Can Securities Analysts Forecast Intangible Firms’ Earnings?

Forthcoming in International Journal of Forecasting

Huong N. Higgins, Ph.D.
Associate Professor
Worcester Polytechnic Institute
School of Business
100 Institute Road
Worcester MA 01609
United States
Voice: 508-831-5626
Fax: 508-831-5720
Email: Hhiggins@wpi.edu

The author appreciates the helpful comments of the Journal’s Associate Editor and anonymous reviewers. A previous draft has benefited from comments by Vedran Capkun, Thomas Jeanjean, Balgobin Nandram, Yoshie Saito, Hervé Stolowy, and other participants to the research workshop at HEC, Paris in December 2007.
Can Securities Analysts Forecast Intangible Firms’ Earnings?

Abstract:

Much evidence in the 1990s cast serious doubt over securities analysts’ ability to produce accurate research of intangible firms. Such evidence contrasted with analysts’ image in the public mind as gatekeepers for the capital markets. This paper assesses the contentious question regarding analysts’ performance in forecasting the future earnings of intangible firms. The assessment is relative to extrapolative time-series models. The paper’s results show that the forecast errors produced by both analysts and extrapolative models are positively associated with intangibles that are above the industry norm, consistent with the difficulty of processing complex intangible information. However, the impact of intangibles on extrapolative models’ forecast errors is stronger than on analysts’. Analysts’ superiority is positively associated with firms’ specific intangibles, and this association increases as the complexity of intangible information increases. These finding are consistent with analysts’ better ability relative to extrapolative models to forecast the earnings of intangible firms.

Keywords: financial analysts, earnings forecasts, intangible assets, information complexity, time series.

JEL: M40, M41.
Can Securities Analysts Forecast Intangible Firms’ Earnings?

1. Introduction

There has been a long-standing agreement among economists and others that technological change is one of the most powerful forces influencing the American economy (Mansfield 1968). Because intangible activities help generate technological change, they have gained much importance over the recent decades. Between 1998 and 2007, business spending on intangible assets accounted for 6.3% of average annual growth in real GDP, compared to spending on commercial structures which accounted for just over 1.3% of average annual growth in real GDP over the same years (Bureau of Economic Analysis 2010).

The importance of intangible assets is accentuated by concerns that even securities analysts (analysts henceforth), who are a class of expert investors, may not understand intangibles. Despite a growing literature consistent with the notion that analysts play an important role in capital markets (Ramnath et al. 2008, Bradshaw et al. 2006, Gu and Chen 2004, Barron et al. 2002, Brown and Sivakumar 2005, Lang and Lundholm 1996, and Brown 1993), intangibles pose a particular challenge to analysts due to their complex information.

To date, there is not much evidence of how well analysts can process intangible information. One exception is a study by Gu and Wang (2005), which investigates the effect of information complexity of intangible assets on analysts’ forecast error. Those authors argue that information complexity of intangibles is primarily attributable to firm-specific intangibles, and accordingly find a positive relation between analysts’ forecast error and the
amount of firms’ intangibles that are above the industry norm. They conclude that information complexity increases the difficulty of forecasting intangible firms’ earnings.

Granted the results by Gu and Wang (2005), it is not clear how well analysts process intangible information. Despite analysts’ soaring public image in the 1980s as gatekeepers for the capital markets (Coffee 2006), and despite analysts’ superior forecast accuracy relative to mechanical models in those earlier years (Brown et al. 1987a-b), analysts’ behavior took a striking shift during the tech boom in the 1990s (Coffee 2006). The shift was manifest in several ways. First, the ratio of “buy” recommendations to “sell” recommendations increased dramatically from 6-to-1 in 1991 to 100-to-1 in 2000, according to a Thomson Financial/First Call survey. ¹ Similarly, 98.4% percent of analyst recommendations were either “buys” or “holds”, with only 1.6 percent being “sells” or “strong sells” between 1996 and 2000, the heart of the tech boom era (Barber et al. 2003). Second, the number of analyst recommendations nearly doubled during these years, from 22,409 in 1996 to 43,248 in 1999 (ibid). The sudden output increase raised questions about the quality of analysts’ work, i.e. whether analysts spent less time per report, because the analyst population could not have grown that rapidly. Third, analysts’ recommendations at the end of the tech boom were perverse: stocks rated by analysts as “buys” in 2000-2001 not only under-performed the market, but also under-performed the relatively few stocks given a “sell” recommendation (ibid). The collapse of the tech bubble in Spring 2000 brought attention to star analysts of the era, who had maintained “buy” recommendations on stocks

that were in free fall.Overall, the manifest shift in analyst behavior bluntly suggests that analysts cannot process intangible information.

Given the contrast between prior research, which showed analyst performance (for example Brown et al. 1987a-b), and the shift in analyst behavior during the tech boom, when intangible information underscored much of analyst research, the question whether analysts can process intangible information is contentious. This paper seeks to address this question via a comparison of analysts’ forecast performance with that of time-series models. The rationale is that the forecast performance reflects the forecaster’s ability to process information for forecasting. Specifically, this paper focuses on earnings forecasts. Analysts are assumed to forecast earnings based on information available to them. Analysts’ information consists of commonly available information in the past and private information processing by analysts themselves (Bradshaw 2009). In the case of intangible firms, analysts’ private information may include information about the firm’s intangible projects.

Time-series models are used as benchmark because they are an important class of earnings forecasting models which are often compared with analyst forecasts (Livnat and Mendenhall 2006, Walther 1997, Wiedman 1996, Lobo 1992, Lobo 1991, Brown et al. 1987a-b, and Givoly and Lakonishok 1984). Specifically, this paper uses extrapolative time-series models (extrapolative models henceforth), which forecast a firm’s earnings number based on its value in the prior year. Extrapolative models are arguably the best of time-series models at forecasting earnings (Brandon et al. 1983a and 1983b, and Watts and Leftwich 1977), and they are often the main time-series models employed in studies of earnings

---

2 For example, see Affidavit of Eric Dinallo, Assistant Attorney General of the State of New York, in Support of the Application for an Order Pursuant to New York General Business Law, s 345.
forecasts (Brown et al. 1987a-b, and Givoly and Lakonishok 1984). Extrapolative models are less effective than analysts because they merely exploit commonly available information in the past time-series without benefiting from human processing of information (Fried and Givoly 1982, and Brown et al. 1987a). If analysts do not perform better than extrapolative models, analysts have not been successful at processing their private and intangible information.

Although a number of extant studies already show that analysts forecast earnings better than time-series models (Brown et al. 1987a-b, and Givoly and Lakonishok 1984), prior research does not focus on intangible firms. Intangible firms need special attention because information about intangible firms is far more complex than information about general firms, thus it is possible that analysts’ resources are not sufficient for processing intangible information. Further, the number of intangible firms and the complexity of intangible information have risen dramatically during the high-tech boom in the 1990s. Therefore, the findings in prior research may not be generalizable to intangible firms. Whether analysts perform better than extrapolative models in forecasting intangible firms remains an empirical question to be addressed in this paper.

This paper finds a significantly positive link between forecast errors, by both analysts and extrapolative models, and firms’ intangibles that are above the industry norm (specific intangibles henceforth). The association between forecast errors and intangibles is consistent with the cost of information complexity, consistent with Gu and Wang (2005). Further, this paper documents a stronger link between extrapolative forecast errors and firms’ specific intangibles than that between analysts’ forecast errors and firms’ specific intangibles. In other words, the impact of specific intangibles on extrapolative models’ forecast errors is
stronger than on analysts’. The paper also documents that analysts’ superiority vis-à-vis extrapolative models is positively associated with firms’ specific intangibles relative to other financial information, and this association increases as the complexity of intangible information increases. These finding are consistent with analysts’ being more able to forecast the earnings of intangible firms than extrapolative forecasting models.

This paper is distinct from Gu and Wang (2005). Those authors point out the cost of processing intangible information on analysts, however they are silent on how analysts perform relative to other benchmarks when forecasting firms with intangible information. Gu and Wang (2005) may be interpreted as analysts’ not doing their jobs right with regards to intangible information. This paper goes further by clarifying that despite the complexity of intangible information, analysts still manage to forecast better than extrapolative models, which are a typical benchmark for assessing analyst performance. This paper’s findings suggest that analysts’ performance is decent with regards to intangible information.

This paper is directly related to prior research on the predictability of intangible firms in securities research. Despite the large literature on securities analysts, investigations of analyst research of intangible firms are relatively rare. For example, a search of over 1,400 publications listed in Thomson Financial Research Bibliography (2007), only about 20 address intangibles (including amortization, goodwill, intangible, intangible asset, research and development, and R&D). Prior research shows that analyst coverage and private information are greater for firms with larger intangibles (Barth et al. 2001 and Barron et al. 2002). Intangibles are associated with analysts’ forecast errors due to information complexity of intangibles (Gu and Wang 2005). Intangibles, simultaneously with dividends and capital expenditures, determine firm value (Guerard et al. 1987 and 1990). The use of an intangible
variable can enhance stockholder wealth models relative to the use of capital expenditures or dividends (Guerard and Mark 2003). This paper contributes to this line of research by documenting analysts’ better ability relative to extrapolative models to forecast the earnings of intangible firms.

This paper adds to prior research that compares the forecast ability of analysts and time-series models. From this line of research, analysts are believed to be better earnings forecasters than time-series models (Guerard and Beidleman 1986, and Givoly and Lakonishok 1984). Particularly, analysts benefit from the dimensionality of the information set (Brown et al. 1987b), and they generate fewer outliers (Collins and Hopwood 1980). However, prior research does not focus on intangible firms, which have more complex information than general firms and are more numerous than could be examined by prior studies. This paper contributes to this line of research by extending its past findings to intangible firms.

This paper is also related to prior research on the pricing of intangibles in the marketplace. The complexity of intangibles raises the possibility of mispricing, namely that stock prices do not fully incorporate the value of intangibles. Numerous studies in economics, finance, and accounting suggest that investors expect intangible investments to produce future benefits (Givoly and Shi 2008, Megna and Klock 1993, Chan et al. 1990, and Mansfield 1968 Chapter 4). However, prices may not fully reflect the information contained in past intangible expenditures, because investors have short time horizons and cannot process intangible information, and so they fail to anticipate the rewards from long-term intangible investments. Further, investors are generally fixated on reported earnings (Sloan 1996), whereas reporting standards for intangibles are inadequate (Lev and Sougiannis 1996,
Aboody and Lev 1998, Lev and Sougiannis 2005, and Wyatt 2005). Mispricing results from investors’ relying on non-persistent earnings information in financial statements (Richardson et al. 2005), and underpricing arises when investors mechanically accept financial statements at face value without adjusting for intangibles’ long-term benefits. Consistent with underpricing, many research studies indicate that R&D is positively associated with subsequent excess stock returns (Chambers et al. 2002, Penman and Zhang 2002, and Lev and Sougiannis 1996). Although this paper’s results highlight the ability of analysts, underpricing suggests that due to the complexity of intangible information, analysts still have room to improve as information intermediaries with respect to intangibles.

