ISSN 2616-5902 Vol. 5, Issue 10: 8-14, DOI: 10.25236/AJBM.2023.051002

# Research on the factors affecting the earning per share (EPS) forecast accuracy

**Zixuan Gu**[1]

[1]*Department of Business and Management, Wuhan Business University, Wuhan, China*

***Abstract:*** *Securities analysts play a vital role in alleviating the asymmetric information in the capital market, through collecting, processing and transmitting information via their expertise. As profit forecasting is gradually beneficial for investors to make decision, increasing number of researchers concentrate on profit forecasting to find out the factors affecting the earning per share (EPS) forecast accuracy. In common, there are three main aspects to influence the EPS forecast accuracy: personal ability of analysts, difficulty of the task and broker size. Since there are still lots of controversies on the influence of broker size on forecast error, this essay will focus on examining broke size by univariate analysis and multivariate analysis with the sample of U.S. automobile industry from 2001 to 2009. This essay hypothesizes that analysts employed by large brokers produce more accurate earnings forecasts than analysts employed by smaller brokers. The results of my article can support my hypothesis.*

***Keywords:*** *EPS, forecast accuracy, broke size*

## 1. Introduction

According to Yi, Wang and Li (2016), profit forecasting is not only an essential part of the analyst's research report, but also a significant reference for investors to make decision.[1] Clement (1999) conducts empirical tests on the factors influencing the analyst's earnings forecast accuracy and summarizes the results into three aspects contenting the personal ability of analysts; difficulty of the task and broker size.[2] And the subsequent studies usually base on the conclusion of Clement. This essay will focus on broke size as there are still controversies.

## 2. Literature Review

Numerous literatures are consistent with the prediction of the impact of personal ability and task difficulty on forecast error. However, there are still lots of controversies on the influence of broker size on forecast accuracy. Clement (1999) discover analysts in large brokerage firms are expected to produce higher EPS forecast accuracy. However, Lin and McNichols (1998) conclude the forecast accuracy of analysts working in large brokerage house is lower.[3] Besides, Clarke (2007) uses the sample, including the data of 216 all-star analysts' job-hopping, to study the changes in analysts' behavior before and after the job changes[4], and does not find a significant change in the analyst's earnings forecast accuracy.

## 3. Hypothesis

Hypothesis: analysts employed by large brokers produce more accurate earnings forecasts than analysts employed by smaller brokers.

## 4. Data and method

Univariate analysis and multivariate analysis will be included in empirical analysis, both involving mean adjusted variables. In univariate analysis, the forecast error of analysts from large broker firm will be compared with the those in small broker firm. Multivariate analysis will test the correlation between EPS forecast error and factors affecting EPS forecast accuracy.

ISSN 2616-5902 Vol. 5, Issue 10: 8-14, DOI: 10.25236/AJBM.2023.051002

### *4.1. Dependent variable*

The method for measuring forecast error, the proportional mean absolute forecast error (PMAFE), is defined by Clement (1999) to avoid time-series and cross-sectional differences.

$$PMAFEijt = DAFEijt / \overline{AFE}jt$$

where DAFEijt = AFEijt - $\overline{AFE}$jt, and AFEijt stands for the absolute difference between the EPS forecasted by analyst i and the actual EPS of firm j in year t, $\overline{AFE}$jt is the mean absolute forecast error of firm j in year t.

### *4.2. Independent variables*

The basic description of these variables defined by Clement (1999) is introduced as follows.

• FEXPijt - The number of years through year t for which analyst i supplied at least one forecast during the first 11 months of the year for firm j.

• GEXPit - The number of years for which analyst i supplied at least one forecast during the first 11 months of the year through year t.

• NCOMit - The number of companies for which the analyst supplied at least one forecast during the current year. It measures portfolio complexity.

