ELSEVIER Journal of Accounting and Economics 27 (1999) 285–303

JOURNAL OF Accounting & Economics

www.elsevier.com/locate/econbase

Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter?☆

Michael B. Clement*

Department of Accounting, College and Graduate School of Business Administration, CBA 4M.202, University of Texas at Austin, Austin, TX 78712-1172, USA

Received 30 May 1997; received in revised form 22 January 1999

Abstract

Prior studies have identified systematic and time persistent differences in analysts' earnings forecast accuracy, but have not explained why the differences exist. Using the I/B/E/S Detail History database, this study finds that forecast accuracy is positively associated with analysts' experience (a surrogate for analyst ability and skill) and employer size (a surrogate for resources available), and negatively associated with the number of firms and industries followed by the analyst (measures of task complexity). The results suggest that analysts' characteristics may be useful in predicting differences in forecasting performance, and that market expectations studies may be

☆I thank my dissertation committee of William Beaver, George Foster, and Peter Reiss for helpful discussions and comments. I am also grateful to S.P. Kothari (editor) and Russell Lundholm (the referee). Helpful comments were provided by Rowland Atiase, Larry Brown, Bryan Cloyd, Ellen Engel, Robert Freeman, Paquita Friday, Ross Jennings, Bill Kinney, Lisa Koonce, Maureen McNichols, Praveen Sinha, Senyo Tse, Beverly Walther, Richard Willis and workshop participants at Baruch College, Boston College, University of Georgia, Stanford University, the University of Texas, and Texas A&M University. I gratefully acknowledge the contribution of I/B/E/S International Inc. for providing earnings per share forecast data, available through the International Brokers Estimate System. These data have been provided as part of a broad academic program to encourage earnings expectation research.

* Tel.: (512) 471-5619; fax: (512) 471-3904.

E-mail address: mclement@mail.utexas.edu (M.B. Clement)

PII: S0165-4101(99)00013-0

improved by modeling these characteristics.

JEL classification: G24; J44

Keywords: Financial analysts; Forecast accuracy; Performance evaluation; Task analysis; Experts

1. Introduction

Prior research provides mixed results about whether systematic differences in financial analysts' forecast accuracy exist. While a number of early studies do not identify significant differences in forecast accuracy, more recent studies document systematic differences. In particular, early studies by Richards (1976), Brown and Rozeff (1980), O'Brien (1987), Coggin and Hunter (1989), O'Brien (1990) and Butler and Lang (1991) do not find systematic differences in forecast accuracy. More recent studies including Stickel (1992) and Sinha, Brown and Das (1997), however, document systematic differences in analysts' forecast accuracy. This paper builds on the more recent studies by investigating analysts' characteristics that are potentially associated with forecast accuracy.

Research that examines the determinants of forecast accuracy is important for at least two reasons. First, given analysts' role as expert financial intermediaries (that is, experts in using and interpreting financial information), accounting researchers are interested in knowing whether differences in forecast accuracy exist and the factors that contribute to the differences. Prior studies may not have identified differences in forecast accuracy because they did not control for factors that contribute to the differences. Second, an understanding of the determinants of forecast accuracy is important to accounting researchers who use analysts' earnings forecasts as a proxy for the capital markets' expectation of earnings. If analysts' forecast accuracy differs predictably, and if the capital markets recognize these differences, then more accurate proxies for earnings expectations can be created by applying larger weights to the forecasts of more accurate analysts.

Stickel (1992) finds that Institutional Investor All-American analysts' forecasts are more accurate than NonAll-Americans' forecasts. He finds that All-Americans are more accurate forecasters, forecast more frequently, and their upward forecast revisions have a greater impact on stock prices than do the upward revisions of NonAll-Americans. The capital market response to upward revisions documented by Stickel suggests that capital market participants believe differences exist in forecast accuracy.

Sinha, Brown and Das (1997) (hereafter SBD) identify systematic differences in forecast accuracy among a larger body of analysts. Using I/B/E/S data from

1984–1990, they extend O'Brien's work. Noting that forecast errors tend to be serially correlated, they include controls in their model for forecast recency. Their results suggest that systematic ex-post differences exist in analysts forecast accuracy. SBD also perform ex-ante tests of forecast accuracy. They find that analysts classified as superior in one period continue to be classified as superior in later periods, but analysts classified as inferior in one period do not continue to be classified as inferior in later periods.

While Stickel and SBD identify differences in forecast accuracy, they do not explain why the differences exist. The objective of these studies is to determine whether systematic differences exist. It is possible that differences in ability, resources and portfolio complexity contribute to differences in forecast accuracy. The objective of the current study is to identify analysts' characteristics that are associated with the differences.

I use cross-sectional regression analysis to investigate whether analysts' forecast errors are associated with the combination of their experience, the number of firms and industries followed, and employer size. The joint analysis of the ability, task complexity and resources is important since the three are potentially correlated. In addition, I also control for firm-year variation in both the dependent and independent variables. This is important because the predictability of a firm's earnings can vary across time (see Clement, 1998).

I find that forecast accuracy is positively associated with general and firm-specific forecasting experience and employer size, and negatively associated with the number of firms and industries followed. Given the magnitudes of the regression coefficients and the distributions of the regression variables, my overall conclusion is that employer size alone, or the other variables taken in combination, may provide information for capital market participants to predict economically meaningful differences in forecast accuracy. For example, analysts who work for large employers have average forecast errors that are 7.7% smaller than other analysts.

