteristic of it is needed in order to properly sum all of the different magnitude reflecting the error made by forecasts, as previously explained in the section relating measure of errors.

Finally when the error is estimated and the forecast is adjusted, the unmodified mean analysts' and the model implemented, forecasts are compared through the use of a simple T test, which verify whether the error made by the new adjusted forecasts are consistently and significantly lower than the one made by the unadjusted predictions.

6.2. Estimating dispersion coefficients

In order to have a better outlook of comparable coefficients all the different linear model from the explanatory section are tested, after the exclusion of the intercept.

1. $PosGroErr_t = \beta_1 BEstNISTDDEV_t + \varepsilon_t$
2. $PosGroErr_t = \beta_1 BEstNISTDDEV_t + \beta_2 NI_g + \beta_3 Quart + \varepsilon_t$
3. $PosGroErr_t = \beta_1 BEstNISTDDEV_t + \beta_2 NI_g + \beta_3 Quart$
$+\beta_3 CNTRYofDMC + \beta_4 IndName + \varepsilon_t$

4. $PosGroErr_t = \beta_1 BEstNISTDDEV_t + \beta_2 NI_g + \beta_3 Quart$
$+\beta_4 CNTRYofDMC + \beta_5 IndName + \beta_6 GovDisc + \beta_7 Esglvl + \varepsilon_t$

The four different linear models come out with different estimation of the dispersion coefficient, which are grouped in the table below.

Table 6.1: σ Coefficients

Lm 1	Lm 2	Lm 3	Lm 4
1.9427	1.6896	1.6751	1.6008

It's easy to see that the more variable are added to the simple model the more the dispersion coefficient lowers in value. Nonetheless, left out for the initial change when there are completely no other variable other than the raw dispersion, the coefficient does not seem to vary much, landing at a lowest value of 1.6. It can be noted that the correlation here measured is significantly more than linear.

6.3. Adjusting the forecast

Now that the error can be estimated, the coefficients calculated are implemented into the initial mean analysts forecasts. At this point the error obtained for each single observation, is either added or subtracted from the original unadjusted measure, depending on which of the two new values better performs in respect to the original value.

For each of the four different model previously proposed, there is an adjusted forecasts, and each of them is compared to the unadjusted prediction through the use of a two sided T- test, which tests the hypothesis that the error made by the two forecasts is not significantly different, an that therefore there is no improvement in the new forecast.
If the T-statistic value is negative, it means that the first value tested is lower than the other, therefore the adjusted measure is put first, and if the T value results in a negative figure it signifies that the new adjusted forecast performs better than the unadjusted one.
The p-value returns the estimated probability of rejecting the null hypothesis of a study question when that hypothesis is true, in our case refers to the probability of rejecting the hypothesis that there is not a clear difference

between the new adjusted forecast and the former one, when in reality that is true.

Table 6.2: T-Test

	Lm 1	Lm 2	Lm 3	Lm4
t.stat	-3.0169	-3.3597	-3.3756	3.4469
p-value	0.00255	0.0007807	0.000737	0.0005675
Mean Error	117.4534	115.2037	115.0984	114.6240

From the results, is possible to notice how the new adjusted measure actually performs better than the unadjusted forecasts. The T-statistics values are all negative and relatively high, while the p-value are all considerably under 1%, meaning that this test shows an accuracy higher than 99% percent. In addition, if compared to the mean error of the original forecast (137.1037), all of the new estimates perform significantly better.

Among all of the new models, the fourth seems to perform better than the others. The last model, is inclusive of all the proxies and variables assumed to be significant, in this way the effect of the standard deviation is slightly reduced since the other factors which add explanatory power to the regression, subtract it from the dispersion coefficient. With this method, the last coefficient for the dispersion (1.6008) seem to better fit the actual data, improving the quality of the forecasts.

6.4. Testing the forecasting model

At this point the next step will regard testing the model on data which are not included in the initial sample from which the standard deviation coefficient is drawn, having therefore the simulation of a real forecast.
This forecasting experiment is carried out with the same proceeding used in the previous section, with the only difference that one of the year is left out from the initial sample, and than the dispersion coefficient is estimated, with the most complete and efficient model, from the data regarding four years ahead of the year removed, and on the same sample of companies; in the end the last year values are predicted as in the former section.