• DTOP10it - A dummy variable set to one if the analyst is employed by a firm ranked in the top 10% during the current year and set to zero otherwise. Brokerage houses are ranked yearly with respect to the number of analysts employed. It proxies for the resources available to the financial analyst. In my sample, top 10% brokerage firms in every year are defined as large brokers while others are small brokers.

• AGEijt - Age (in days) of analyst i's forecast for firm j's earnings at time t.

• IEXPint - The number of years through year t for which analyst i supplied at least one forecast during the first 11 months of the year for industry n.

### *4.3. Multivariate analysis*

Based on the results of Clement (1999), the following equation will be adopted as estimated model to test hypothesis stated above:

$$\begin{gathered} PMAFE_{ijt} = \beta_1 DNTOP10_{ijt} + \beta_2 DFEXP_{ijt} + \beta_3 DAGE_{ijt} + \\ (-) \qquad\qquad (-) \qquad\qquad (+) \\ \beta_4 DNCOM_{ijt} + \beta_5 DIEXP_{int} + \eta_{ijt} \\ (+) \qquad\qquad (-) \end{gathered} \tag{1}$$

where the prefix D added each variable means all independent variables are mean adjusted and predicted signs of coefficients are presented in parentheses below each factor. The negative proportional mean absolute forecast error means that analyst i's forecast error for company j is below the average level in year t and vice versa. The larger the forecast error, the lower the level of forecast accuracy is. Finally, cross-sectional regression analysis by year separately will be run, then the regression put together and the results compared.

In terms of sensitivity analysis, in order to test the robustness of equation (1), I add an additional variable, FREV, which reflects the frequency of EPS forecast revisions. The study of Holden and Stuerke (2008) indicates that FREV has positive contribution of forecast error as special financial phenomenon and particular managerial adjustment.[5] Here is the new model:

ISSN 2616-5902 Vol. 5, Issue 10: 8-14, DOI: 10.25236/AJBM.2023.051002

$$PMAFE_{ijt} = \beta_1 DNTOP10_{ijt} + \beta_2 DFEXP_{ijt} + \beta_3 DAGE_{ijt} + \beta_4 DNCOM_{ijt} + \beta_5 DIEXP_{int} + \beta_6 DFREV_{ijt} + \eta_{ijt} \quad (2)$$

(-) (-) (+) (+) (-) (-)

### *4.4. Sample selection*

The data, collected from Institutional Broker Estimate System (I/B/E/S), covers earnings forecasts of automobile industry in the US over the period 2001-2009. Table 1 describes summary statistics of analyst forecasts on I/B/E/S.

*Table 1: Summary statistics of analyst forecasts on I/B/E/S*

| | No. firms | No. analysts | No. brokers | No. forecasts |
|---|---|---|---|---|
| Section A: Initial sample of EPS forecasts | | | | |
| Total | 61 | 388 | 143 | 10517 |
| | | | | |
| Section B: Number of observations deleted due to filters | | | | |
| Imposing forecast horizon between 30 and 365 | 0 | 3 | 1 | 377 |
| Requiring firms followed by at least three analysts | 6 | 16 | 9 | 229 |
| | | | | |
| Section C: Final sample after applying all the filters | | | | |
| 2001 | 12 | 53 | 111 | 1234 |
| 2002 | 11 | 55 | 108 | 1136 |
| 2003 | 8 | 40 | 72 | 941 |
| 2004 | 13 | 42 | 66 | 1022 |
| 2005 | 10 | 56 | 97 | 983 |
| 2006 | 17 | 59 | 84 | 1103 |
| 2007 | 13 | 55 | 85 | 1036 |
| 2008 | 9 | 66 | 112 | 1272 |
| 2009 | 33 | 63 | 145 | 1184 |
| Total | 55 | 369 | 133 | 9911 |

## 5. Results

### *5.1. Univariate analysis*

*Table 2: Univariate test for variables (t-test)*

| Broker size group | Large | Small |
|---|---|---|
| Broker size range | Top 10% | The other 90% |
| Mean | -0.580 | -0.486 |
| Observations | 1430 | 8481 |
| P (T<=t) one-tail | 0.125 | |
| t Critical one-tail | 1.645 | |

In brief, as showed in Table2, analysts from large-scale brokerage house might not produce more accurate EPS forecast than those in small-scale broker firm, as the p-value is more than 10%. Since the sample size is not large enough, only 1430 forecasts of large-scale broker firms and 8481 forecasts of small-scale broker firm involved, more tests should be done in future.