The remainder of this paper is organized as follows. Section 2 develops the paper’s Hypotheses. Section 3 discusses the data sources. Section 4 describes the empirical designs. Section 5 reports the summary data. Section 6 reports the empirical results. Section 7 presents additional discussions. And Section 8 concludes the paper.

2. Hypotheses Development

Analysts are professionals making their living doing securities research. Many analysts are specialized in particular industries and firms, resulting in highly industry-specific and firm-specific knowledge (Coffee 2006, and Schipper 1991). Analysts prepare detailed research reports on stocks they follow, which incorporate not only analyses of the company’s published financial information but also discussions with management and industry sources, including customers and suppliers (Coffee 2006). From these discussions, analysts gain insight beyond commonly available information. Competition among analysts for forecast accuracy helps create incentive to do research well and forecast accurately (Mikhail 1999). Analysts also benefit from their institutional and professional networks to continuously
improve their information processing skills. Analysts belonging to brokerage houses have strong networks (i.e., other analysts and supporting staff) for collecting, disseminating, and interpreting information (Jacob et al. 1999).

Many empirical studies show that analysts generally are good information processors. Analysts are assumed to be among the primary users of commonly available information (Bradshaw 2009), but analysts also acquire their own private idiosyncratic information (Barron et al. 2002). Analysts increasingly process their own earnings definitions and measurements that are different from those from standard financial reports (Bradshaw 2003, and Bhattacharya et al. 2003). Analysts do a good job of processing their own information, given that their earnings information is more persistent and has higher valuation multiples than financial statements’ information they choose to exclude (Gu and Chen 2004). Overall, the information processing hypothesis posits that analysts are expected to tackle intangibles well, because analysts have dedicated resources for processing information.

However, the information hypothesis is not without caveats. The first caveat is that analysts are subject to conflicts of interest due to analysts’ personal and careerist concerns. Analysts often personally own stocks that they research (McTague 2001). Analysts employed by investment banks often inflate their forecasts of the banks’ clients to help their employers win lucrative investment bank deals (Dugar and Nathan 1995). Inexperienced analysts tend to herd, particularly to avoid a downward deviation, because an individual mistake is deemed far worse professionally than a collective mistake (Clement and Tse 2005, and Hong et al. 2000). Analysts’ conflicts of interests tend to lead to optimistic bias in forecasts, which could yield results contrary to the information hypothesis. Consistent with this caveat, Hong and Kubik (2003) find that relative optimism outranks even overall accuracy in determining
career advancement of analysts, and optimism is most important in the case of analysts who cover stocks underwritten by their own firms.

The second caveat is the complexity of intangible information. Although analysts’ resources should help, these resources may not be sufficient to process complex information. There are two sources of complexity about intangibles, namely their inherent risks and prescriptive accounting standards. Intangibles have inherent risks, due to the fact that investments are at a most uncertain stage of the innovation process (Lev 2001), property rights are less well defined for intangible than for tangible assets (Lev 2001), and intangible values often cannot be independently discerned or owned by any given firm (Basu and Waymire 2008). These inherent risks make it difficult for firms to provide and for analysts to understand information about intangibles. Under accounting standards that are based on classes of transactions (i.e., purchased versus internally developed), intangible investments are not consistently reported with regards to their economic fundamentals. The prescriptive standards create inconsistent information and further complicate understanding of intangibles (Wyatt 2008 and 2005).

Due to the above two caveats, analysts may process intangible information no better than extrapolative models do, despite analysts’ dedicated resources for information processing. Ultimately, an empirical examination as this study is necessary to determine whether analysts can indeed process intangible information. Following are the paper’s hypotheses formulated in the alternative.

If the information hypothesis outweighs the caveats while analysts process intangible information, analysts’ forecast errors are less associated with intangible information than are
extrapolative forecast errors, and analysts’ superiority over extrapolative models should stem from analysts’ ability to process intangible information:

**Hypothesis 1:** The association between firms’ specific intangibles and extrapolative forecast errors is stronger than the association between firms’ specific intangibles and analysts’ forecast errors.

**Hypothesis 2:** Analysts’ earnings forecasting superiority vis-a-vis extrapolative models is positively associated with firms’ specific intangibles.

If analysts’ superiority over extrapolative models stems from analysts’ ability to process intangible information, analysts’ superiority should increase as intangible information becomes increasingly complex:

**Hypothesis 3:** The evidence supporting Hypothesis 2 is greater as the complexity of intangibles increases.

It is important to note that analysts’ advantage over extrapolative models is due to an information advantage, namely analysts can benefit from available information from multiple sources (Brown et al. 1987b). Thus, if analysts actually can process intangible information, analysts’ advantage over extrapolative models from intangible information is in addition to the advantage from other information. Therefore, the paper’s Hypotheses should be tested relative to other information than intangible information.

3. **Data Sources**

Analysts’ earnings and forecasts are retrieved from I/B/E/S. Extrapolative earnings and forecasts are assumed to stem from published financial reports, specifically continuing
operating income (COI) as reported on the income statement. This is because of all published reports, earnings information on the income statement has the best content for predicting future earnings, and of all earnings components presented on the income statement, COI best captures recurring earnings, and therefore has the best information content for predicting future earnings and stock price (Lipe 1986, Ou 1990, Penman 1989a and 1989b, Fairfield et al. 1996, DeChow and Ge 2005, and Fairfield et al. 2009).

Two types of extrapolative models are used because of their parsimony. The first is the no-change extrapolative model which uses the most recent earnings information to forecast future earnings. The no-change extrapolative model remains one of the best of tested time-series models to predict corporate earnings (Brandon et al. 1983a and 1983b, and Watts and Leftwich 1977). Alternatively and for robustness, the second type of extrapolative model is the extrapolative model incorporating growth, which uses the most recent earnings and earnings growth information to forecast future earnings.

In earnings forecasting, parsimonious time-series models have advantage over more sophisticated ones. Many studies have generally agreed that the earnings process is a random walk (Brandon et al. 1983a and 1983b, Watts and Leftwich 1977, Albrecht et. al. 1977, Ball and Watts 1972, and Brealey 1969). Because past information of a random walk is not easily exploitable, parsimonious models should perform as well as any more sophisticated time-series methods. Yet parsimonious models are more efficient due to their simplicity. Indeed, empirical evidence by Watts and Leftwich (1977) and Brandon et al. (1983a and 1983b) shows that Box-Jenkins models are outperformed by parsimonious extrapolative models. Therefore, if sophisticated time series models were used instead, time series forecasts would
most probably not be improved by much or at all, and the paper’s results would not be changed.

Data are drawn from the entire I/B/E/S U.S. Summary and Compustat databases from 1982 to February 2006. The main analyses use I/B/E/S one-year-ahead (FYR1) forecasts dated six months before the end of the previous year, to capture the forecast contemporaneous to the actual earnings announcement by a firm, while allowing some time for analysts to process their private information. Annual forecasts are used instead of quarterly forecasts to avoid seasonal effects, and because annual reports are subject to more comprehensive audit scrutiny than quarterly financial reports.\(^3\) The median of FYR1 forecasts is used to parallel Gu and Wang (2005), but the results are consistent when the mean (consensus) forecast is used.\(^4\)

The sample firm-years must meet the following requirements:

1) FYR1 forecasts are available on I/B/E/S. Forecasts must be dated 6 months (between 165 and 195 days) after the end of the previous year.

2) Compustat data are available for stock price, book value, total asset, continuing operating income and net income.

\(^3\) Many studies use quarterly data, however, as Lambert (2004) discusses, they do not address seasonal effects. Lambert (2004) further states that quarterly earnings are not “that important” in valuation, namely their value relevance and predictive ability are not that high.

\(^4\) A number of replications are performed besides the main reported analyses. First, replications are performed using FYR2 instead of FYR1 forecasts. Second, the first forecast issued for a given period end date is used, instead of the last forecast. The replicated results are consistent with the main reported results.
3) Firms that are Utilities (SIC: 4900-4999), Banks (SIC: 6000-6411), and Insurance and Real Estate (SIC: 6500-6999) are excluded due to their special regulations.

4) Intangibles’ data are available on Compustat.

The resulting sample consists of 29,023 firm-years. Due to restriction 4), the sample contains firms with intangible activities and therefore with relatively greater analyst following.

4. Empirical Designs

4.1 Hypothesis 1

Hypothesis 1 predicts that the association between specific intangibles and extrapolative error is stronger than that between specific intangibles and analyst error. To find support for this Hypothesis, a multivariate multiple regression is used to test the joint linear effect of intangible variables on the set of two dependent variables, namely analysts’ and extrapolative forecast errors. The purpose is to compare the associations between analysts’ and extrapolative earnings forecast errors and intangibles. The advantage of this multivariate multiple regression over a system of two separate multiple regressions is to capture the covariance structure across all observations.


6 The covariance structure among securities is important enough to affect key decisions in finance, such as determining risk (Rosenberg 1974), and constructing optimal portfolios (Elton et al. 1979).