Table 6.3: T-Test on applied forecasts

	Lm 4
t.stat	-2.0758
p-value	0.03794
Mean Errors	126.9717 - 153.0606

The results outline again a significant improvement in respect to the initial model with a confidence interval higher than 95%. The error mean made by the adjusted forecast is again significantly lower than the one made by the unmodified mean analysts forecasts (153.0606).
These outcomes only strengthen the original hypothesis of a strong correlation between the dispersion of analyst on accuracy of the mean forecasts, and hint to a possible application of an estimation of this dispersion coefficient, which could be effective in the real income forecasting sector.

6.5. Comparing with the Random Walk Forecasts

Having tested the implemented model on the forecasting accuracy, this new method and its precision are now compared to the actual existing and highly considered "Random Walk Forecasting Models". The RWF models basic idea, is that stocks, or in our case revenues, follow unpredictable patterns and cannot be foreseen using past data, therefore the best forecast for the future is just the current present value.

In the next two sided T-tests, the RWF (or more simply the previous year net income values) is first compared to the mean analysts' forecasts and than to the adjusted mean analysts' forecast, using the data where one year is taken out and the coefficients are evaluated on the past four years, so basically the test is carried out on only the one last years observations, and with the methodology of a real forecasting situation.

Table 6.4

	Mean Analysts F. - RWF
t.stat	-5.6676
p-value	1.495e-08
Mean Errors	153.0606 - 241.1423

Table 6.5

	Adj. Mean Analysts' F. - RWF
t.stat	-7.5826
p-value	1.495e-08
Mean Errors	126.9717 - 241.1423

The two T-tests both significantly outline the difference in the accuracy made by the forecasters and the random walk model. Looking at the figures for mean errors made by all the forecasts, the analysts' significantly beat the random walk model, comparatively making almost just half of the error made by the random walk. This results just improves when the analysts' forecasts are adjusted with the dispersion coefficient, reducing additionally the error volume.

6.6. Limits of the model

The model previously introduced, although showing how different factors can affect the accuracy of mean analysts' forecasts, has a considerable limit in its forecasting capabilities.
As stated earlier, the *PosGroErr* is just a positive measure of the error, and after its estimation must *either be added or subtracted* from the current forecast, since sometimes the predictions are short and more pessimistic in respect to the real value, and at times they are more optimistic and are greater than the actual future net income.
The problem in reality with future forecasts, is that there is not a precise way to guess whether this error must be added or subtracted to the current mean analysts' forecasts in order to get its value nearer, and more precise in respect to the future real figures of the net income.

Possible solution to the limits

A possible solution to this problem is to analyse the correlation between some particular factor and the fact that forecasts are too optimistic or pessimistic. In this way, a system to decide whether to add or subtract the estimated error from the unadjusted forecasts, could be implemented on the basis of empirical past prevalence of whether forecasts were too high or too low according to some variable.

A new dummy variable is created, representing the fact that forecasts in respect of the real net income value of a certain year, were either too low if the dummy equals to zero, or too high if it equals to one. A regression is run analysing the effect of the previously analysed variables on this new factor.[1]

$$Binary = \beta_0 + \beta_1 RelBEstNISTDDEV + \beta_2 Quart + \beta_3 GovnDisc + \\ \beta_4 EsgLvl + EV + \beta_5 Nig + \beta_6 IndName + \beta_7 CNTRYofDMC$$

Econometric analysis shows how the probability of making an *optimistic* forecast, and therefore higher than the real future value, is affected by some variables. In particular, belonging to an industry type or another seems to be highly significant in the matter, together with the relative size of the company.

Once accounted for factors as relevant as industry type and relative size, the problem whether to add or subtract the error estimated could be solved by taking into account information from explanatory regression as the one just run, and a case by case study on the past performances of the single entities.
Although this paper does not provide these in depth study on the matter, a complementary research on this factor related to when a forecast is excessively optimistic or pessimistic, together with the results obtained with the models here proposed, could significantly improve the quality of mean analysts' forecasts, providing a real improvement in the revenue forecasting sector.