Concurrent studies by Mikhail, Walther and Willis (1997) (hereafter MWW) and Jacob, Lys and Neale (1997) (hereafter JLN) also examine factors that may contribute to analysts' forecast accuracy. MWW use a time series approach and find forecast accuracy and forecasting experience to be related. Like Stickel, MWW's results may not be generalizable because MWW examine the performance of a small group of analysts. In order to estimate time series parameters, they restrict their sample to analysts who have 32 continuous quarters of forecasts for a company. This requirement causes 97% of the potential sample to be excluded. The incremental contributions of my study over MWW are that I use a larger sample and a broader set of variables (i.e., resources and task complexity) to investigate differences in forecast accuracy.

JLN examine the contributions of experience and brokerage house variables on analyst forecasting attributes including forecast accuracy, frequency and horizon. Consistent with my study they find employer size is associated with forecast accuracy. They also find that forecast accuracy is positively associated with the brokerage house's degree of industry specialization and negatively associated with brokerage house turnover. Unlike my study and MWW, JLN do not find evidence that forecast accuracy improves with experience. The contrast in results may be due to differences in research design. For example, I control for firm-year fixed effects while JLN do not. On the other hand, JLN include variables not included in my study such as forecast frequency.

Section 2 outlines the hypotheses. Section 3 describes the measurement of variables. Section 4 reports the empirical results and Section 5 concludes.

2. Hypotheses

2.1. Hypotheses for ability

While ability is unobservable, we may observe indicators of ability through the way the analyst labor market functions. The analyst labor market is assumed to function as a tournament in which stronger performers continue, while the weaker performers are forced out of the profession.[1] According to Milgrom and Roberts (1992), the only performance information needed or used in a tournament is the relative, ordinal information about who did better. This is consistent with the methods used by brokerage firms and clients to evaluate analysts. For example, when money managers are asked to rank analysts for inclusion on the All-American Team, they are asked to rank analysts against other analysts in the same industry.[2] If these ordinal rankings affect analysts' tenure, then we should expect more experienced analysts to be more accurate. (It is possible, however, that an analyst may survive long term due to luck rather than skill.) Another reason that analysts with greater experience may supply more accurate forecasts is that their general skills and knowledge improve with time. For example, analysts may become better at analyzing financial statements or recognizing economic trends as they gain experience.

[1] It is possible that the stronger performers are promoted to other positions (e.g., mutual fund managers). This argument is inconsistent, however, with reports in the financial press. For example, the June 19, 1997 *Wall Street Journal* article "Sixteen All-Stars Excel for Fifth Time", acknowledges the performance of 16 security analysts who placed on the All-Star team 5 years in a row. The story notes, "At big firms, compensation of $300,000 to $600,000 for top analysts is common and $1 million is possible." Similarly, the June 20, 1996 *Wall Street Journal* article "A few analysts shine year after year", notes that 80 of 440 All-Stars were on the team 3 of the previous 4 years. The implication of these stories is that strong performers have an incentive to stay in the profession.

[2] Forecast accuracy is only one of several criteria used to rank analysts. Other criteria include stock picking ability, quality of written research reports and client service. The study assumes analysts' performance is positively correlated along these dimensions. That is, a good forecaster is likely to be a good stock picker who writes good reports and provides good client service.

In addition to acquiring general skills, analysts also acquire firm-specific skills over time (i.e., skills specific to the firms they follow). For example, these skills might include a better understanding of the idiosyncrasies of a particular firm's reporting practices or the analyst might establish better relationships with insiders and thereby gain better access to managers' private information.

Given these observations, the hypotheses for ability are as follows:

H_1: Holding resources (employer size) and portfolio complexity (number of firms and industries followed) constant, forecast accuracy increases with forecasting experience.

H_2: Holding resources (employer size) and portfolio complexity (number of firms and industries followed) constant, forecast accuracy increases with firm-specific forecasting experience.

2.2. *Hypotheses for portfolio complexity*

An analyst's portfolio complexity is also likely to be associated with forecast accuracy. The numbers of firms and industries followed are used as proxies for portfolio complexity based on the assumption that it is more difficult to follow a larger set of firms and industries. In other words, larger portfolios allow the analyst to devote less attention to each individual firm. Furthermore, Lees (1981) concludes that there are economies of scale to following firms in a particular industry. In a setting where analysts face identical effort constraints and diminishing returns to effort, the magnitude of forecast errors will increase as analysts increase the number of companies and industries followed. These observations lead to the following hypotheses regarding portfolio complexity:

H_3: Holding ability and resources constant, forecast accuracy decreases with the number of firms followed.

H_4: Holding ability and resources constant, forecast accuracy decreases with the number of industries followed.

2.3. *Hypotheses for resources*

Large brokers may provide superior resources that contribute to the forecast accuracy of their analysts. For example, analysts employed by large brokerage firms may have access to better data sets and administrative support.[3] They may

[3] Commenting on the role resources play in performance, Andrew Melnick, director of global equity research at Merrill Lynch said, "Today the senior analyst needs strong associates [assistants], and part of their job is to make the star look like a star." Melnick further noted that the growth in Merrill's research staff has come not so much from senior analysts as from these associates. "They are critical to maintaining our research franchise and the franchise of the superstars", he added. (See Institutional Investor Magazine, 1996, p. 52.)

also have better access to the private information of managers at the firms they follow. Stickel (1995) provides evidence that capital market participants respond more to the buy and sell recommendations of analysts employed by large brokerage houses relative to other analysts. He attributes this difference to the larger firms' more advanced distribution networks, which allow large firms to better disseminate their analysts' recommendations into the capital markets. If larger brokerage firms provide superior resources for distributing buy and sell recommendations, they may also provide superior resources for analysts in performing their research. These observations suggest that when ability and portfolio complexity are held constant, analysts employed by large brokerage firms supply more accurate forecasts than other analysts. This hypothesis is stated formally as:

H_5: Holding ability and portfolio complexity constant, analysts who are employed by large brokers supply more accurate forecasts than other analysts.