In an alternative test that does not consider covariance across observations, the following two separate multiple regressions are performed using the same variables as Equation (1):
Specifically, the multivariate multiple regression has the following form:

\[
\begin{pmatrix}
\text{AnalystError}_{t+1} \\
\text{ExtrapolativeError}_{t+1}
\end{pmatrix} = a_0 + a_1 \text{RD}_{it} + a_2 \text{AD}_{it} + a_3 \text{BI}_{it} + a_4 \text{LMV}_{it} + a_5 \text{DISP}_{it} + a_6 \text{DAYS}_{it} + \\
a_7 \text{STDE}_{it} + a_8 \text{LOSS}_{it} + a_9 \text{MTB}_{it} + a_{10} \text{COV}_{it} + a_{11} \text{MKT}_{t} + \Sigma \gamma_k YR_k + \epsilon_{it}
\]

(1)

In Equation (1), Analyst and Extrapolative Errors are computed based on three common forecast error metrics\(^7\), namely absolute error scaled by price (MESP), absolute error (MAFE)\(^8\), and absolute percentage error (MAPE). Analyst Errors are labeled with the

\[
\text{AnalystError}_{t+1} = a_0 + a_1 \text{RD}_{it} + a_2 \text{AD}_{it} + a_3 \text{BI}_{it} + a_4 \text{STDE}_{it} + a_5 \text{LOSS}_{it} + a_6 \text{MTB}_{it} + \\
a_7 \text{LMV}_{it} + a_8 \text{COV}_{it} + a_9 \text{MKT}_{t} + \epsilon_{it}
\]

\[
\text{ExtrapolativeError}_{t+1} = \beta_0 + \beta_1 \text{RD}_{it} + \beta_2 \text{AD}_{it} + \beta_3 \text{BI}_{it} + \beta_4 \text{STDE}_{it} + \beta_5 \text{LOSS}_{it} + \beta_6 \text{MTB}_{it} + \\
\beta_7 \text{LMV}_{it} + \beta_8 \text{COV}_{it} + \beta_9 \text{MKT}_{t} + \epsilon_{it}
\]

The paper’s Hypothesis predicts that \(\beta_1 > \alpha_1\), \(\beta_2 > \alpha_2\), and \(\beta_3 > \alpha_3\). The above inequalities may be tested using Z statistics on the regression coefficients, computed as

\[
Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}
\]

where \(\bar{x}_1\) and \(\bar{x}_2\) are the regression coefficients, \(\sigma_1\) and \(\sigma_2\) are their standard deviations, and \(n_1\) and \(n_2\) are the sample sizes. Z tests are often used in research on securities analysts and stock valuations to compare the impacts of factors on different dependent variables in different regressions (for example, see Brown and Sivakumar 2005, and Ely and Pownall 2002). Replications based on this alternative test are consistent with the tabulated results.

\(^7\) With respect to each firm-year of earnings, MESP = abs[forecast-actual]/price; MAFE = abs[forecast – actual], and MAPE = abs[(forecast – actual)/actual]. MESP is used by Gu and Wang (2005). MAFE and MAPE are often used as the means of a sample to evaluate forecast models.

\(^8\) It should be noted that MAFE is not scaled and therefore could be problematic due to scale differences, however its use should provide additional robustness to the paper. Scale differences arise when large (small) firms have large (small) values of variables. If the magnitudes of the differences are unrelated to the research question (forecast error), they result in biased regression coefficients. For example, a one-cent MAFE is a relatively small forecast error for a firm with EPS of a dollar per share, but a relatively great error for a firm with EPS of five cents per share. Lo and Lys (2000) and Barth and Kallapur (1996) argue
prefix A (AMESP, AMAFE, and AMAPE), while Extrapolative Errors are labeled with the prefix E (EMESP, EMAFE, and EMAPE). RD, AD, and BI are intangible variables. LMV, DISP, DAYS, STDE, LOSS, MTB, COV, and MKT are control variables that, as motivated by prior studies, capture the information environment of a firm, to consider analyst advantage stemming from information other than intangible information. And YRk are year dummies. All variables of Equation (1) are discussed in full details in the remainder of Section 4.1.

AnalystError is measured via AMESP, AMAFE, and AMAPE. AMESP_{t+1} measures analysts’ earnings forecast error for year t+1, defined as the absolute difference between the median I/B/E/S analysts’ forecast of earnings per share of year t+1 and I/B/E/S actual earnings per share of the same year. The absolute difference is deflated by stock price. The forecast is issued six months after the end of fiscal year t. This time line is chosen so that the release of analyst forecast of year t+1 is close to the actual earnings announcement of year t to facilitate comparisons between analyst information and financial statement information, while allowing some time for analysts to generate their private information beyond information from the recently released financial statements. AMESP_{t+1} is deflated by the stock price (PRICE) taken from I/B/E/S and dated the quarter before the release of the selected I/B/E/S/ forecast. Thus, the timing of stock price falls between the release of

that scale differences result in heteroscedastic regression error variances, and if severe enough may lead to opposite coefficient signs in regression models. In this paper, the only results that are contrary to sign expectations are based on this metric.

\footnote{In unreported replications, PRICE_{it} is also selected as Compustat monthly price item PRCC taken at three months after the fiscal year end for year t, and scale AMESP, EMESP1 and EMESP2 by this stock price measure. This selection also aims at a time between the release of published reports for year t and the release of analyst forecasts for year t+1. All results are robust when this Compustat price data is used instead of I/B/E/S data.}
published reports for year t and the release of analyst forecast for year t+1. AMAFE is 
analyst absolute error, and AMAPE is analyst absolute percentage error, both computed 
based on the same data as AMESP.

\[
\text{AMESP}_{it+1} = \frac{\text{ForecastIBES-EPS}_{it+1} - \text{ActualIBES-EPS}_{it+1}}{\text{PRICE}_{it}} \quad (2)
\]

\[
\text{AMAFE}_{it+1} = \text{abs}[\text{ForecastIBES-EPS}_{it+1} - \text{ActualIBES-EPS}_{it+1}] \quad (3)
\]

\[
\text{AMAPE}_{it+1} = \text{abs}[(\text{Forecast IBES-EPS}_{it+1} - \text{Actual IBES-EPS}_{it+1})/\text{Actual IBES-EPS}_{it+1}] \quad (4)
\]

In Equation (1), \textit{ExtrapolativeError} is measured via no-change extrapolative time-
series models (namely EMESP1, EMAFE1, and EMAPE1), and via extrapolative time-series 
models incorporating growth (namely EMESP2, EMAFE2, and EMAPE2). EMESP1\textsubscript{it+1} is 
defined as the absolute difference between actual earnings per share of year t and year t+1, 
both as reported in financial statements under income from continuing operations per share 
(COI, Compustat Data 58).\textsuperscript{10} This definition assumes a no-change extrapolative forecast 
model, forecasting COI in year t+1 as its value in year t. EMESP1\textsubscript{it+1} is deflated by the stock 
price as of the quarter before the release of analysts’ earnings forecast. EMAFE1 is no-
growth extrapolative absolute error, and EMAPE1 is no-growth extrapolative absolute 
percentage error, both computed based on the same data as EMESP1.

\[
\text{EMESP1}_{it+1} = \text{abs}[\text{COIpershare}_{it} - \text{COIpershare}_{it+1}]/\text{PRICE}_{it} \quad (5)
\]

\[
\text{EMAFE1}_{it+1} = \text{abs}[\text{COIpershare}_{it} - \text{COIpershare}_{it+1}] \quad (6)
\]

\[
\text{EMAPE1}_{it+1} = \text{abs}[(\text{COIpershare}_{it} - \text{COIpershare}_{it+1})/\text{COIpershare}_{it+1}] \quad (7)
\]

\textsuperscript{10} Compustat Data 58 is a per share number, therefore it is adjusted for stock splits and stock 
dividends via Compustat Data 27. In fact, all per share measures in this paper are adjusted for 
stock splits and stock dividends in this way.
EMESP2_{it+1} is extrapolative models’ earnings forecast error for year t+1, defined as the absolute difference between actual earnings per share of year t multiplied by the earnings growth experienced between year t-1 and year t and actual earnings per share of year t+1. Earnings per share is as reported in financial statements under income from continuing operations per share (COI, Compustat Data 58). This definition assumes that extrapolative models incorporate earnings growth, forecasting COI in year t+1 as its value in year t times an expected growth factor which is determined by the change rate in COI from year t-1 to year t. To obtain meaningful measures of earnings growth, EMESP2_{it+1} is computed only for firms that have positive COI in year t-1. EMESP2_{it+1} is also deflated by the stock price as of the quarter before the release of analysts’ earnings forecast. EMAFE2 is with-growth extrapolative absolute error, and EMAPE2 is with-growth extrapolative absolute percentage error, computed based on the same data as EMESP2.

\[
\text{FACTOR}_{it} = 1 + \frac{(\text{COIpershare}_{it} - \text{COIpershare}_{it-1})}{\text{COIpershare}_{it-1}} \quad (8)
\]

\[
\text{EMESP2}_{it+1} = \text{abs}[\text{COIpershare}_{it} \cdot \text{FACTOR}_{it} - \text{COIpershare}_{it+1}] / \text{PRICE}_{it} \quad (9)
\]

\[
\text{EMAFE2}_{it+1} = \text{abs}[\text{COIpershare}_{it} \cdot \text{FACTOR}_{it} - \text{COIpershare}_{it+1}] \quad (10)
\]

\[
\text{EMAPE2}_{it+1} = \text{abs}[(\text{COIpershare}_{it} \cdot \text{FACTOR}_{it} - \text{COIpershare}_{it+1}) / \text{COIpershare}_{it+1}] \quad (11)
\]

Summing the Extrapolative Errors, two extrapolative forecast models are used, one with no change or without growth (EMESP1, EMAFE1, and EMAPE1) and the other incorporating growth (EMESP2, EMAFE2, and EMAPE2). A large body of the literature that examines the predictability of earnings has concluded that the no-change extrapolative model is the best time-series models to forecast earnings (Watts and Leftwich 1977). Therefore,
EMESP1 (EMAFE1, and EMAPE1) is supposed to perform better, however EMESP2 (EMAFE2, and EMAPE2) is also used for robustness.

Continuing with the variables in Equation (1), RD, AD, and BI measure the relative degree to which a firm is comprised of intangible assets: R&D expenses (RD, Compustat Data 46), brand promotions and advertising expenses (AD, Compustat Data 45), and intangibles recognized on the firm’s balance sheet (BI, Compustat Data 33). All three proxies are deflated by equity market value, measured as stock price multiplied by the number of shares, both taken from I/B/E/S at the same timing as for computing AMESP.11 These three proxies constitute the major categories of intangible information (Wyatt 2008), are used by Gu and Wang (2005) and Barron et al. (2002) as proxies of intangibles, and are computed similarly to Gu and Wang (2005). The accounting treatments and economic characteristics underlying the above three measures may differ substantially, however higher levels of these variables generally indicate greater intangible investments (Barron et al. 2002).

Turning to the control variables, Equation (1) has information variables as motivated by Brown et al. (1987b). Firm size (LMV), which denotes the dimensionality of the information environment, is measured as the natural log of equity market value at the end of year t, where equity market value is measured as I/B/E/S stock price multiplied by I/B/E/S number of shares outstanding. Prior analyst dispersion (DISP), which denotes the variance in information observations, is measured as the cross-sectional standard deviation of upcoming annual forecasts deflated by the absolute value of mean forecast across analysts. Number of days (DAYS) between the end of the previous fiscal year and the date of the analyst

11 The results are unchanged when PRICE and number of shares are taken from Compustat at three months after the fiscal year end for year t.
forecast\textsuperscript{12} is included to denote analysts’ time advantage. As argued by Brown et al. (1987b)\textsuperscript{13}, larger measures for LMV and DAYS are associated with better analyst information processing; on the contrary, a larger measure for DISP denotes a challenge to analyst information processing.