### *5.2. Descriptive statistics*

Table 3 describes the descriptive statistics for the raw (non-mean adjusted) variables. Most analysts

ISSN 2616-5902 Vol. 5, Issue 10: 8-14, DOI: 10.25236/AJBM.2023.051002

can provide a forecasted EPS rather close to accurate EPS and the others need to improve their EPS forecast accuracy. As for independent variable, broker size, Table 3 indicates the significant difference of broker size among brokers measured by the number of analysts employed.

*Table 3: Descriptive statistics for raw variables*

| | Mean | Median | SD | Min | Max | Quartile1 | Quartile3 |
|---|---|---|---|---|---|---|---|
| Forecast error (%) | 0.106 | 0.008 | 1.093 | 0 | 32.642 | 0.002 | 0.035 |
| Number of analysts employed | 84.937 | 68 | 69.245 | 1 | 442 | 29 | 121 |
| Firm experience (years) | 3.633 | 3 | 2.499 | 0 | 13 | 2 | 5 |
| Forecast age (days) | 184.399 | 168 | 92.338 | 30 | 352 | 92 | 256 |
| Number of companies followed | 14.545 | 14 | 7.210 | 1 | 104 | 11 | 17 |
| Industry experience (years) | 5.568 | 5 | 2.913 | 0 | 14 | 3 | 7 |
| N | 9911 | | | | | | |

### *5.3. Pearson Correlations*

Table 4 provides the correlation coefficients of both dependent variable and independent variables contained in regression. As shown in the previous section, firm experience, forecast age, the number of firms followed by the analyst, industry experience and broker size will influence forecast error. The results suggest that the magnitudes of DAGE is the highest. Thus, it is worth to test the relationship between forecast error and these independent variables. Besides, only DNTOP10 and DFEXP have negative correlation with forecast error.

Specifically, forecast errors are negatively correlated with broker size (DNTOP10),$-0.012$. It points out large-scale brokerage firms may have benefits on improving EPS forecast accuracy in automobile industry to some extent, as the correlation of broker size and forecast error is negatively.

*Table 4: Correlation coefficients between demeaned variables*

| | PMAFE | DNTOP10 | DFEXP | DAGE | DNCOM | DIEXP |
|---|---|---|---|---|---|---|
| PMAFE | 1 | | | | | |
| DNTOP10 | -0.012 | 1 | | | | |
| DFEXP | -0.008 | -0.039 | 1 | | | |
| DAGE | 0.063 | 0.018 | 0.005 | 1 | | |
| DNCOM | 0.020 | 0.136 | 0.137 | -0.005 | 1 | |
| DIEXP | 0.010 | -0.047 | 0.698 | 0.009 | 0.192 | 1 |

### *5.4. Multivariate analysis*

Equation (1) is applied to test my hypothesis. The estimated coefficients and p values for each regression model based on the sample of pooled year 01-09 are demonstrated in Table 5.

Table 5 involves two main statistic norms to judge the regression model: p value and adjusted R-square ($\overline{R^2}$). Panel A presents these five independent variables: DNTOP10, DFEXP, DAGE and DNCOM have significant coefficients and are related regressors to explain the forecast error.