2.4. Correlation of ability, complexity and resources

The joint analysis of ability, complexity and resources is important because the three factors are likely to be correlated. Omission of the correlated variables from the analysis would result in biased estimates of the relationship between performance and the individual variables. The following discussion illustrates some settings in which the variables would be correlated.

Ability and complexity would be correlated if the most able analysts follow the most complex portfolios. Ability would be correlated with resources if high ability analysts work in environments that provide greater resources. Similarly, the two factors would be correlated if large firms hire more able analysts. Finally, complexity would be correlated with resources if analysts who work for large employers follow a smaller number of firms and industries than do analysts who work for other employers.

The preceding discussion suggests that ability, portfolio complexity, and resources should be jointly analyzed to prevent biased estimates of the relationships between each of the factors and forecast accuracy. These observations also raise the possibility that the factors affecting forecast accuracy may be jointly determined. For example, the most able analysts may choose, or be assigned, to follow the most complex portfolios. Or the most able analysts may have the greatest labor market opportunities and may choose to work for employers who provide the greatest resources. If the factors that determine forecast accuracy are jointly determined, then the regression coefficients estimated in the following analyses may suffer from simultaneity bias. To address this issue a model of brokerage firm resource allocation is required. Development of such a model is beyond the scope of this study and is a potential area for future research.

3. Data and sample selection

3.1. Measurement of variables

3.1.1. Dependent variable

I measure performance by comparing the analyst's absolute forecast error to the average absolute forecast error of other analysts following the same stock during the same time period. The performance measure used here is the proportional mean absolute forecast error (PMAFE) and is calculated as:

$$PMAFE_{ijt} = DAFE_{ijt}/\overline{AFE_{jt}}, \tag{1}$$

where $DAFE_{ijt} = AFE_{ijt} - \overline{AFE_{jt}}$, AFE_{ijt} is the absolute forecast error for analyst i's forecast of firm j for year t, and $\overline{AFE_{jt}}$ is the mean absolute forecast error for firm j for year t.

Clement (1998) finds that controlling for firm-year effects increases the likelihood of identifying systematic differences in analysts' forecast accuracy relative to a model that controls for firm fixed effects and year fixed effects. Firm-year effects result from factors that make a firm's earnings easier or more difficult to predict in some years than others. Examples of events that may give rise to firm-year effects are voluntary management disclosures, mergers and strikes. $PMAFE$ controls for firm-year effects by subtracting from the absolute forecast error its related firm-year mean.[4] Clement (1998) also shows that large EPS firms have greater variation in their DAFE's than do small EPS firms, and that deflating $DAFE$ by $\overline{AFE}$ reduces heteroscedasticity. The model in Section 4.2 below was also estimated using $DAFE_{ijt}$ as the dependent variable and the inferences remained unchanged.

$PMAFE$ is calculated using a 30-day minimum forecast horizon and can be interpreted as analyst i's fractional forecast error relative to the average of the analysts' absolute forecast errors for firm j at year t. Negative values of $PMAFE$ represent better than average performance while positive values of $PMAFE$ represent worse than average performance.

3.1.2. Proxies for ability

H_1 predicts analysts' forecasts become more accurate with increases in forecasting experience. The proxy used for (general) forecasting experience is calculated as:

$GEXP_{it}$ = number of years for which analyst i supplied at least one forecast during the first 11 months of the year through year t.

[4] Analyst i is included in the calculation of $\overline{AFE_{jt}}$ since he too is affected by the firm-year effect.

The 11-month requirement is imposed based on the assumption that active analysts would supply forecasts for the firms they follow during this period. An analyst who only releases forecasts more than 12 months prior to period end is not likely to be following companies very closely. Similarly, an analyst who only releases forecasts less than 30 days prior to period end is more likely to be mimicking the forecasts of other analysts rather than following the companies himself. The 11-month requirement is also imposed on the other variables in the model.[5]

H_2 predicts that forecast accuracy increases with firm-specific experience. The proxy for the number of years of firm-specific experience is calculated using the analyst's experience following a particular firm as follows:

$FEXP_{ijt}$ = 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.

3.1.3. Proxies for portfolio complexity

H_3 and H_4 state that when ability and resources are held constant, forecast accuracy decreases with the number of companies and industries followed. The number of firms followed is determined by counting the number of ticker symbols for which the analyst made forecasts in a given year as follows:

$NCOS_{it}$ = number of firms for which analyst i supplied at least one forecast during the first 11 months of year t.

Two-digit SICs are used as the proxy for industries and the variable used for the number of industries followed is calculated as:

$NSIC2_{it}$ = number of two-digit SICs for which analyst i supplied at least one forecast during the first 11 months of year t.