[1]Results are shown in table A.4

Chapter 7

Conclusions

7.1. General Results

Focussing on the study of net income mean analysts' forecasts relative accuracy, the empirical evidence provided by this paper, finds that there are various significant factors affecting the precision of this particular forecast of net income values.
These empirical results, strengthen the former hypotheses put forward by different research studies, concerning the factors which affect the presence and the relative magnitude of the error made by mean analysts' forecasts. In particular introducing empirically a new important element of study, the dispersion of forecasts in the mean value, on which previous research is missing or lack of significant empirical data.
Dispersion, estimated by calculating the standard deviation on each observation of mean forecasts, is tested with the use of linear models, and is found to be significantly correlated with the size of the relative error made by each prediction.

With the findings obtained by the explanatory regressions, a model of forecasting is built using the estimated coefficient of the dispersion on the error, to forecast the effective mistake made by the mean analysts' forecasts on the real future value of a firm's net income.
The model seems to work properly and its results are significant, providing a consistent improvement in respect to the normal and non modified values of mean analysts' forecasts, which are in this sample, already found to be superior to Random Walk Forecasting.
Nonetheless the forecasting model presents a major problem which hampers its applicability on the current market.

The predictions improvement are based on the addition or subtraction of an

estimated error to the current value of the forecasts, unfortunately this choice of arithmetic sum sign cannot be currently carried out with a good degree of certainty.
Although this may look like a limit, the problem could be solved by a joint study on two factors:

- A case by case analysis of the past forecasting errors made on each firm, analysing whether the analysts tend to be optimistic or pessimistic on their forecasts for a given company in the past years.
- An econometric study on the general factors affecting whether forecast are higher or lower than the actual net income value of the predicted period. In this paper, industry type and company size are already found to be significantly correlated with the fact that forecasts are either too high or too low.[1]

7.2. Applicability of this study

The main contributions of this paper regard the empirical results obtained on the significance of the forecasts dispersion, and of the already previously studied factors.
Results underline an important correlation between size, macroeconomics fundamentals, speed of revenues growth, industry type, information disclosure level and forecasts dispersion, with the accuracy of mean analysts' predictions on net income.

The informations stemming from this findings may have various use in the actual forecasting sector.
Investors and, in general everyone making use of the mean analysts' forecasts future predictions on net income, can now pay more attention to whether the forecasts their are looking at are in a country, or are made for an industry, where forecasts tend to make larger errors. They may add into their analysis, observations regarding the variables here studied, creating their own degree of reliability to a forecast and improving their understanding of it.

The new model introduced for adjusting forecasting, although not being ready for a professional application, is efficient in its objective, and it may hold the key for the development of a new system for better forecasting reliability.

[1] See page - Possible solution to the limits.

Further research could be focussed on whether the error estimated with the forecast dispersion should be added or subtracted from the mean analysts' forecast, basing the choice on some signal from other possible related factors, as showed in this paper already, or on a case by case study.

Based on the data provided, evidence hint on the importance of establishing a degree of reliability for analysts' predictions, the empirical results suggest that this degree of reliability is obtainable and might be built basing its fundamentals on the factors here explained, of which assumption are supported by other researches, focussing mainly on the dispersion of the mean analysts' forecast.