As shown in Panel A, the coefficients of these variables all display the same sign from regression (1) to regression (5). It indicates that the intercorrelations between independent variables are not very significant. As predicted, the results are in accord with the prior researches that forecast errors are negatively correlated with the regressors DNTOP10 and DFEXP, and positively correlated with the regressor DAGE and DNCOM. However, contrary of prediction, the coefficients of DIEXP are unexpectedly keeping positive. Specifically, EPS forecast accuracy could be slightly improved by around 0.11 unit when the analysts work in top 10% of brokerage firms or have unique firm-specific experience for one more year than others. The p value of DNTOP10 and DFEXP are 0.004 and 0.008 respectively, showing the coefficients are all significant at 1% level.

ISSN 2616-5902 Vol. 5, Issue 10: 8-14, DOI: 10.25236/AJBM.2023.051002

*Table 5: Results of pooled year 01-09 regression and annual regressions*

| Panel A. Results of pooled year 01-09 regressions by using forward selection | | | | | |
|---|---|---|---|---|---|
| Dependent variable: PMAFE | | | | | |
| Regressor | (1) | (2) | (3) | (4) | (5) |
| DNTOP10 | -0.093 | -0.096 | -0.106 | -0.133 | -0.126 |
| P>\|t\| | 0.011** | 0.011** | 0.005*** | 0.003*** | 0.004*** |
| DFEXP | | -0.036 | -0.035 | -0.049 | -0.118 |
| P>\|t\| | | 0.146 | 0.126 | 0.049** | 0.008*** |
| DAGE | | | 0.361 | 0.362 | 0.361 |
| P>\|t\| | | | 0*** | 0*** | 0*** |
| DNCOM | | | | 0.138 | 0.124 |
| P>\|t\| | | | | 0.002*** | 0.006*** |
| DIEXP | | | | | 0.131 |
| P>\|t\| | | | | | 0.021** |
| Summary Statistics | | | | | |
| Adg.R2 | 0 | 0 | 0.004 | 0.004 | 0.005 |
| F-statistic | 6.5 | 3.28 | 18.97 | 15.56 | 14.5 |
| N | 9911 | 9911 | 9911 | 9911 | 9911 |

| Panel B. Annual regressions | | | | | | | |
|---|---|---|---|---|---|---|---|
| Year | DNTOP10 | DFEXP | DAGE | DNCOM | DIEXP | Adg.R2 | N |
| 2001 | -0.111 | -0.12 | 0.393 | 0.034 | 0.209 | 0.052 | 1234 |
| P>\|t\| | 0.044** | 0.176 | 0*** | 0.327 | 0.126 | | |
| 2002 | -0.081 | -0.057 | 0.129 | 0.02 | 0.096 | 0.06 | 1136 |
| P>\|t\| | 0*** | 0.012** | 0*** | 0.408 | 0.011** | | |
| 2003 | -0.069 | -0.079 | 0.174 | -0.079 | 0.164 | 0.111 | 941 |
| P>\|t\| | 0.031** | 0*** | 0*** | 0*** | 0*** | | |
| 2004 | -0.001 | -0.213 | 0.219 | 0.025 | 0.218 | 0.04 | 1022 |
| P>\|t\| | 0.996 | 0*** | 0*** | 0.557 | 0*** | | |
| 2005 | -0.422 | -0.575 | 0.386 | 0.438 | 0.329 | 0.018 | 983 |
| P>\|t\| | 0.003*** | 0.001*** | 0.018** | 0.058* | 0.147 | | |
| 2006 | -0.979 | -0.45 | 0.724 | 1.034 | 0.139 | 0.002 | 1103 |
| P>\|t\| | 0.008*** | 0.237 | 0.066* | 0.066* | 0.787 | | |
| 2007 | -0.145 | 0.03 | 0.3 | 0.139 | -0.076 | 0.001 | 1036 |
| P>\|t\| | 0.056* | 0.22 | 0*** | 0.113 | 0.026** | | |
| 2008 | 0.171 | 0.06 | 0.467 | -0.027 | -0.029 | 0.077 | 1272 |
| P>\|t\| | 0.006*** | 0.163 | 0*** | 0.337 | 0.548 | | |
| 2009 | 0.17 | 0.11 | 0.568 | -0.78 | -0.061 | 0.214 | 1184 |
| P>\|t\| | 0*** | 0*** | 0*** | 0.05** | 0.042** | | |
| Pre. | - | - | + | + | - | | |
| Sign | | | | | | | |

Panel B demonstrates the results of annual regressions from 2001 to 2009. Among all the five regressors, only DAGE keeps the same sign in the whole sample process whereas the others - DNTOP10, DFEXP, DNCOM and DIEXP present different signs during the whole period.