3.1.4. Proxy for available resources

H_5 predicts that analysts employed by large brokers will supply more accurate forecasts than other analysts will. Like Stickel (1995), I assume there are two levels of resources provided by employers: one level provided by large employers and another level provided by all other employers. He uses a dummy variable to identify analysts employed by large brokers. I use a similar approach and calculate employer size as:

$DTOP10_{it}$ = a dummy variable set to 1 if analyst i is employed by a firm in the top size decile during year t, and set to 0 otherwise. Size deciles are calculated based on the number of analysts employed in year t.

[5] Since the data are left censored, analysts appearing in the 1983 sample are excluded from the calculation of all variables, including AFE_{jt}.

This specification assumes there is an elite set of brokerage firms that provides a higher level of resources than other firms. These elite firms are generally referred to as "bulge bracket firms" in the investment-banking community.

3.1.5. Control variables

Results from prior research suggest that forecast age, firm effects, and year effects should be controlled for when evaluating differences in analysts' forecasting ability. *PMAFE* controls for firm and year effects by adjusting absolute forecast errors by their related firm-year means. The model's independent variables are also adjusted by their related firm-year means to properly control for firm-year effects.[6] Consequently, the model takes the form $Y_{ijt} - \bar{Y}_{jt} = (X_{ijt} - \bar{X}_{jt})\beta$. (See Greene (1991) for a more detailed description of using mean adjusted data to control for fixed effects.)

I control for forecast age by including the forecast's age in days as an independent variable. Forecast age is measured as:

AGE_{ijt} = age (in days) of analyst *i*'s forecast for firm *j*'s earnings at time *t*.

3.2. Description of the I/B/E/S data set and sample selection method

The data for this study were obtained from the Institutional Broker Estimate System (I/B/E/S) Detail History tape. Table 1 reports summary statistics for the I/B/E/S data set. The data set covers the period from 1983 to 1994 and contains over 1 million forecasts for the annual earnings of more than 9,500 companies made by over 7,500 analysts. Analyst codes are used to identify analysts on the academic tape. These codes remain with an analyst as he moves from broker to broker. Some entries to the data set are forecasts supplied by individual analysts and others are supplied by teams of analysts. The analyst codes on the usual academic tapes do not distinguish between individuals and teams, and this makes it impossible to identify teams of analysts. The I/B/E/S broker translation file is used to eliminate teams from the sample.[7] Panel A of Table 1 reports sample statistics after eliminating analyst teams. This sample is referred to as the initial sample.

The sample is restricted to forecasts supplied during the first 11 months of the fiscal year in an attempt to capture the forecasts of active analysts. Once forecasts from the first 11 months are identified, forecasts with a minimum

[6] The approach used here is analogous to estimating a model using firm-year dummy variables to control for firm-year effects.

[7] Teams were removed from the sample by excluding analysts who had a slash (/), an ampersand (&), the word "and", or the word "group" in their names from the sample.

forecast horizon of 30 days are selected. Panel B of Table 1 reports descriptive statistics after imposing the 11-month requirement and minimum forecast horizon. This sample is referred to as the intermediate sample. Relative to the initial sample, the intermediate sample has 4% (246/6468) fewer analysts and 73% (890,429/1,219,979) fewer forecasts.

Since the I/B/E/S data set is left censored, it is not possible to tell how much experience analysts have prior to the first year of available data. To mitigate this problem, analysts who appear in the data set in the initial year (1983) are excluded from the sample. Forecasts from 1984 are also excluded from the sample since there would be little variation in the experience variables that year

Table 1
I/B/E/S data set summary statistics

Year	No. analysts	No. forecasts	No. brokers	No. firms
Panel A. Initial sample of annual earnings forecasts				
1983	1,548	59,147	101	2,939
1984	1,810	78,428	126	3,444
1985	2,023	98,693	141	3,829
1986	2,199	104,035	145	3,861
1987	2,398	114,295	162	4,296
1988	2,405	114,362	185	4,435
1989	2,408	105,662	189	4,238
1990	2,329	108,711	198	4,077
1991	2,191	107,727	190	4,047
1992	2,332	109,871	205	4,294
1993	2,712	113,973	221	4,725
1994	2,824	105,075	233	5,249
Total	6,468	1,219,979	357	9,707
Panel B. Sample after imposing the 11-month and 30-day minimum forecast horizon requirements				
1983	1,403	19,344	89	2,306
1984	1,659	21,658	112	2,672
1985	1,913	25,021	136	2,971
1986	2,075	27,885	141	3,239
1987	2,270	30,144	152	3,733
1988	2,263	28,895	171	3,794
1989	2,310	28,872	185	3,669
1990	2,214	28,720	190	3,655
1991	2,109	27,573	188	3,570
1992	2,212	28,724	189	3,797
1993	2,548	31,902	215	4,185
1994	2,584	30,812	223	4,274
Total	6,222	329,550	349	8,324

Table 1. (continued)

Panel C. Final sample after controlling for left censoring

Year	No. analysts	No. forecasts	No. brokers	No. firms
1985	962	10,596	125	2,499
1986	1,255	14,712	135	2,808
1987	1,518	17,846	145	3,307
1988	1,597	18,518	165	3,446
1989	1,695	19,706	182	3,371
1990	1,669	20,176	184	3,434
1991	1,600	19,720	175	3,375
1992	1,740	20,905	180	3,609
1993	2,088	23,931	208	3,960
1994	2,144	23,529	215	4,084
Total	4,758	189,639	324	7,540