Equation (1) also includes other information variables motivated by other studies. Volatility in firm earnings (STDE) is the standard deviation of ROA over the ten years preceding year $t+1$, where ROA is measured as the ratio between net income (Compustat Data 172) and total asset (Compustat Data 6). Fluctuations of a firm’s information environment reflect the inherent variability of the firm and the variance of its information observations, making it more difficult to understand its information and forecast its future earnings (Lang and Lundholm 1996). Firm earnings loss (LOSS) is a dummy variable equal to 1 if net income (Compustat Data 172) in year $t$ is negative, and 0 otherwise. Due to the inherent variability of loss firms, earnings loss is difficult to forecast (Hwang et al. 1996). Firm growth (MTB) is measured by the ratio of market value of equity to book value of equity (Compustat Data 199 * Data 25 / Data 60) at the end of year $t$. Analysts tend to cover firms of higher growth (Barth et al. 2001), thus firm growth denotes rich information environment, but conversely, growth may denote fluctuations in the information environment due to the inherent variability of growth firms. Analyst coverage (COV) is measured as the number of analysts contributing to the I/B/E/S median earnings estimate. As analysts tend to

\textsuperscript{12} IBES forecasts are a consensus of individual forecasts which are made on different dates. Thus, this procedure for capturing time advantage is only an approximation.

\textsuperscript{13} Brown et al. (1987b) also control for industry and forecast horizon. In this paper, industry effects are controlled by adjusting all variables by industry medians, and only the FYR1 forecast horizon is included in the main sample, while the FYR2 forecast horizon is replicated separately.
cover growth firms (Barth et al. 2001), analyst coverage may denote both rich information environment and information fluctuations.

Finally, Equation (1) includes MKT and YRk to control for macro-economic conditions. MKT measures overall annual market return (MKT), defined as the annual return of the Dow Jones 5000 Index reported by Datastream. MKT may denote market information environment and thus may help reduce forecast errors, but may also capture market fluctuations and thus may help increase forecast errors. As in Brown et al. (1987b), year dummies (YRk) are included to capture other effects associated with specific years.

Industry effects are controlled by measuring all variables as deviations from the three-digit SIC medians before entering them into the regressions. As discussed in Gu and Wang (2005), measuring intangible variables as deviations from the three-digit SIC industry medians helps capture intangibles specific to the given firms that are above the industry norms. As a result, the estimated coefficients of the intangible variables inform whether within industry forecast errors are related to a firm’s intangibles relative to its industry.

Hypothesis 1 predicts that $\alpha_{1,\text{Analyst}} < \alpha_{1,\text{Extrapolative}}; \alpha_{2,\text{Analyst}} < \alpha_{2,\text{Extrapolative}}; \alpha_{3,\text{Analyst}} < \alpha_{3,\text{Extrapolative}}$ in Equation (1). The inequalities are tested via F statistics\(^\text{14}\) produced by the multivariate multiple regression procedure to compare the analyst error coefficients with the extrapolative error coefficients.

\(^{14}\) The F statistic approximates the likelihood ratio for testing that the Hypothesis comparison is zero. The formulation of the F statistic is described by Everitt and Dunn (1991, page 219):

\[
F = \frac{(n_1 + n_2 - p - 1) \times \frac{n_1 n_2}{(n_1 + n_2 - 2) p} \times (\bar{x}_1 - \bar{x}_2)' S^{-1}(\bar{x}_1 - \bar{x}_2)}{S},
\]

where $\bar{x}_1$ and $\bar{x}_2$ are the regression coefficients, $n_1$ and $n_2$ are the sample sizes, $p$ is the number of independent variables, and $S$ is the pooled estimate of the within-groups variance-covariance matrix. The statistic is
4.2. Hypothesis 2.

Hypothesis 2 predicts that analyst superiority is associated with specific intangibles. To find support for Hypothesis 2, the following regression is used:

\[
ASUP_{it} = \alpha_0 + \alpha_1 RD_{it} + \alpha_2 AD_{it} + \alpha_3 BI_{it} + \alpha_4 LMV_{it} + \alpha_5 DISP_{it} + \alpha_6 DAYS_{it} + \alpha_7 STDE_{it} + \\
\alpha_8 LOSS_{it} + \alpha_9 MTB_{it} + \alpha_{10} COV_{it} + \alpha_{11} MKT_{t} + \Sigma \gamma_k YR_k + \epsilon_{it}
\] (12)

where ASUP proxies for analyst superiority over extrapolative models, and is measured based on the ratio between extrapolative forecast error and analyst forecast error as in Brown et al. (1987b):

\[
ASUP_{1it} = \text{Ln}(EMESP_{1it}^2/AMESP_{it}^2)
\] (13)

\[
ASUP_{2it} = \text{Ln}(EMAFE_{1it}^2/AMAFE_{it}^2)
\] (14)

\[
ASUP_{3it} = \text{Ln}(EMAPE_{1it}^2/AMAPE_{it}^2)
\] (15)

All other variables are defined as for Hypothesis 1.\(^{15}\) All firm variables are adjusted for industry medians to control for industry effects. Hypothesis 2 predicts that \(\alpha_1 > 0; \ \alpha_2 > 0; \ \text{and} \ \alpha_3 > 0.\)

4.3. Hypothesis 3.

distributed as an F-variate with \(p\) and \((n1 + n2 - p - 1)\) d.f. under the null hypothesis of no difference.

\(^{15}\) Only no-change extrapolative forecast errors are used in measuring analyst superiority, as subsequent analyses reveal that those incorporating growth have obviously larger errors, which makes it easier to find analyst superiority. Replications based on extrapolative forecasts incorporating growth are similar and not reported.
Hypothesis 3 predicts that the association between analyst superiority and specific intangibles increases as the complexity of intangibles increases. Because this complexity rises in recent time and with intensive intangible activity, the association should strengthen in recent time and in cases of intensive intangible activity.

Using variables defined for Hypotheses 1-2, the two following regressions incorporate these two measures (namely, recent time and intensity of intangible activity) in interaction with specific intangibles:

\[
\text{ASUP}_{it} = \alpha_0 + \alpha_1 \text{REC}_{it} + \alpha_2 \text{BAR}_{it} + \alpha_3 \text{LMV}_{it} + \alpha_4 \text{DISP}_{it} + \alpha_5 \text{STDE}_{it} + \alpha_6 \text{LOSS}_{it} + \alpha_7 \text{MTB}_{it} + \alpha_8 \text{COV}_{it} + \alpha_9 \text{MKT}_{it} + \alpha_{10} \text{TENS}_{it} * \text{REC}_{it} + \alpha_{11} \text{TENS}_{it} * \text{BAR}_{it} + \alpha_{12} \text{TENS}_{it} * \text{LMV}_{it} + \alpha_{13} \text{TENS}_{it} * \text{DISP}_{it} + \alpha_{14} \text{TENS}_{it} * \text{STDE}_{it} + \alpha_{15} \text{TENS}_{it} * \text{LOSS}_{it} + \epsilon_{it} \\
\]

In Equation (15), recent time (REC) is a dummy equal 1 if the year is larger than 1998, and 0 otherwise. The year 1998 is selected as it splits the sample in approximately half and half based on time. Hypothesis 3 predicts that \(\alpha_{12} > 0; \ \alpha_{13} > 0; \ \alpha_{14} > 0; \ \text{and} \ \alpha_{15} > 0.\)

In Equation (16), intensity of intangible activity (TENS) is a dummy equal 1 if adjusted RD is larger than the industry’s median, and 0 otherwise. Hypothesis 3 also predicts that that \(\alpha_{12} > 0; \ \alpha_{13} > 0; \ \alpha_{14} > 0; \ \text{and} \ \alpha_{15} > 0\) in this Equation.

5. **Summary data**

\[\text{Replications using many years in the 1990s instead of 1998 show consistent results and are not reported. Year dummies are not used here to avoid excessive information overlap among the time related variables.}\]
Panel A of Table 1 reports the summary statistics of the variables of interest after winsorizing the 1% outliers and before adjusting for industry medians, while Panel B reports these statistics after industry adjustment. From both Panels, all mean firm values are higher than the medians, indicating concentration of a subset of firms that have large values for the variables, except for DAYS which has a smaller mean than its median.

From Panel A, analyst earnings forecast errors are smaller than extrapolative models’ forecast errors on average. Comparing the means, AMESP = 0.026 versus EMESP1 = 0.071 and EMESP2 = 0.190. Comparing the medians, AMESP = 0.007 versus EMESP1 = 0.021 and EMESP2 = 0.028. The AMESP mean and median are identical to those reported by Gu and Wang (2005) who use the same computation method. However, the standard deviation of analyst forecast error reported by Gu and Wang (2005), which is 0.125, is larger than the one reported in this paper, which is 0.082. This difference reflects smaller variability in this paper’s sample, which is larger and consists of relatively better covered firms.

The means of absolute errors and absolute percentage errors also show analysts’ better accuracy: AMAFE = 0.276, versus EMAFE1 = 0.758 and EMAFE2 = 1.991; and AMAPE = 0.554, versus EMAPE1 = 1.351 and EMAPE2 = 3.084. Similarly, the median values also show analysts’ better accuracy: AMAFE = 0.100, versus EMAFE1 = 0.313 and EMAFE2 = 0.445; AMAPE = 0.122, versus EMAPE1 = 0.366 and EMAPE2 = 0.535. Likewise, Theil’s Inequality Coefficients (TICs), defined as the ratios of analyst root mean square errors and extrapolative root mean square errors, are lower than 1, indicating that

---

17 Gu and Wang (2005) report a mean of 0.026 and a median of 0.007 for analyst earnings forecast error.
analyst forecast error is smaller than extrapolative forecast error.\(^{18}\) Consistent with analyst superiority over extrapolative models, the averages of ASUP1-ASUP3 are positive (respectively 2.201, 2.205, and 2.156), and so are their medians (respectively 2.173, 2.172, and 2.097). Overall, all the utilized metrics are consistent in showing that analyst forecasts are more accurate than extrapolative forecasts.

<Table 1 about here>

It should be noted that the reported statistics of forecast errors are based on absolute differences, therefore their magnitudes are quite larger than if they were based on signed differences. It is also notable that forecasts based on the no-change extrapolative model of earnings are better on average than those based on the extrapolative model incorporating earnings growth. For example, based on the means: EMESP1 = 0.071 < EMESP2 = 0.190; EMAFE1 = 0.758 < EMAFE2 = 1.991; and EMAPE1 = 1.351 < EMAPE2 = 3.084. This result is consistent with the argument that the corporate earnings process is a random walk (Brandon et al. 1983a and 1983b, and Watts and Leftwich 1977), and the captured growth is not persistent enough to be useful for forecasting.