As the same in my prediction, the coefficient of DAGE is positive and significant at 1% level in every year. Even though DNTOP10 keeps its coefficient being significant all the years expect 2004, the negative coefficient turns to positive at the last two years - 2008 and 2009.

Nevertheless, the coefficients of DFEXP, DNCOM and DIEXP present different signs and are not always significant during the night years. It is a wonder that the directions of all the independent variables are changed except DAGE in year 2008 and 2009. The reason might be the influence of global financial crisis during 2008 and 2009[6], increasing the instability of the regression model (Schipper, 2015). Besides, what is out of my expectation is the coefficient of DIEXP keeps positive in

ISSN 2616-5902 Vol. 5, Issue 10: 8-14, DOI: 10.25236/AJBM.2023.051002

most years rather than negative in my prediction.

DNTOP10, presents negative coefficient all the time as predicted except in the last two years - 2008 and 2009. Besides, it keeps significant in annual years other than year 2004. As DNTOP10 is significant at 1%, its coefficient in 2001-2009 pooled regression model ($-0.126$) indicates forecast errors will decrease 0.126 unit if the analyst is hired by the top 10% broker firm. This result is consistent with the research of Clement (1999), Livnat and Mendenhall (2006) suggests investors to relay more on the analysts working in large-scale broker firm.[7] In all, the findings show that broker size is positive correlated with EPS forecast accuracy of automobile industry in U.S., which is consistent with my hypothesis.

### *5.5. Sensitivity test*

*Table 6: Sensitivity test*

| Panel A. Sensitivity test by comparing in different year period | | | |
|---|---|---|---|
| | Prediction | Year 01--05 | Year 06--09 |
| DNTOP10 | - | -0.113 | -0.149 |
| P>\|t\| | | 0*** | 0.116 |
| DFEXP | - | -0.234 | -0.015 |
| P>\|t\| | | 0*** | 0.825 |
| DAGE | + | 0.25 | 0.517 |
| P>\|t\| | | 0*** | 0*** |
| DNCOM | + | 0.084 | 0.184 |
| P>\|t\| | | 0.003*** | 0.053* |
| DIEXP | - | 0.225 | -0.057 |
| P>\|t\| | | 0*** | 0.496 |
| Adg.R2 | | 0.016 | 0.004 |
| F-statistic | | 11.79 | 7.35 |
| N | | 5316 | 4595 |

| Panel B. Sensitivity test by adding DFREV | | | |
|---|---|---|---|
| | Prediction | (1) | (2) |
| DNTOP10 | - | -0.126 | -0.036 |
| P>\|t\| | | 0.004*** | 0.210 |
| DFEXP | - | -0.118 | -0.061 |
| P>\|t\| | | 0.008*** | 0.124 |
| DAGE | + | 0.361 | 0.372 |
| P>\|t\| | | 0*** | 0*** |
| DNCOM | + | 0.124 | 0.084 |
| P>\|t\| | | 0.006*** | 0.034** |
| DIEXP | - | 0.131 | 0.135 |
| P>\|t\| | | 0.021** | 0.018** |
| DFREV | | | 0.293 |
| P>\|t\| | | | 0*** |
| Adg.$R^2$ | | 0.005 | 0.013 |
| F-statistic | | 14.5 | 13.34 |
| N | | 9911 | 9911 |