Notes: This table reconciles the full I/B/E/S data set to the sample used in the study. Panel A shows descriptive statistics for annual forecasts in the data set after removing teams. Panel B shows descriptive statistics for the last forecast for each analyst-firm pair during the first 11 months of the fiscal year (i.e., a minimum forecast horizon of 30 days). Panel C is the final sample. It shows descriptive statistics after removing 1983 analysts and 1984 forecasts to control for left censoring. *No. analysts* represents the number of analysts in the sample. *Total analysts* represents the unique number of analysts in the sample. *No. forecasts* represents the number of annual earnings forecasts in the sample. *No. brokers* represents the number of brokers (analyst employers) in the sample. *Total brokers* represents the unique number of brokers in the sample. *No. firms* represents the number of firms in the sample. *Total firms* represents the unique number of firms in the sample.

(i.e., the experience variables can take on values of 0 or 1 in 1984). Panel C reports descriptive statistics after imposing these restrictions. This is the final sample. Relative to the intermediate sample, the final sample has 24% (1,646/6,222) fewer analysts and 42% (139,911/329,550) fewer forecasts. Panels B and C suggest that most of the reduction takes place during early years of the sample. For example, relative to the initial sample, the final sample has 42% (59,439/140,817) fewer forecasts in 1985–1989, but only 27% (39,347/147,731) fewer forecasts for 1990–1994.

The study also uses the I/B/E/S Actuals file. The Actuals file contains the company ticker, a measure indicator, a periodicity indicator, the fiscal period end date, the actual value and the report date. I/B/E/S adjusts the Actuals file so that all forecasts and reported earnings are stated on the same basis.[8] Most analysts provide forecasts of income from continuing operations and

[8] Some companies are followed on a primary EPS basis and others are followed on a fully diluted EPS basis. I/B/E/S makes any necessary adjustments to forecasts so they are on the same basis that the firm is normally followed.

accordingly I/B/E/S backs out non-operating items (e.g., restructuring charges) when reporting actual earnings.

4. Results

4.1. Descriptive univariate statistics

Table 2 shows correlation coefficients and distributions of the regression variables. Panel A shows that forecast errors are negatively correlated with

Table 2
Correlation coefficients and distributions of regression variables

$N = 189{,}639$

	PMAFE	*DAGE*	*DGEXP*	*DFEXP*	*DNCOS*	*DNSIC2*	*DNTOP10*
Panel A. Pearson correlation coefficients							
PMAFE	1.000						
DAGE	0.347	1.000					
DGEXP	− 0.019	0.002	1.000				
DFEXP	− 0.022	0.004	0.633	1.000			
DNCOS	0.009	− 0.030	0.191	0.114	1.000		
DNSIC2	0.031	0.128	0.093	0.034	0.511	1.000	
DNTOP1	− 0.055	− 0.270	0.081	0.057	− 0.101	− 0.186	1.000
Panel B. Distributions of regression variables							
Q1	− 0.51	− 55.00	− 1.04	− 0.78	− 6.90	− 1.50	− 0.36
Median	− 0.08	− 12.50	0.00	0.00	− 1.25	− 0.33	0.00
Q3	0.27	41.06	1.00	0.57	4.47	1.00	0.40

Notes: Panel A shows the Pearson Correlation Coefficients for the regression variables. All non-zero correlations in Panel A are statistically significant at the 1% level. Panel B shows the distributions of the regression variables. Means are not reported in Panel B because by construction the means are zero. *PMAFE* = difference between the absolute forecast error for analyst *i* for firm *j* at time *t* and the mean absolute forecast error for firm *j* at time *t* scaled by the mean absolute forecast error for firm *j* at time *t*. *DAGE* = the age of analyst *i*'s forecast minus the age of the average analyst's forecast following firm *j* at time *t*, where age is the age of the forecast in days at the minimum forecast horizon date. *DGEXP* = the number of years (including time *t*) that analyst *i* appeared in the data set minus the average number of years analysts following firm *j* at time *t* appeared in the data set. *DFEXP* = the number of years (including time *t*) that analyst *i* supplied a forecast for firm *j* minus the average number of years analysts following firm *j* had supplied forecasts. *DNCOS* = the number of companies followed by analyst *i* at time *t* minus the average number of companies followed by an analyst following firm *j* at time *t*. *DNSIC2* = the number of 2 digit SICs followed by analyst *i* at time *t* minus the average number of two-digit SICs followed by an analyst following firm *j* at time *t*. *DNTOP10* = dummy variable with value of 1 if analyst works at a top decile size firm (and 0 otherwise) minus the mean value of dummy variable for analysts following firm *j* at time *t*.

general and firm-specific forecasting experience (surrogates for ability), positively correlated with the number of firms and industries followed (surrogates for portfolio complexity), and negatively correlated with employer size (a surrogate for resources). Each of these correlations has the predicted sign, but as previously noted, the joint tests in the following section are more appropriate for investigating the hypotheses. As suggested in the preceding section, the surrogates for ability, portfolio complexity, and resources are correlated. The ability surrogate is positively correlated with both the portfolio complexity and resource surrogates, while the portfolio complexity and resource surrogates are negatively correlated.

Panel B of Table 2 shows distributions of the regression variables. Means are not reported because by construction the means are zero. The negative median values for *PMAFE*, *DAGE*, *DNCOS*, and *DNSIC*2 indicate that the distributions for these variables are skewed. The variation reported in the table is consistent with a setting in which portfolio complexity and resources potentially affect accuracy.