Appendix A

Regression Tables

Table A.1

	Dependent variable:	
	PosGroErr	SymErr
	(1)	(2)
BEstNISTDDEV[1]	1.819***	
	(0.032)	
RelBEstNISTDDEV[2]		$2.429e{-}04$ ***
		(0.00005)
Constant	70.921***	0.361***
	(4.646)	(0.003)
Observations	28,168	28,168
Residual Std. Error (df = 28166)	754.748	0.542
F Statistic (df = 1; 28166)	3,204.953***	23.769***
Note:	*p<0.1; **p<0.05; ***p<0.01	

1. BEstNISTDDEV referst to: "Bloomberg Estimate Net Income Standard Deviation"

2. RelBEstNISTDDEV refers to: "Relative Bloomber Estimate Net Income Standard Deviation"

Table A.2

	Dependent variable:		
	PosGroErr		
	(1)	(2)	(3)
Constant	8.954	−74.430	−92.057
	(12.846)	(81.634)	(118.604)
BEstNISTDDEV	1.690***	1.675***	1.601***
	(0.033)	(0.034)	(0.041)
QuartQ2	4.590	14.810	4.745
	(15.860)	(16.186)	(25.335)
QuartQ3	23.032	28.539*	7.645
	(15.264)	(16.033)	(24.670)
QuartQ4	166.962***	166.372***	96.634***
	(15.112)	(16.300)	(25.400)
Nig	0.533***	0.537***	0.564***
	(0.099)	(0.098)	(0.110)
GovnDisc			−0.498
			(0.873)
EsgLvl			3.614***
			(0.629)
CNTRYofDMCAU		148.966*	110.895
		(82.019)	(114.108)
CNTRYofDMCCH		160.223*	99.879
		(85.780)	(120.206)
CNTRYofDMCDE		180.696**	155.974
		(83.027)	(117.113)
CNTRYofDMCFR		178.514**	138.115
		(82.745)	(117.469)
CNTRYofDMCGB		326.050***	320.735***
		(81.359)	(114.207)
CNTRYofDMCHK		192.551**	196.806*
		(83.665)	(116.928)
CNTRYofDMCIT		158.150*	136.424
		(87.541)	(123.514)
CNTRYofDMCNL		178.354*	151.390
		(91.199)	(128.430)
CNTRYofDMCUS		140.999*	124.310
		(79.458)	(111.846)
IndNameConsumer Goods		−48.936**	−22.982
		(20.136)	(25.780)
IndNameConsumer Services		−66.892***	−35.934
		(19.642)	(25.061)
IndNameFinancials		−67.458***	−40.386
		(20.211)	(26.588)
IndNameIndustrials		−67.029***	−54.556**
		(18.259)	(23.430)
IndNameTelecommunications		198.280***	282.132***
		(33.541)	(40.304)
IndNameUtilities		−25.194	−56.518*
		(27.376)	(34.250)
Observations	28,167	28,167	20,499
F Statistic	705.085***	46.766***	33.451***

Note: *p<0.1; **p<0.05; ***p<0.01

*Only variables showing significance are presented in the table.

Table A.3

	Dependent variable:		
	SymErr		
	(1)	(2)	(3)
Constant	0.423***	0.336***	0.412***
	(0.009)	(0.057)	(0.075)
RelBEstNISTDDEV	0.0002***	0.0002***	0.003***
	(0.00005)	(0.00005)	(0.0003)
QuartQ2	−0.013	−0.005	−0.006
	(0.011)	(0.011)	(0.016)
QuartQ3	−0.068***	−0.068***	−0.083***
	(0.011)	(0.011)	(0.016)
QuartQ4	−0.110***	−0.126***	−0.148***
	(0.011)	(0.011)	(0.016)
Nig	−0.0003***	−0.0003***	−0.0002***
	(0.0001)	(0.0001)	(0.0001)
GovnDisc			−0.001***
			(0.001)
EsgLvl			0.001***
			(0.0004)
Constant	0.423***	0.336***	0.412***
	(0.009)	(0.057)	(0.075)
CNTRYofDMCAR		−0.009	−0.051
		(0.120)	(0.126)
CNTRYofDMCAT		0.107	0.139
		(0.070)	(0.087)
CNTRYofDMCAU		0.283***	0.256***
		(0.057)	(0.072)
CNTRYofDMCBE		0.258***	0.167*
		(0.065)	(0.086)
CNTRYofDMCBM		0.440***	0.499***
		(0.068)	(0.086)
CNTRYofDMCBR		0.170***	0.083
		(0.060)	(0.075)
CNTRYofDMCBS		0.166	0.135
		(0.221)	(0.221)
CNTRYofDMCCA		0.295***	0.283***
		(0.057)	(0.073)
CNTRYofDMCCH		0.182***	0.182**
		(0.060)	(0.076)
CNTRYofDMCCL		0.024	−0.009
		(0.080)	(0.093)
CNTRYofDMCCN		0.021	−0.013
		(0.057)	(0.071)
CNTRYofDMCCO		0.031	−0.012
		(0.110)	(0.117)
CNTRYofDMCCY		0.235*	−0.296*
		(0.120)	(0.164)
CNTRYofDMCCZ		0.030	−0.175
		(0.135)	(0.222)
CNTRYofDMCDE		0.143**	0.129*
		(0.058)	(0.074)
CNTRYofDMCDK		0.165**	0.104
		(0.070)	(0.090)
CNTRYofDMCEG		0.109	1.019***
		(0.090)	(0.222)