The result of Table 6 Panel A shows the coefficients of all the variables in 2001-2005 pooled regression model are significant and their signs are the same as in 2001-2009 pooled regression model. However, most coefficients in 2006-2009 pooled regression model are not significant except that of DAGE and DNCOM, and the sign of DIEXP even turns to negative which is different from that of original model. It means the 2001-2005 pooled regression model can explain sample more accurate as the coefficients are all significant. Surprisingly, its $\overline{R^2}$ (0.016) is far more than $\overline{R^2}$of year 2006-2009 pooled regression (0.004). These results indicate the original regression model is not as constant as in year 2006-2009 than in the first five years of data set, affecting by the extreme economic fluctuation-global financial crisis. In other words, the independent variables are sensitive when the sample year changes.

In the second method, a new variable, DFREV, representing the frequency of revision, which is

ISSN 2616-5902 Vol. 5, Issue 10: 8-14, DOI: 10.25236/AJBM.2023.051002

discussed to have negative effect on improving forecast accuracy in the study of Holden and Stuerke (2008), is involved. The results in Panel B demonstrate frequency of revision has great impact on increasing forecast error, as its coefficient is highest (0.293) and significant at 1% level. Besides, the signs, figures and p-values of coefficients of all independent variables remain nearly unchanged except that of DFEXP and DNTOP10 when DFREV added. It shows that frequency of revision may have influence on firm-specific experience and broker size. Collectively, DFEXP and DNTOP10 are more sensitive than other variables as they are easier to be effected in sensitivity test. In later study, more researches on the influence of firm-specific experience and broker size on forecast accuracy are suggested to be done.

## 6. Conclusion

This essay aims at examining if the EPS forecast accuracy of analysts hired by large-scale (top 10%) broker firm is higher than that of small-scale brokerage house for U.S. automobile companies. It focuses on the annual earnings forecasts produced by approximately 400 analysts from 143 brokers for nearly 60 U.S. automobile companies from 2001 to 2009.

I reach an overall conclusion that analysts employed by large brokers produce more accurate earnings forecasts than analysts employed by smaller brokers, as the coefficient of DNTOP10 is positive and significant in multivariate analysis. Besides, the analysts' forecast performance in automobile firms will be improved if more information related to the observed companies can be obtained. According to the 2001-2009 pooled regression, broker size and firm-specific experience both have positive influence on improving forecast accuracy. Further studies are suggested to focus on the influence of other forecast error related factors on broker size and firm-specific experience, as these two factors are most sensitive in sensitivity test.

## Acknowledgement

Project Name: 2021Wuhan Business University scientific project (2021KY025).

## References

*[1] Yi Zhihong, Wang Xinbin and Li Ying (2016). Broker Size and Analysts' Earnings Forecast Accuracy--Empirical Evidence Based on Analyst Turnover. Journal of Shanxi University of Finance and Economics, 38 (1), 36-43.*

*[2] Clement M., (1999). Analyst forecast accuracy: do ability, resources, and portfolio complexity matter? Journal of Accounting and Economics 27, 285−303.*

*[3] Lin H. and M. F. McNichols, (1998). Underwriting Relationships, Analysts' Earnings Forecasts and Investment Recommendations, Journal of Accounting and Economics, 25, 101-127.*

*[4] Clarke J., Khorana A., Patel A. and Rau P.R., 2007. The impact of all-star analyst job changes on their coverage choices and investment banking deal flow. Journal of Financial Economics, 84(3), 713-737.*

*[5] Holden C. and Stuerke P. (2008). The Frequency of Financial Analysts' Forecast Revisions: Theory and Evidence. Social Science Electronic Publishing, 35(7-8), 860–888.*

*[6] Schipper K. 2015. Commentary on analysts' forecasts. Accounting Horizons 5, 105–121.*

*[7] Livnat J., and Mendenhall R., (2006). Comparing the Post–Earnings Announcement Drift for Surprises Calculated from Analyst and Time Series Forecasts Journal of Accounting Research 44, 177–205.*