4.2. Results from tests of hypotheses

Recall from Section 3.1.5 that both the dependent and independent variables are mean adjusted to control for firm-year effects. Eq. (2) below is used to test the study's hypotheses:

$$\begin{aligned} PMAFE_{ijt} = {} & \beta_1 DAGE_{ijt} + \underset{(-)}{\beta_2}\, DGEXP_{ijt} + \underset{(-)}{\beta_3}\, DFEXP_{ijt} \\ & + \underset{(+)}{\beta_4}\, DNCOS_{ijt} + \underset{(+)}{\beta_5}\, DNSIC2_{ijt} \\ & + \underset{(-)}{\beta_6}\, DNTOP10_{ijt} + \eta_{ijt}, \end{aligned} \tag{2}$$

where all variables are firm-year mean adjusted (the *D* preceding each variable stands for differenced) and the predicted signs are included below each coefficient.

Note that the original hypotheses are stated in terms of forecast accuracy, yet the empirical tests investigate analysts' forecast errors. Small forecast errors represent a high level of accuracy. Also note that the regression equation does not require an intercept since means are subtracted from all variables.

Results from the estimation of Eq. (2) are reported in Table 3.[9] Panel A shows the results for the individual annual regressions. As predicted, the coefficient for

[9] The model was checked for serial correlation by regressing residuals on (firm-year) lagged residuals. The results suggested serial correlation was present, but the effect was small (the serial correlation coefficient was less than 0.02). The model was then re-estimated after correcting for serial correlation and there was no change in conclusions.

$DNTOP10$ is negative and significant each year. Similarly, the coefficients for $DFEXP$ are predominately negative (as predicted), though not always significant. The coefficients for $DNCOS$ and $DNSIC2$ are predominately positive (as predicted), though not always significant. The coefficients for $DGEXP$ are

Table 3
Tests of ability, resource and portfolio variables

$$PMAFE_{ijt} = \beta_1 DAGE_{ijt} + \beta_2 DGEXP_{ijt} + \beta_3 DFEXP_{ijt} + \beta_4 DNCOS_{ijt} + \beta_5 DNSIC2_{ijt} + \beta_6 DNTOP10_{ijt} + v_{ijt}$$

Panel A. Annual regressions

Year	*DAGE*	*DGEXP*	*DFEXP*	*DNCOS*	*DNSIC2*	*DNTOP10*	Adj. R^2	*N*
1985	0.36	2.94	0.94	0.13	0.23	− 8.04	0.13	10,596
	38.80	1.48	0.41	1.97	0.84	4.99		
1986	0.20	4.33	− 3.08	− 0.04	0.00	− 8.93	0.04	14,712
	23.25	3.27	2.06	0.64	0.00	5.67		
1987	0.29	2.82	− 1.45	0.06	0.36	− 2.74	0.09	17,846
	41.58	3.32	1.50	1.17	1.45	1.96		
1988	0.27	0.23	− 0.14	0.03	0.60	− 4.30	0.07	18,518
	37.39	0.36	0.19	0.67	2.45	3.06		
1989	0.29	− 0.09	− 1.01	0.09	− 0.14	− 9.14	0.11	19,706
	47.30	0.20	1.76	1.88	0.70	7.29		
1990	0.37	− 0.06	− 1.47	− 0.03	0.37	− 10.87	0.15	20,176
	59.02	0.15	3.10	0.73	1.84	8.54		
1991	0.44	− 1.58	0.11	0.06	0.41	− 9.32	0.18	19,720
	63.44	4.38	0.25	1.31	1.94	6.92		
1992	0.38	0.36	− 1.07	0.08	0.78	− 7.51	0.13	20,905
	56.14	1.17	2.73	1.71	3.50	5.97		
1993	0.42	− 0.91	− 0.87	0.06	0.65	− 7.51	0.18	23,931
	71.32	3.54	2.47	1.37	3.08	6.36		
1994	0.40	− 0.44	− 1.07	0.11	0.48	− 6.00	0.13	23,529
	59.31	1.81	3.26	2.31	2.18	5.01		
Pred. sign		−	−	+	+	−		

Panel B. Pooled regression coefficients

Year	*DAGE*	*DGEXP*	*DFEXP*	*DNCOS*	*DNSIC2*	*DNTOP10*	Adj. R^2	*N*
Pooled10	0.35	− 0.40	− 0.94	0.07	0.41	− 7.72	0.12	189,639
	160.66	3.35	6.00	4.50	5.64	18.44		
Pooled5	0.40	− 0.53	− 0.90	0.06	0.55	− 8.05	0.16	108,261
	138.80	4.06	5.26	2.93	5.72	14.45		

Table 3 (continued)

Panel C. Time series average regression coefficients

Year	*DAGE*	*DGEXP*	*DFEXP*	*DNCOS*	*DNSIC2*	*DNTOP10*	*n*
TS10	0.34	0.76	− 0.91	0.06	0.37	− 7.44	10
	14.21	1.25	2.67	3.18	4.14	9.53	
TS5	0.40	− 0.53	− 0.87	0.06	0.54	− 8.24	5
	32.56	1.56	3.30	2.36	7.01	9.78	