Regarding the intangible variables, the means of RD, AD and BI are 0.027, 0.010, and 0.105, respectively. Their medians are 0.000, except for BI which is 0.008. Compared to the results reported by Gu and Wang (2005), the sample firms in this paper have twice the

\(^{18}\) Root mean square error (RMSE) = \(\sqrt{\frac{\sum_{k=1}^{n} (\text{Forecast} - \text{Actual})^2}{n}}\)

\[\text{TIC}_{\text{analyst / no-growth-extrapolative}} = \frac{\text{RMSE}_{\text{analyst}}}{\text{RMSE}_{\text{nogrowth extrapolative}}} = 0.42\]

\[\text{TIC}_{\text{analyst / with-growth-extrapolative}} = \frac{\text{RMSE}_{\text{analyst}}}{\text{RMSE}_{\text{withgrowth extrapolative}}} = 0.04\]
rate of balance sheet intangibles, but only half the rate of R&D expenses, and only one third the rate of advertisement expenses. These comparisons are consistent with a tendency by companies to switch from expensing to capitalizing intangibles during the height of the tech boom.

Turning to the control variables, the average natural log of market value (LMV) is 6.268, corresponding to about 400 Millions USD. Prior analyst dispersion (DISP) averages 0.158, larger than the statistic reported by Brown et al. (1987b) which is equivalent to 0.069, perhaps because this paper’s sample consists of more intangible firms with more information variance. DAYS, the number of days between the end of the previous fiscal year and the date of the analyst forecast, averages 190.387. STDE, the standard deviation of ROA over the years preceding the forecast, averages 0.126, which denotes a substantial variation in profitability among sample firms. The average of LOSS is 0.198, which approximates the fraction of firms with negative earnings before extraordinary items. MTB averages 2.996, denoting the capitalization rate of corporate book assets. COV, analyst coverage, averages 7.832. The LMV and COV statistics denote about the same average firm size but better average analyst coverage than the sample used by Gu and Wang (2005), who report a mean log market value of around 6.113 and a mean analyst coverage of 3.105. The difference in analyst coverage likely stems from the difference in the examination periods, as this paper’s sample covers more recent years that are not included by Gu and Wang (2005). Some of these years (between 1998 and 2006) are the height of the tech boom, with very intensive analyst activity. MKT averages 0.283, which denotes the annual rate of return of the Dow Jones 5000 index over the examination period.
Table 2 reports the Pearson correlation coefficients among the variables of interest after adjusting for industry medians.\textsuperscript{19} The correlations are consistent with the expectations set forth in Section 4.1. Analysts’ forecast errors are mostly positively correlated with intangibles, consistent with Gu and Wang’s (2005) argument that intangible information complicates the forecasting process. Extrapolative forecast errors are also mostly positively correlated with intangibles, consistent with the notion that intangibles are associated with fluctuations that make it difficult to use extrapolative earnings to forecast future earnings.

The first measure of analyst superiority (ASUP1) is positively correlated with R&D expense (RD) and balance sheet intangibles (BI), consistent with the notion that analysts’ advantage stems from their ability to process intangible information. All three measures of analyst superiority (ASUP1-3) are positively correlated with firm size (LMV), earnings loss (LOSS) and annual market return (MKT), consistent with the notion that analysts’ advantage stems from the information environment. Forecast errors are positively associated with prior analyst dispersion (DISP), the volatility of historical earnings (STDE), and the status of loss firms (LOSS), consistent with the notion that forecast errors are exacerbated by the fluctuations of firm information environment. Forecast errors are mostly negatively associated with firm size (LMV), growth (MTB) and analyst coverage (COV), consistent with the notion that forecast errors are reduced by the richness of firm information environment. Forecast errors are both negatively and positively associated with annual market return (MKT), suggesting that this variable denotes both the richness and fluctuations in the market information environment.

\textit{<Table 2 about here>}

6. Empirical Results

\textsuperscript{19} Coefficients of unadjusted variables and Spearman coefficients reveal similar information and are not reported.
6.1. **Hypothesis 1**

Table 3 reports the results of Equation (1), which jointly estimates the regressions of analyst and no-change extrapolative forecast errors. Three models correspond to the three metrics of forecast error: MESP in Model 1, MAFE in Model 2, and MAPE in Model 3. All models include the three intangible variables (RD, AD, and BI) and all the control variables of Equation (1). Except for MKT which is a market variable and YR_k which are year dummies, all variables are adjusted by their industry medians to control for industry effects. As a result, significant effects from industry-adjusted variables can be interpreted as above the industry norms.

<Table 3 about here>

The sign expectations for the analyst regressions are as follows. The three intangible variables are expected to be positive, consistent with the notion that the complexity of intangible information exacerbates forecast errors. LMV and DAYS are expected to be negative, consistent with the notions that the richness of firm information environment (LMV) and time advantage (DAYS) reduces forecast errors. DISP, STDE and LOSS are expected to be positive, consistent with the notion that the fluctuations of firm information environment exacerbate forecast errors. MTB, COV, and MKT are expected to be either negative or positive, because they denote both rich and fluctuating firm information environment.

Except for DAYS, the sign expectations for the extrapolative regressions are the same as those for the analyst regressions. The reason is that factors that enhance the information environment, which reduces analyst forecast error, are also those that underlie earnings
stability, which mechanically reduces extrapolative forecast error. Vice versa, factors that create variance in information observations, which increases analyst forecast error, are also those that underlie earnings fluctuations, which mechanically increase extrapolative forecast error. The variable DAYS, which relates to analysts’ time advantage but is not related to extrapolative processes, has no sign expectation in the extrapolative regressions.

As seen from Panel A of Table 3, the results are mostly consistent with the sign expectations. DISP and LOSS are significantly positive in all regressions, while MTB and LMV are significantly negative in most regressions. DAYS is significantly negative in two analyst regressions, but insignificant in the extrapolative regressions. The sign of COV is as expected but not consistent throughout all the regressions: Model 2 shows a different sign than Models 1 and 3. The one unexpected result is LMV, which is significantly positive instead of negative in the extrapolative regression of Model 2. The inconsistent sign of COV and the unexpected sign of LMV are perhaps due to scaling differences in the measure of MAFE\(^20\) in Model 2.

Focusing on RD, AD, and BI, their coefficients are significantly positive as expected in all regressions of Models 1-3 (all p-values <= 0.0065). The positive sign of the intangible variables is consistent with the notion that complexity in intangible information increases forecast error, and consistent with Gu and Wang (2005).

To seek support for Hypothesis 1, the three intangible coefficients shown in Panel A are compared between the analyst and extrapolative regressions in each of Models 1-3, with statistical tests shown in Panel B. In Model 1, the RD coefficient in the analyst regression

---

\(^{20}\) As discussed in Footnote 8, MAFE is not scaled and could be problematic.
(α₁_{Analyst}) is 0.1611, while that in the extrapolative regression (α₁_{Extrapolative}) is 0.5548. This comparison is statistically significant (F-test = 138.80, p-value < .0001). The AD coefficient in the analyst regression (α₂_{Analyst}) is 0.1117, while that in the extrapolative regression (α₂_{Extrapolative}) is 0.2993, a comparison that is statistically significant (F-test = 12.00, p-value = .0005). The BI coefficient in the analyst regression (α₃_{Analyst}) is 0.0295, while that in the extrapolative regression (α₃_{Extrapolative}) is 0.0804, a comparison that is also statistically significant (F-test = 66.24, p-value < .0001). Summing Model 1, all three intangible coefficients are statistically significantly smaller in the analyst regression than in the extrapolative regression, consistent with Hypothesis 1. Likewise, the comparisons in Models 2 and 3 also show statistically significantly smaller intangible coefficients in the analyst regressions (all the corresponding F-statistics have p-values < 0.0009). These results also provide support for Hypothesis 1.

Table 4 reports the results of Equation (1) which jointly estimates the regressions of analyst and with-growth extrapolative forecast errors. Table 4 is structured identically to Table 3, except that the Table 4 extrapolative regression is based on extrapolative forecasts incorporating growth, whereas that of Table 3 is based on no-growth extrapolative forecasts.

From Panel A of Table 4, the signs and statistical significance of all control variables are similar to Table 3. The intangible variables, RD, AD, and BI, are significantly positive in eight of nine regressions of Models 1-3, consistent with the notion that the complexity of
intangible information exacerbates forecast errors.\textsuperscript{21} Consistent with Hypothesis 1, Panel A shows that the intangibles’ coefficients in the analyst regressions are smaller than their equivalents in the extrapolative regressions in all three models. Panel B of Table 4 provides formal tests for comparing the intangibles’ coefficients between the analyst and extrapolative regressions. As shown, eight of nine\textsuperscript{22} F tests are statistically significant (p-values $<$0.0008), consistent with Hypothesis 1. In sum, the results reported in Tables 3-4 provide support for Hypothesis 1.

6.2. Hypothesis 2

Table 5 reports the results of Equation (12) via Models 1-3, which correspond to three measures of analyst superiority, ASUP1-ASUP3, respectively. Hypothesis 2 predicts that the intangible variables (RD, AD, and BI) in the Equation are positive because analyst superiority should stem from their ability to process intangible information. The signs of other variables are also formed in accordance with an information interpretation of analyst superiority. LMV, MTB, COV, and MKT are expected to be positive due to analysts’ information advantage, as these variables denote a rich information environment. But, as discussed in Section 4.1, MTB, COV, and MKT are also expected to be negative as these variables also denote information fluctuations which pose a challenge to analysts’ information processing. DAYS is expected to be positive due to analysts’ time advantage. DISP is expected to be negative due to variance in information observed by analysts. STDE and LOSS are expected to be positive as analysts are expected to address to some extent the

---

\textsuperscript{21} The exception is the MAFE extrapolative regression. As discussed in Footnote 8, compared to the other measures of forecast error, MAFE is a noisier measure due to a lack of scaling.

\textsuperscript{22} See Footnote 21.
information variance underlying these variables, while this variance mechanically induces large extrapolative errors.

As shown in Table 5, most significant control variables are in their expected signs: LMV, and LOSS are significant in all three models, STDE and MTB in Model 1, and DISP and DAYS in Model 3.

More importantly, RD is significantly positive in Model 1 (p-value = 0.0059), and BI is significantly positive in Models 1-3 (p-values = 0.0003, 0.0555, and 0.0477, respectively). These results are consistent with the notion that analyst superiority stems from their ability to process intangible information, and are consistent with Hypothesis 2.