Table A.3 Contd.

	Dependent variable:		
	SymErr		
	(1)	(2)	(3)
CNTRYofDMCES		0.236***	0.192**
		(0.062)	(0.079)
CNTRYofDMCFI		0.279***	0.216***
		(0.063)	(0.080)
CNTRYofDMCFR		0.199***	0.185**
		(0.058)	(0.074)
CNTRYofDMCGB		0.306***	0.289***
		(0.057)	(0.072)
CNTRYofDMCGI		0.785***	
		(0.161)	
CNTRYofDMCGR		0.332***	0.362***
		(0.070)	(0.090)
CNTRYofDMCHK		0.157***	0.134*
		(0.059)	(0.074)
CNTRYofDMCHR		−0.163	−0.217
		(0.221)	(0.222)
CNTRYofDMCHU		0.219	−0.052
		(0.161)	(0.221)
CNTRYofDMCID		0.017	−0.077
		(0.065)	(0.081)
CNTRYofDMCIE		0.367***	0.326***
		(0.072)	(0.087)
CNTRYofDMCIL		0.322***	0.094
		(0.083)	(0.102)
CNTRYofDMCIM		−0.006	0.209
		(0.161)	(0.221)
CNTRYofDMCIN		0.116**	0.086
		(0.058)	(0.073)
CNTRYofDMCIS		−0.067	
		(0.221)	
CNTRYofDMCIT		0.299***	0.290***
		(0.061)	(0.078)
CNTRYofDMCJE		−0.127	−0.138
		(0.221)	(0.221)
CNTRYofDMCJO		0.029	−0.002
		(0.221)	(0.221)
CNTRYofDMCJP		0.048	0.004
		(0.056)	(0.071)
CNTRYofDMCKR		0.199***	0.167*
		(0.077)	(0.097)
CNTRYofDMCKW		0.062	0.004
		(0.162)	(0.165)
CNTRYofDMCKY		0.030	
		(0.221)	
CNTRYofDMCLU		0.272***	0.450***
		(0.083)	(0.111)
CNTRYofDMCMA		−0.051	−0.107
		(0.221)	(0.222)
CNTRYofDMCMO		−0.021	−0.036
		(0.161)	(0.164)

Table A.3 Contd.