Notes: Coefficient values are reported as percentages with t-statistics below. *PMAFE* = difference between the absolute forecast error for analyst i for firm j at time t and the mean absolute forecast error for firm j at time t scaled by the mean absolute forecast error for firm j at time t. *DAGE* = the age of analyst i's forecast minus the age of the average analyst's forecast following firm j at time t, where age is the age of the forecast in days at the minimum forecast horizon date. *DGEXP* = the number of years (including time t) that analyst i appeared in the data set minus the average number of years analysts following firm j at time t appeared in the data set. *DFEXP* = the number of years (including time t) that analyst i supplied a forecast for firm j minus the average number of years analysts following firm j had supplied forecasts. *DNCOS* = the number of companies followed by analyst i at time t minus the average number of companies followed by an analyst following firm j at time t. *DNSIC2* = the number of 2 digit SICs followed by analyst i at time t minus the average number of two-digit SICs followed by an analyst following firm j at time t. *DNTOP10* = dummy variable with value of 1 if analyst works at a top decile size firm (and 0 otherwise) minus the mean value of dummy variable for analysts following firm j at time t.

Pooled10 represents pooled regression coefficients estimated from the 1985–1994 sample. Pooled5 represents pooled regression coefficients estimated from the 1990–1994 sample. TS10 represents time series average regression coefficients estimated from the 1985–1994 sample. TS5 represents time series average regression coefficients estimated from the 1990–1994 sample.

The t-statistics in Panel C are calculated as $\bar{x}/(s_x/\sqrt{n})$ where the x's are the regression coefficients for the individual years, s_x is the standard deviation of the regression coefficients across the years, and n is the number of years.

positive and significant (not negative as predicted) in the early years and negative in the later years. Note that the model appears to fit the data better in the latter half of the sample than it does in the earlier half of the sample. The average R^2 for the model is 0.09 during 1984–1989 and 0.15 during 1990–1994.[10] Note also that the model's coefficients generally have the predicted signs and are significant during 1990–1994. The model's weaker performance in 1984–1989 may be the result of the sample size reduction required to control for left censoring (see Panels B and C in Table 1). In other words, there is likely greater

[10] The model's R^2 represents the variation in the dependent variable explained after removing firm-year effects (all variables in the model are adjusted by subtracting firm-year means). This is similar to regressing absolute forecast errors on residuals from a regression of absolute forecast errors on firm-year dummies. Since a large source of variation (i.e., firm-year effect) has been removed from the dependent variable, the R^2 is significantly lower than it would be otherwise.

variation in experience and less colinearity between general and firm specific experience in the latter half of the sample.[11]

Panel B reports the results of the pooled regression. Pooled10 represents the regression coefficients when the model is estimated during 1985–1994 and Pooled5 represents the regression coefficients when the model is estimated during 1990–1994. The following discussion is for Pooled10 since the results for the two models are generally consistent. Coefficients for all variables in the pooled regression have the predicted signs and are significant at conventional levels. While the hypotheses make no predictions about *DAGE*, prior research suggests that its coefficient should be positive. The coefficients for *DAGE* are consistently positive and significant. The pooled regression results suggest that *relative* absolute forecast errors increase at the rate of 0.35% per day, or 10.5% per month. The implication for researchers is that careful controls for age are needed when comparing forecasts.

The pooled regression coefficient for *DGEXP* is -0.40%. This suggests that an analyst who has seven years of forecasting experience (the 90th percentile value) will have an expected absolute forecast error that is 2.4% smaller than an analyst who has one year of forecasting experience (the 10th percentile value).

The pooled regression coefficient for *DFEXP* is -0.94% suggesting that an analyst with five years of firm experience (the 90th percentile value) will on average have an absolute forecast error that is 3.8% smaller than an analyst with one year of firm experience (the 10th percentile value). The magnitude of the coefficient appears reasonable relative to MWW who find that forecast accuracy improves by 3% as firm-specific forecast experience doubles. The findings suggest that absolute forecast errors improve faster per year of firm-specific forecasting experience than per year of general forecasting experience.

As predicted, absolute forecast errors increase with the number of firms and industries followed. The pooled regression coefficient for *DNCOS* is 0.07%. An analyst who follows 23 firms (the 90th percentile value) has an expected absolute forecast error that is 1.5% larger than an analyst who follows two firms (the 10th percentile value).

The pooled regression coefficient of 0.41% for *DNSIC2* suggests an analyst who follows eight industries (the 90th percentile value) has an expected forecast error that is 2.9% larger than an analyst who follows one industry (the 10th percentile value). The relative magnitudes of the coefficients for *DNCOS* and *DNSIC2* are intuitively appealing since it seems reasonable that forecast

[11] Note from Table 3 that *GEXP* and *FEXP* are highly correlated (Pearson Correlation Coefficient = 0.63). This condition reduces the model's ability to estimate the unique effects of the two variables. When the model is estimated using *GEXP* as the only experience measure the coefficient is negative and significant in 5 of 10 years compared to 2 of 10 years when both variables are included. This result suggests the model does not completely disentangle the effects of the two variables.

accuracy would deteriorate more with the addition of an industry than with the addition of a firm. The coefficient for *DNSIC2* might also be interpreted as a measure of the benefit of industry specialization. The coefficient tells us the change in forecast accuracy per industry followed holding constant the number of companies followed.