6.3. Hypothesis 3

Panel A of Table 6 reports the results for Equation (15) via Models 1-3, which correspond to the three measures of analyst superiority, ASUP1-ASUP3, respectively. In this Equation, Hypothesis 3 predicts that REC (the dummy for recent time) and its interaction terms with the intangible variables (REC*RD, REC*AD, and REC*BI) are positive. This is because under the information hypothesis, analyst superiority over extrapolative models should increase with increased complexity of intangible information, and this complexity rises over time. The sign expectations for all other variables are as for Table 5, although the inclusion of REC and its interaction terms is expected to decrease the statistical significance of the stand-alone intangible variables.
As shown in Panel A of Table 6, most significant control variables are in their expected signs: LMV, and LOSS are significant in all three models, STDE and MTB in Model 1, DISP, DAYS, and MKT in Model 3. The exceptional result is DAYS in Model 2, which is significant in the unexpected sign.

Turning to the intangible variables, BI remains significantly positive in Model 1 (p-value = 0.0033). More importantly, as predicted by Hypothesis 3, REC is significantly positive in Model 1 (p-value = 0.0578) and Model 2 (p-value = 0.0197), and REC*BI is significantly positive in Model 1 (p-value = 0.0312), Model 2 (p-value < 0.0001), and Model 3 (p-value = 0.0074). Overall, the Panel A results are consistent with the notion that analyst superiority increases over time, and consistent with Hypothesis 3.

Panel B of Table 6 reports the results for Equation (16) also via three models corresponding to ASUP1-ASUP3, respectively. In this Equation, Hypothesis 3 predicts that TENS (the dummy for intensity of intangible activity) and its interaction terms with the intangible variables (TENS*RD, TENS*AD, and TENS*BI) are positive, because analyst superiority over extrapolative models should increase with increased complexity of intangible information, and this complexity rises with the intensity of intangible activity. The sign expectations for all other variables are as for Panel A.

As shown, most significant control variables are in their expected signs: LMV, and LOSS are significant in all three models, STDE and MTB in Model 1, and DISP, DAYS, and COV in Model 3. Regarding the intangible variables, BI remains significantly positive in Model 1 (p-value = 0.0099). More importantly, as predicted by Hypothesis 3, TENS*BI is

---

23 As discussed before, Model 2 is based on MAFE, which could be problematic due to scaling differences. See more discussions in Footnote 8.
significantly positive in all three models (p-value = 0.0388, 0.0001, and 0.0263, respectively). These results are consistent with the notion that analyst superiority increases as the intensity of intangible activity increases, and consistent with Hypothesis 3.

In sum, all three Hypotheses of this paper are supported by the empirical evidence presented in this Section.

7. Additional Discussions

7.1. Replications

Several replications are performed besides the main reported analyses to assess analysts’ early forecasts. First, replications are performed using FYR2 instead of FYR1 forecasts, as a number of authors also examine FYR2 alongside FYR1 forecasts (for example, Elton et al. 1981). The results show FYR2 forecasts on average (AMESP averages 0.053) are more inaccurate than FYR1 forecasts, but more accurate than extrapolative forecasts. Accordingly, analyst superiority based on FYR2 is reduced compared to those based on FYR1 (ASUP1 averages 0.032 based on FYR2). Nonetheless, replications of Tables 3-6 based on FYR2 forecasts are overall consistent with the tabulated reports, providing support for the paper’s Hypotheses.

Second, the paper’s analyses are replicated with the first analyst forecast issued for a given period end date. The first analyst forecasts are also more accurate on average than extrapolative forecasts (AMESP averages 0.038, and ASUP1 averages 0.927). Replications of Tables 3-6 using the first analyst forecasts are overall consistent with the tabulated reports. While the above two replications are consistent with prior findings that the accuracy of analysts’ forecasts improves as the end of the fiscal year is approached (Crichfield et al.
they show that early analysts’ forecasts are quite good relative to extrapolative models, consistent with the information hypothesis.

Finally, replications are performed using I/B/E/S mean (consensus) forecasts instead of median forecasts, and using Compustat stock price instead of I/B/E/S stock price. These replications also yield consistent results with the tabulated reports. Overall, all replications provide further support for the paper’s Hypotheses.

7.2. **Smoothing Behaviors**

The following considers several behaviors by market participants and their potential effects on the reported results. If any behavior results in reduced variation in forecast error, this could cause the explanatory variables to not fully register in the regression coefficient, which would distort comparisons of these coefficients.

Particularly, analyst herding is well documented in the literature. Analysts are known for modifying their forecasts from their own true beliefs to publish forecasts that are closer to their peers’. The purpose of herding is to reduce deviation from analyst consensus forecasts. However, herding behavior by analysts actually results in increased analysts’ forecast bias (Clement and Tse 2005, Hong and Kubik 2003, and Hong et al. 2000), which generally means increased analyst forecast error and variation in forecast error. Therefore, analyst herding should lead to results contrary to the paper’s Hypotheses, making its conclusion more conservative.

A different and potential behavior by analysts is termed adaptive expectation (Brown and Rozeff 1979), whereby analysts raise (or lower) their forecasts of future earnings when they have underpredicted (overpredicted) current earnings per share. However, this adaptive behavior is not frequently documented empirically. More importantly, one cannot determine
whether this behavior should result in smoothed analyst forecast error, because forecast error depends on both analyst forecast and actual earnings, whereas adaptive behavior generally does not impact actual earnings. Therefore, it does not seem plausible for this behavior to drive the paper’s results.

Finally, a smoothing behavior that should be considered is by company management. According to the accounting literature, company managers tend to smooth their reported earnings to make these numbers more predictable (Graham et al. 2005). This smoothing behavior effectively mitigates extrapolative forecast errors, inflating analyst superiority and causing the explanatory variables to not fully register in the coefficients of the extrapolative regressions. Therefore, this smoothing behavior should lead to results contrary to the paper’s Hypotheses, making its conclusion more conservative.

7.3. Pricing of Intangibles

A related question to consider is whether the market is efficient with respect to intangible information. In an efficient market, a firm’s stock price impounds the value of the firm’s intangibles, so there should be no association between intangibles and future stock returns (Chan et al. 2001). However, the complexity of intangibles raises the possibility that stock prices do not fully incorporate their value. A number of recent studies are consistent with R&D undervaluation, showing that R&D is positively associated with subsequent excess stock returns (Chambers et al. 2002, Penman and Zhang 2002, and Lev and Sougiannis 1996).

An additional analysis of this paper investigates evidence of market inefficiency with respect to intangibles during the examination period. Specifically, excess returns are measured in the month of earnings announcement to investigate any association between
excess returns and intangible variables. Excess returns are computed based on the market model on an estimation period of 12 months prior to the month of earnings announcement.\textsuperscript{24} To adjust returns for market risk, the market model is estimated by firm portfolios that correspond to deciles of firm market capitalization. Excess return is the difference between the actual return in the month of earnings announcement and the returns computed from the market model.

From this investigation, excess returns of intangible-intensive firms, defined as firms with larger industry-adjusted RD than the sample median, are significantly positive (mean = 1.18\%, at p-value < 0.0001), but the returns of other firms are insignificant (mean = -0.2\%, at p-value = 0.1566). Furthermore, there is a significantly positive association between excess returns and intangible intensive firms (RD p-value < 0.0001; AD p-value =0.0705; BI insignificant) after controlling for earnings surprises from both analyst and no-growth extrapolative forecast models and all control variables.\textsuperscript{25} These investigation results are consistent with prior studies that show R&D undervaluation. This pricing investigation provides perspective for viewing the paper’s main results: although analysts are able

\textsuperscript{24} The method to compute excess returns is as in Elton et al. (1981), except that this paper uses 12 instead of 60 months to estimate the parameters of the market model. The shorter estimation period is deemed more suitable for the life cycle of intangibles. The monthly period is short enough to provide information content, yet long enough so that irregularities in the trading process (mispricing of prices due to infrequent trades, or disturbance in prices due to buying or selling pressure in thin markets) are small relative to the true price movements in the period (Rosenberg and Marathe 1979).

\textsuperscript{25} The regression model has the following form:

\[ \text{ExcessReturn}_{it} = \alpha_0 + \alpha_1 \text{RD}_{it} + \alpha_2 \text{AD}_{it} + \alpha_3 \text{BI}_{it} + \alpha_4 \text{LMV}_{it} + \alpha_5 \text{DISP} + \alpha_6 \text{DAYS} + \alpha_7 \text{STDE}_{it} + \alpha_8 \text{LOSS}_{it} + \alpha_9 \text{MTB}_{it} + \alpha_{10} \text{LMV}_{it} + \alpha_{11} \text{COV}_{it} + \alpha_{12} \text{MKT}_{t} + \alpha_{13} \text{ESURP}_{t} + \alpha_{14} \text{ASURP}_{t} + \varepsilon_{it} \]

ESURP is earnings surprise based on the no-growth extrapolative forecast of earnings per share, where earnings surprise = (actual – forecast )/price, ASURP is earnings surprise based on the analyst forecast, and the remaining variables are as in Equation (1).
processors of intangible information, they still have room to improve as information intermediaries due to the complexity of intangible information.