	Dependent variable:		
	SymErr		
	(1)	(2)	(3)
CNTRYofDMCMT		−0.064	
		(0.221)	
CNTRYofDMCMX		0.061	0.034
		(0.068)	(0.082)
CNTRYofDMCMY		0.023	0.020
		(0.059)	(0.077)
CNTRYofDMCNL		0.442***	0.359***
		(0.064)	(0.081)
CNTRYofDMCNO		0.375***	0.325***
		(0.064)	(0.079)
CNTRYofDMCNZ		0.214***	0.054
		(0.065)	(0.091)
CNTRYofDMCOM		−0.004	−0.188
		(0.120)	(0.164)
CNTRYofDMCPA		−0.017	−0.052
		(0.221)	(0.221)
CNTRYofDMCPE		0.085	0.386*
		(0.123)	(0.221)
CNTRYofDMCPH		−0.021	−0.023
		(0.070)	(0.083)
CNTRYofDMCPK		−0.166*	−0.226*
		(0.090)	(0.117)
CNTRYofDMCPL		0.105	0.125
		(0.068)	(0.090)
CNTRYofDMCPR		−0.023	−0.061
		(0.221)	(0.221)
CNTRYofDMCPT		0.080	0.051
		(0.079)	(0.111)
CNTRYofDMCQA		−0.048	−0.082
		(0.120)	(0.140)
CNTRYofDMCRO		−0.139	−0.158
		(0.221)	(0.221)
CNTRYofDMCRU		0.337***	0.293***
		(0.064)	(0.079)
CNTRYofDMCSA		0.041	0.046
		(0.072)	(0.094)
CNTRYofDMCSE		0.201***	0.172**
		(0.061)	(0.078)
CNTRYofDMCSG		0.187***	0.099
		(0.062)	(0.079)
CNTRYofDMCSI		−0.015	−0.074
		(0.221)	(0.221)
CNTRYofDMCTH		0.074	0.003
		(0.060)	(0.080)
CNTRYofDMCTR		0.097	0.023
		(0.062)	(0.081)
CNTRYofDMCTW		−0.081	−0.123*
		(0.058)	(0.073)
CNTRYofDMCUA		0.209	0.178
		(0.161)	(0.164)
CNTRYofDMCUS		0.256***	0.229***
		(0.056)	(0.070)

Table A.3 Contd.

	Dependent variable:		
	SymErr		
	(1)	(2)	(3)
CNTRYofDMCVN		−0.159	−0.069
		(0.098)	(0.221)
CNTRYofDMCZA		0.187***	0.145*
		(0.063)	(0.077)
IndNameConsumer Goods		−0.156***	−0.142***
		(0.014)	(0.016)
IndNameConsumer Services		−0.141***	−0.126***
		(0.014)	(0.016)
IndNameFinancials		−0.038***	−0.012
		(0.014)	(0.017)
IndNameHealth Care		−0.182***	−0.160***
		(0.016)	(0.019)
IndNameIndustrials		−0.141***	−0.136***
		(0.013)	(0.015)
IndNameOilAndGas		0.045***	0.033*
		(0.016)	(0.019)
IndNameTechnology		−0.022	0.036**
		(0.015)	(0.018)
IndNameTelecommunications		−0.073***	−0.053**
		(0.023)	(0.025)
IndNameUtilities		−0.102***	−0.097***
		(0.019)	(0.022)
Observations	28,167	28,167	20,499
F Statistic	44.722***	26.427***	23.719***
Note:			*p<0.1; **p<0.05; ***p<0.01

Table A.4

	Dependent variable:
	Bin
Constant	0.556***
	(0.070)
RelBEstNISTDDEV	0.0002
	(0.0003)
QuartQ2	0.007
	(0.015)
QuartQ3	−0.030**
	(0.015)
QuartQ4	−0.043***
	(0.015)
GovnDisc	0.001**
	(0.001)
EsgLvl	0.0001
	(0.0004)
EV	0.000
	(0.000)
Nig	−0.00004
	(0.0001)
IndNameConsumer Goods	−0.051***
	(0.015)
IndNameConsumer Services	−0.052***
	(0.015)
IndNameFinancials	−0.167***
	(0.016)
IndNameHealth Care	−0.001
	(0.017)
IndNameIndustrials	−0.072***
	(0.014)
IndNameOilAndGas	−0.068***
	(0.018)
IndNameTechnology	0.076***
	(0.017)
IndNameTelecommunications	−0.010
	(0.024)
IndNameUtilities	−0.087***
	(0.020)
CNTRYofDMCAR	−0.209*
	(0.119)
CNTRYofDMCAU	0.163**
	(0.068)
CNTRYofDMCCH	0.119*
	(0.071)
CNTRYofDMCCZ	−0.390*
	(0.209)
CNTRYofDMCHK	−0.154**
	(0.069)
CNTRYofDMCJO	0.404*
	(0.209)
CNTRYofDMCKR	0.292***
	(0.091)
CNTRYofDMCNL	0.141*
	(0.076)
CNTRYofDMCPR	−0.448**
	(0.209)
Observations	20,499

Note: *p<0.1; **p<0.05; ***p<0.01

* Variables related to countries, not showing significance, are not shown.