The pooled regression coefficient for *DNTOP10* suggests that, as predicted, analysts who work for large brokers have smaller absolute forecast errors than other analysts. The pooled regression results suggest large broker analysts' errors are 7.7% smaller than other analysts.[12] The magnitude of the coefficient appears reasonable relative to Stickel (1992) who finds that All-American Analysts supply forecasts that are $0.027 more accurate than other analysts. Given that the mean absolute forecast error in the current study is $0.594, this suggests that analysts who work for large firms supply forecasts that are approximately $0.045 more accurate than other analysts. This is consistent with a setting in which analysts at large firms have better resources and more access to managers' private information at the firms they cover. On the other hand, it is possible that large brokers have greater financial resources available to compensate their analysts and therefore employ higher ability analysts. Regardless of which explanation is more likely, this finding provides a possible explanation for Stickel's (1995) finding that the capital markets respond more to the revisions of large broker analysts. The capital markets may respond more to the forecasts of these analysts because large broker analysts are perceived to be more accurate. Combining the results from the tests of the experience and employer size variables yields a prediction that an analyst employed at a large broker, with five years experience would have an expected absolute forecast error that is approximately 10% smaller than a first year analyst employed at a small broker.[13]

Given the large number of observations used to estimate the pooled regression and the considerable variation in the sign and magnitudes in the annual regression coefficients, it is possible that the effects are not as strong as the pooled regression suggests. Time-series averages of the annual regression coefficients are therefore reported in Panel C. TS10 represents the regression coefficients when the model is estimated during 1985–1994 and TS5 represents the regression coefficients when the model is estimated during 1990–1994. Except

[12] The model was estimated using the continuous variable (number of analysts employed by the broker) instead of the dummy variable and the coefficient on the continuous variable was negative and significant. The model was also estimated using both the dummy variable and the continuous variable and both coefficients were negative and significant.

[13] This does not reflect the improvement in forecast accuracy that is associated with firm-specific experience. Forecast accuracy improves by about 1% per year with firm specific-experience. Approximately 5% of the 1994 sample were analysts who were employed by large brokers and had five years of general experience.

for the coefficient for *DGEXP*, the inferences drawn from the time-series averages are consistent with those drawn from the pooled regression. The coefficient for *DGEXP* is positive (not negative as predicted) for TS10 and negative and marginally significant for TS5. Coefficients for the remaining variables have *t*-statistics that are smaller than the pooled regression, but still significant at conventional levels, and the coefficients are generally of a similar magnitude as the pooled regression.

5. Conclusion

I investigate whether analysts' ability, portfolio complexity and resources explain systematic differences in forecast accuracy. The three factors are jointly analyzed since they are likely to be correlated. The study uses a cross-sectional approach since the objective is to explain cross-sectional differences in forecast accuracy.

Forecast accuracy is found to increase with experience (a surrogate for ability) and employer size (a surrogate for resources), and decrease with the number of firms and industries followed (surrogates for portfolio complexity). My overall conclusion is that employer size alone, or the other labor market and portfolio variables taken in combination, may provide information for capital market participants to predict economically meaningful differences in forecast accuracy. For example, knowing an analysts employer size, or knowing the number of firms and industries followed may be sufficient information for capital market participants to predict economically meaningful differences in forecast accuracy. Note, however, that even small systematic differences in forecast accuracy may yield economically meaningful benefits for large investors.

The results imply that analysts' characteristics may be useful in predicting forecast accuracy, and that market expectations studies may be improved by modeling analysts' characteristics. A potential area for future research is to investigate whether capital market participants consider analysts' ability, resource and portfolio complexity variables in forming earnings expectations.

References

Butler, K.C., Lang, L.H., 1991. The forecast accuracy of individual analysts: Evidence of systematic optimism and pessimism. Journal of Accounting Research 29, 150–156.

Brown, L.D., Rozeff, M.S., 1980. Analysts can forecast accurately! Journal of Portfolio Management 6, 31–34.

Clement, M., 1998. Some considerations in measuring analysts' forecasting performance. Working Paper, University of Texas, Austin, TX.

Coggin, D. T. and J. E. Hunter, 1989. Analysts forecasts of EPS growth decomposition of error, relative accuracy and relation to return. Working paper. Michigan State University, East Lansing, MI.

Greene, W., 1991. Econometric Analysis. MacMillan Publishing Company, New York, NY.
Jacob, J., Lys, T., Neale, M., 1997. Expertise in forecasting performance of security analysts. Working paper. Northwestern University, Evanston, IL.
Lees, F., 1981. Public Disclosure of Corporate Earnings Forecasts. Conference Board, New York, NY.
Mikhail, M., Walther, B., Willis, R., 1997. Do security analysts improve their performance with experience? Journal of Accounting Research 35, 131–166.
Milgrom, P., Roberts, J., 1992. Economics, Organization & Management. Prentice Hall, Englewood Cliffs, NJ.
O'Brien, P., 1990. Forecast accuracy of individual analysts in nine industries. Journal of Accounting Research 28, 286–304.
O'Brien, P., 1987. Individual forecasting ability, Managerial Finance 13, 3–9.
Richards, R.M., 1976. Analysts' performance and the accuracy of corporate earnings forecasts. The Journal of Business 49, 350–357.
Sinha, P., Brown, L.D., Das, S., 1997. A re-examination of financial analysts' differential earnings forecast accuracy. Contemporary Accounting Research 14, 1–42.
Stickel, S., 1992. Reputation and performance among security analysts. Journal of Finance 47, 1811–1836.
Stickel, S., 1995. The anatomy of the performance of buy and sell recommendations. Financial Analysts Journal 51, 25–39.