8. Conclusion

This paper continues prior research that investigates the relationship between analyst forecasts and intangible investments, such as research and development, brand promotions and advertising expenses, and recognized intangible assets. The complexity of intangible information and concerns over analysts’ conflicts of interest prompt questions about analysts’ ability to process intangible information. From the paper’s results, the impact of intangibles on extrapolative models’ errors is stronger than on analysts’, analyst superiority over extrapolative models is positively associated with firms’ specific intangibles, and this association increases with increasing complexity of intangibles. These results are relative to many variables that proxy for firm and market information environments, and are controlled for industry effects. The overall findings are consistent with analysts’ ability to process intangible information.
References


Wiedman, C. (1996). The relevance of characteristics of the information environment in the selection of a proxy for the market's expectations for earnings: An extension of


Table 1: Descriptive Statistics

Panel A: Unadjusted

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>25%</th>
<th>Median</th>
<th>75%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analyst Forecast Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMESP</td>
<td>29,023</td>
<td>0.026</td>
<td>0.082</td>
<td>0.002</td>
<td>0.007</td>
<td>0.021</td>
</tr>
<tr>
<td>AMAFE</td>
<td>29,023</td>
<td>0.276</td>
<td>0.517</td>
<td>0.035</td>
<td>0.100</td>
<td>0.270</td>
</tr>
<tr>
<td>AMAPE</td>
<td>28,962</td>
<td>0.554</td>
<td>1.403</td>
<td>0.040</td>
<td>0.122</td>
<td>0.388</td>
</tr>
<tr>
<td>No-Growth Extrapolative Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMESP1</td>
<td>29,023</td>
<td>0.071</td>
<td>0.301</td>
<td>0.008</td>
<td>0.021</td>
<td>0.058</td>
</tr>
<tr>
<td>EMAFE1</td>
<td>29,023</td>
<td>0.758</td>
<td>1.321</td>
<td>0.121</td>
<td>0.313</td>
<td>0.780</td>
</tr>
<tr>
<td>EMAPE1</td>
<td>28,965</td>
<td>1.351</td>
<td>3.391</td>
<td>0.157</td>
<td>0.366</td>
<td>0.986</td>
</tr>
<tr>
<td>With-Growth Extrapolative Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMESP2</td>
<td>22,292</td>
<td>0.190</td>
<td>1.316</td>
<td>0.011</td>
<td>0.028</td>
<td>0.072</td>
</tr>
<tr>
<td>EMAFE2</td>
<td>22,292</td>
<td>1.991</td>
<td>6.205</td>
<td>0.161</td>
<td>0.445</td>
<td>1.161</td>
</tr>
<tr>
<td>EMAPE2</td>
<td>22,254</td>
<td>3.084</td>
<td>11.648</td>
<td>0.210</td>
<td>0.535</td>
<td>1.130</td>
</tr>
<tr>
<td>Analyst Superiority</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASUP1</td>
<td>28,020</td>
<td>2.201</td>
<td>3.106</td>
<td>0.290</td>
<td>2.173</td>
<td>4.129</td>
</tr>
<tr>
<td>ASUP2</td>
<td>28,020</td>
<td>2.205</td>
<td>3.099</td>
<td>0.296</td>
<td>2.172</td>
<td>4.127</td>
</tr>
<tr>
<td>ASUP3</td>
<td>27,929</td>
<td>2.156</td>
<td>3.228</td>
<td>0.101</td>
<td>2.097</td>
<td>4.145</td>
</tr>
<tr>
<td>Intangible Variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RD</td>
<td>29,023</td>
<td>0.027</td>
<td>0.064</td>
<td>0.000</td>
<td>0.000</td>
<td>0.026</td>
</tr>
<tr>
<td>AD</td>
<td>29,023</td>
<td>0.010</td>
<td>0.039</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>BI</td>
<td>29,023</td>
<td>0.105</td>
<td>0.310</td>
<td>0.000</td>
<td>0.008</td>
<td>0.094</td>
</tr>
<tr>
<td>Control Variables/Other Information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LMV</td>
<td>29,023</td>
<td>6.268</td>
<td>1.842</td>
<td>4.94</td>
<td>6.129</td>
<td>7.470</td>
</tr>
<tr>
<td>DISP</td>
<td>24,949</td>
<td>0.158</td>
<td>0.514</td>
<td>0.020</td>
<td>0.045</td>
<td>0.114</td>
</tr>
<tr>
<td>DAYS</td>
<td>29,023</td>
<td>190.387</td>
<td>10.919</td>
<td>195.000</td>
<td>196.000</td>
<td>197.000</td>
</tr>
<tr>
<td>STDE</td>
<td>29,023</td>
<td>0.126</td>
<td>1.575</td>
<td>0.030</td>
<td>0.059</td>
<td>0.126</td>
</tr>
<tr>
<td>LOSS</td>
<td>29,023</td>
<td>0.198</td>
<td>0.398</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>MTB</td>
<td>28,480</td>
<td>2.996</td>
<td>5.394</td>
<td>1.368</td>
<td>2.045</td>
<td>3.312</td>
</tr>
<tr>
<td>COV</td>
<td>29,023</td>
<td>7.832</td>
<td>7.332</td>
<td>2.000</td>
<td>5.000</td>
<td>11.000</td>
</tr>
<tr>
<td>MKT</td>
<td>29,023</td>
<td>0.283</td>
<td>0.139</td>
<td>0.130</td>
<td>0.292</td>
<td>0.420</td>
</tr>
</tbody>
</table>
Panel B: Adjusted by Industry Medians

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>25%</th>
<th>Median</th>
<th>75%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analyst Forecast Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMESP</td>
<td>29,023</td>
<td>0.018</td>
<td>0.081</td>
<td>-0.005</td>
<td>-0.000</td>
<td>0.013</td>
</tr>
<tr>
<td>AMAFE</td>
<td>29,023</td>
<td>0.175</td>
<td>0.517</td>
<td>-0.065</td>
<td>0.000</td>
<td>0.170</td>
</tr>
<tr>
<td>AMAPE</td>
<td>28,962</td>
<td>0.431</td>
<td>1.403</td>
<td>-0.082</td>
<td>0.000</td>
<td>0.265</td>
</tr>
<tr>
<td>No-Growth Extrapolative Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMESP</td>
<td>29,023</td>
<td>0.048</td>
<td>0.301</td>
<td>-0.012</td>
<td>-0.000</td>
<td>0.035</td>
</tr>
<tr>
<td>EMAFE1</td>
<td>29,023</td>
<td>0.445</td>
<td>1.320</td>
<td>-0.192</td>
<td>0.000</td>
<td>0.468</td>
</tr>
<tr>
<td>EMAPE1</td>
<td>28,965</td>
<td>0.985</td>
<td>3.391</td>
<td>-0.208</td>
<td>0.000</td>
<td>0.620</td>
</tr>
<tr>
<td>With-Growth Extrapolative Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMESP2</td>
<td>22,292</td>
<td>0.155</td>
<td>1.280</td>
<td>-0.017</td>
<td>-0.000</td>
<td>0.040</td>
</tr>
<tr>
<td>EMAFE2</td>
<td>22,292</td>
<td>1.517</td>
<td>6.177</td>
<td>-0.274</td>
<td>0.000</td>
<td>0.682</td>
</tr>
<tr>
<td>EMAPE2</td>
<td>22,254</td>
<td>2.519</td>
<td>11.522</td>
<td>-0.310</td>
<td>0.000</td>
<td>0.577</td>
</tr>
<tr>
<td>Analyst Superiority</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASUP1</td>
<td>28,852</td>
<td>1.834</td>
<td>3.330</td>
<td>0.133</td>
<td>1.776</td>
<td>3.471</td>
</tr>
<tr>
<td>ASUP2</td>
<td>28,951</td>
<td>3.174</td>
<td>9.907</td>
<td>0.269</td>
<td>2.028</td>
<td>3.700</td>
</tr>
<tr>
<td>ASUP3</td>
<td>28,920</td>
<td>1.758</td>
<td>3.447</td>
<td>-0.101</td>
<td>1.735</td>
<td>3.447</td>
</tr>
<tr>
<td>Intangible Variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RD</td>
<td>29,023</td>
<td>0.021</td>
<td>0.061</td>
<td>0.000</td>
<td>0.000</td>
<td>0.018</td>
</tr>
<tr>
<td>AD</td>
<td>29,023</td>
<td>0.008</td>
<td>0.037</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>BI</td>
<td>29,023</td>
<td>0.091</td>
<td>0.308</td>
<td>-0.009</td>
<td>0.000</td>
<td>0.079</td>
</tr>
<tr>
<td>Control Variables/Other Information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LMV</td>
<td>29,023</td>
<td>0.126</td>
<td>1.779</td>
<td>-1.110</td>
<td>0.000</td>
<td>1.260</td>
</tr>
<tr>
<td>DISP</td>
<td>24,949</td>
<td>0.110</td>
<td>0.512</td>
<td>-0.024</td>
<td>-0.000</td>
<td>0.066</td>
</tr>
<tr>
<td>DAYS</td>
<td>29,023</td>
<td>-4.825</td>
<td>11.088</td>
<td>-1.000</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>STDE</td>
<td>29,023</td>
<td>0.063</td>
<td>1.574</td>
<td>-0.027</td>
<td>-0.000</td>
<td>0.061</td>
</tr>
<tr>
<td>LOSS</td>
<td>29,023</td>
<td>0.197</td>
<td>0.398</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>MTB</td>
<td>28,480</td>
<td>0.972</td>
<td>5.326</td>
<td>-0.579</td>
<td>0.074</td>
<td>1.256</td>
</tr>
<tr>
<td>COV</td>
<td>29,023</td>
<td>2.620</td>
<td>7.186</td>
<td>-2.000</td>
<td>0.000</td>
<td>5.000</td>
</tr>
<tr>
<td>MKT</td>
<td>29,023</td>
<td>0.283</td>
<td>0.139</td>
<td>0.130</td>
<td>0.292</td>
<td>0.420</td>
</tr>
</tbody>
</table>

AMESP is analyst earnings forecast error with respect to earnings of year $t+1$. It is defined as the absolute difference between median I/B/E/S analysts’ forecast of earnings per share for year $t+1$ and I/B/E/S actual earnings per share of that year. I/B/E/S forecast is issued six months after the end of fiscal year $t$. AMESP is deflated by stock price as of the quarter before the release of analysts’ earnings forecast. The timing of stock price falls between the release of published reports for year $t$ and the release of analyst forecast for year $t+1$. $AMESP_{it+1} = \text{abs}[\text{Forecast I/B/E/S}_{it+1} - \text{Actual I/B/E/S}_{it+1}]/\text{PRICE}_{it}$. AMAFE is analyst absolute error, measured by $\text{AMAFE}_{it+1} = \text{abs}[\text{Forecast I/B/E/S}_{it+1} - \text{Actual I/B/E/S}_{it+1}]$. 

48
ActualIBES-EPS_{t+1}. AMAPE is analyst absolute percentage error, measured by \(\text{abs}[(\text{Forecast IBES-EPS}_{t+1} - \text{Actual IBES-EPS}_{t+1})/ \text{Actual IBES-EPS}_{t+1}]\).

EMESP1 is extrapolative models’ earnings forecast error with respect to earnings of year t+1, defined as the absolute difference between actual earnings per share of year t and year t+1, both as reported in financial statements under income from continuing operations (COI), deflated by the stock price as of the quarter before the release of analysts’ earnings forecast. EMESP1_{it+1} = \text{abs}[\text{COIpershare}_{it} - \text{COIpershare}_{it+1}]/\text{PRICE}_{it}.

EMAFE1 is no-growth extrapolative absolute error, measured by \(\text{abs}[\text{COIpershare}_{it} - \text{COIpershare}_{it+1}]\). EMAPE1 is no-growth extrapolative absolute percentage error, measured by \(\text{abs}[(\text{COIpershare}_{it} - \text{COIpershare}_{it+1})/\text{COIpershare}_{it+1}]\).