References

[1] Slack N., Chambers, S., Jhonston R., (2004). Operations Management, fourth Edition, Prentice Hall

[2] Wallace, T.F., Stahl, R.A., (2008) Sales & Operations planning The How-To Hand-book, Third Edition. TF Wallace & Co.

[3] Givoly, Dan, and Josef Lakonishok. "The quality of analysts' forecasts of earnings." *Financial Analysts Journal* 40.5 (1984): 40-47.

[4] Fried, Dov, and Dan Givoly. "Financial analysts' forecasts of earnings: A better surrogate for market expectations." *Journal of Accounting and Economics* 4.2 (1982): 85-107.

[5] Bradshaw, Mark T., et al. A re-examination of analysts' superiority over time-series forecasts. Working paper, Boston College, The Ohio State University, and University of Arkansas, 2011.

[6] Kim, Oliver, Steve C. Lim, and Kenneth W. Shaw. "The Use of Forecast Revision in Reducing Built-In Biases in Mean Analyst Forecasts." (1998).

[7] Dispersion in Analysts' Earnings Forecasts as a Measure of Uncertainty Orie E. BarronPamela S. Stuerke** Journal of Accounting, Auditing & Finance, 07/1998, Volume 13, Fascicolo 3 *Journal of Accounting, Auditing & Finance* July 1998 13: 245-270

[8] Imhoff Jr, Eugene A., and Gerald J. Lobo. "The effect of ex ante earnings uncertainty on earnings response coefficients." *Accounting Review* (1992): 427-439.

[9] Daley, Lane A., David W. Senkow, and Robert L. Vigeland. "Analysts' forecasts, earnings variability, and option pricing: Empirical evidence." *Accounting Review* (1988): 563-585.

[10] Abarbanell, Jeffery, and Reuven Lehavy. "Biased forecasts or biased earnings? The role of reported earnings in explaining apparent bias and

over/underreaction in analysts' earnings forecasts." *Journal of Accounting and Economics* 36.1 (2003): 105-146.

[11] Hope, Ole _ Kristian. "Disclosure practices, enforcement of accounting standards, and analysts' forecast accuracy: An international study." *Journal of accounting research* 41.2 (2003): 235-272.

[12] Bhat, Gauri, Ole_ Kristian Hope, and Tony Kang. "Does corporate governance transparency affect the accuracy of analyst forecasts?." *Accounting & Finance* 46.5 (2006): 715-732.

[13] Buettner, Thiess, and Bjoern Kauder. "Revenue Forecasting Practices: Differences across Countries and Consequences for Forecasting Performance*." *Fiscal Studies* 31.3 (2010): 313-340.

[14] Kwon, Sung S. "Financial analysts' forecast accuracy and dispersion: High-tech versus low-tech stocks." *Review of Quantitative Finance and Accounting* 19.1 (2002): 65-91.

[15] Das, Somnath, Carolyn B. Levine, and Ki Sivaramakrishnan. "Earnings predictability and bias in analysts' earnings forecasts." *Accounting Review* (1998): 277-294.

[16] Lundholm, Russell, Sarah McVay, and Taylor Randall. "Forecasting sales: A model and some evidence from the retail industry." Unpublished working paper. *University of British Columbia and University of Washington* (2010).

[17] Lang, Mark H., and Russell J. Lundholm. "Corporate disclosure policy and analyst behavior." *Accounting review* (1996): 467-492.

[18] Armstrong, J. Scott, and Fred Collopy. "Error measures for generalizing about forecasting methods: Empirical comparisons." *International journal of forecasting* 8.1 (1992): 69-80.

[19] Makridakis, Spyros, and Michele Hibon. "The M3-Competition: results, conclusions and implications." *International journal of forecasting* 16.4 (2000): 451-476.

[20] Bloomberg official website - http://www.bloomberg.com/