EMESP2 is also extrapolative models’ earnings forecast error with respect to earnings of year t+1. The computation is as EMESP1, except that actual earnings per share of year t is multiplied by the earnings growth experienced between year t-1 and year t. FACTOR_{it} = 1+ (\text{COIpershare}_{it} - \text{COIpershare}_{it-1})/\text{COIpershare}_{it-1}. EMESP2_{it+1} = \text{abs}[\text{COIpershare}_{it} \times \text{FACTOR}_{it} - \text{COIpershare}_{it+1}]/\text{PRICE}_{it}. EMAFE2 is with-growth extrapolative absolute error, measured by EMAFE2_{it+1} = \text{abs}[\text{COIpershare}_{it} \times \text{FACTOR}_{it} - \text{COIpershare}_{it+1}]. EMAPE2 is with-growth extrapolative absolute percentage error, measured by EMAPE2_{it+1} = \text{abs}[(\text{COIpershare}_{it} \times \text{FACTOR}_{it} - \text{COIpershare}_{it+1})/\text{COIpershare}_{it+1}].

ASUP1-ASUP3 are proxies of analyst superiority based on the ratio of industry-adjusted extrapolative and analyst forecast errors. ASUP1 = \text{Ln}(\text{EMESP1}_{2}/\text{AMESP}^2), \text{ASUP2} = \text{Ln}(\text{EMAFE1}_{2}/\text{AMAFE}^2), \text{and ASUP3} = \text{Ln}(\text{EMAPE1}_{2}/\text{AMAPE}^2).

RD is R&D expense, AD is advertising expense, and BI is recognized intangibles on the balance sheet, all deflated by equity market value as of the end of year t.

LMV is (the log of) firm size measured as equity market value at the end of year t, DISP is forecast dispersion which is measured as the cross-sectional standard deviation of upcoming annual forecasts deflated by the absolute value of the mean forecast across analysts, DAYS is the number of days between the last fiscal year end date and the forecast date, STDE is the standard deviation of profitability measured as ROA over the years preceding year t+1, LOSS is a dummy variable for loss measured as a negative income before extraordinary items at the end of year t, MTB is market to book value of equity at the end of year t, COV is analyst coverage measured as the number of estimates underlying the I/B/E/S median forecast, and MKT is the annual rate of return of the Dow Jones 5000 index.

In Panel A, all variables are unadjusted for industry. In Panel B, all variables are deviations from the three-digit SIC medians to adjust for industry effects, except for MKT which is a market variable.
Table 2: Correlation Coefficients

<table>
<thead>
<tr>
<th></th>
<th>AMESP</th>
<th>AMAPE</th>
<th>EMESP1</th>
<th>EMAPE1</th>
<th>EMESP2</th>
<th>EMAPE2</th>
<th>ASUP1</th>
<th>ASUP2</th>
<th>ASUP3</th>
<th>RD</th>
<th>AD</th>
<th>BI</th>
<th>LMV</th>
<th>DISF</th>
<th>DAYS</th>
<th>STDE</th>
<th>LOSS</th>
<th>MTR</th>
<th>COV</th>
<th>MKT</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMESP</td>
<td>0.52</td>
<td>0.30</td>
<td>0.24</td>
<td>0.21</td>
<td>0.14</td>
<td>0.07</td>
<td>-0.16</td>
<td>-0.17</td>
<td>0.10</td>
<td>0.12</td>
<td>-0.23</td>
<td>0.08</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.14</td>
<td>-0.04</td>
<td>-0.12</td>
<td>-0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMAPE</td>
<td>0.30</td>
<td>0.48</td>
<td>0.08</td>
<td>0.11</td>
<td>0.26</td>
<td>0.07</td>
<td>-0.21</td>
<td>-0.13</td>
<td>-0.23</td>
<td>0.08</td>
<td>0.04</td>
<td>-0.08</td>
<td>0.09</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.08</td>
<td>-0.04</td>
<td>-0.06</td>
<td>-0.04</td>
<td></td>
</tr>
<tr>
<td>EMESP1</td>
<td>0.12</td>
<td>0.21</td>
<td>0.38</td>
<td>0.14</td>
<td>0.06</td>
<td>0.21</td>
<td>-0.15</td>
<td>-0.08</td>
<td>-0.27</td>
<td>0.06</td>
<td>0.04</td>
<td>0.02</td>
<td>0.18</td>
<td>0.07</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.09</td>
<td>-0.05</td>
<td>-0.11</td>
<td>-0.03</td>
</tr>
<tr>
<td>EMAPE1</td>
<td>0.16</td>
<td>0.10</td>
<td>0.49</td>
<td>0.42</td>
<td>0.15</td>
<td>0.02</td>
<td>0.05</td>
<td>0.15</td>
<td>0.07</td>
<td>0.11</td>
<td>-0.13</td>
<td>0.07</td>
<td>-0.02</td>
<td>0.00</td>
<td>0.12</td>
<td>-0.01</td>
<td>-0.06</td>
<td>-0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMESP2</td>
<td>0.38</td>
<td>0.23</td>
<td>0.50</td>
<td>0.26</td>
<td>0.26</td>
<td>0.08</td>
<td>0.14</td>
<td>0.11</td>
<td>0.04</td>
<td>0.09</td>
<td>0.00</td>
<td>0.10</td>
<td>-0.03</td>
<td>0.01</td>
<td>0.18</td>
<td>-0.02</td>
<td>-0.01</td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMAPE2</td>
<td>0.29</td>
<td>0.21</td>
<td>0.64</td>
<td>0.16</td>
<td>0.03</td>
<td>0.29</td>
<td>0.09</td>
<td>0.04</td>
<td>0.04</td>
<td>-0.10</td>
<td>0.09</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.19</td>
<td>-0.02</td>
<td>-0.05</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMESP2</td>
<td>0.02</td>
<td>0.16</td>
<td>0.11</td>
<td>0.04</td>
<td>0.05</td>
<td>-0.08</td>
<td>0.10</td>
<td>-0.00</td>
<td>0.01</td>
<td>0.23</td>
<td>-0.01</td>
<td>-0.04</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMAPE1</td>
<td>0.03</td>
<td>0.06</td>
<td>0.13</td>
<td>0.03</td>
<td>0.07</td>
<td>-0.03</td>
<td>0.11</td>
<td>0.00</td>
<td>0.02</td>
<td>0.22</td>
<td>-0.01</td>
<td>-0.03</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASUP1</td>
<td>0.33</td>
<td>0.04</td>
<td>0.16</td>
<td>0.11</td>
<td>0.04</td>
<td>0.05</td>
<td>-0.08</td>
<td>0.10</td>
<td>-0.00</td>
<td>0.01</td>
<td>0.23</td>
<td>-0.01</td>
<td>-0.04</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASUP2</td>
<td>0.20</td>
<td>0.49</td>
<td>0.03</td>
<td>-0.00</td>
<td>0.03</td>
<td>0.04</td>
<td>0.01</td>
<td>0.00</td>
<td>0.01</td>
<td>0.09</td>
<td>0.01</td>
<td>0.01</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASUP3</td>
<td>0.17</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.01</td>
<td>0.02</td>
<td>-0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.03</td>
<td>0.00</td>
<td>0.01</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RD</td>
<td>0.01</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.06</td>
<td>0.00</td>
<td>0.02</td>
<td>0.00</td>
<td>0.07</td>
<td>0.01</td>
<td>0.04</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AD</td>
<td>0.04</td>
<td>0.03</td>
<td>-0.19</td>
<td>0.05</td>
<td>0.01</td>
<td>0.00</td>
<td>0.23</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BI</td>
<td>0.09</td>
<td>-0.08</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>-0.02</td>
<td>-0.04</td>
<td>-0.06</td>
<td>-0.03</td>
<td>-0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LMV</td>
<td>-0.03</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>-0.03</td>
<td>-0.04</td>
<td>-0.04</td>
<td>-0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DISF</td>
<td>-0.09</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.28</td>
<td>-0.01</td>
<td>-0.04</td>
<td>-0.03</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAYS</td>
<td>-0.00</td>
<td>0.02</td>
<td>0.02</td>
<td>-0.02</td>
<td>0.02</td>
<td>-0.02</td>
<td>0.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STDE</td>
<td>0.02</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.01</td>
<td>0.03</td>
<td>-0.02</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOSS</td>
<td>0.03</td>
<td>-0.12</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTR</td>
<td>0.06</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MKT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
AMESP is analyst earnings forecast error with respect to earnings of year t +1. It is defined as the absolute difference between median I/B/E/S analysts’ forecast of earnings per share for year t+1 and I/B/E/S actual earnings per share of that year. I/B/E/S forecast is issued six months after the end of fiscal year t. AMESP is deflated by stock price as of the quarter before the release of analysts’ earnings forecast. The timing of stock price falls between the release of published reports for year t and the release of analyst forecast for year t+1. AMESP_{it+1} = abs[Forecast IBES-EPS_{it+1} - Actual IBES-EPS_{it+1}] / PRICE_{it}. AMAFE is analyst absolute error, measured by AMAFE_{it+1} = abs[Forecast IBES-EPS_{it+1} - Actual IBES-EPS_{it+1}]. AMAPE is analyst absolute percentage error, measured by abs[(Forecast IBES-EPS_{it+1} - Actual IBES-EPS_{it+1}) / Actual IBES-EPS_{it+1}].

EMESP1 is extrapolative models’ earnings forecast error with respect to earnings of year t+1, defined as the absolute difference between actual earnings per share of year t and year t+1, both as reported in financial statements under income from continuing operations (COI), deflated by the stock price as of the quarter before the release of analysts’ earnings forecast. EMESP1_{it+1} = abs[COIpershare_{it} - COIpershare_{it+1}] / PRICE_{it}. EMAFE1 is no-growth extrapolative absolute error, measured by abs[COIpershare_{it} - COIpershare_{it+1}]. EMAPE1 is no-growth extrapolative absolute percentage error, measured by abs[(COIpershare_{it} - COIpershare_{it+1}) / COIpershare_{it+1}].

EMESP2 is also extrapolative models’ earnings forecast error with respect to earnings of year t+1. The computation is as EMESP1, except that actual earnings per share of year t is multiplied by the earnings growth experienced between year t-1 and year t. FACTOR_{it} = 1+ (COIpershare_{it} - COIpershare_{it-1}) / COIpershare_{it-1}. EMESP2_{it+1} = abs[COIpershare_{it} * FACTOR_{it} - COIpershare_{it+1}] / PRICE_{it}. EMAFE2 is with-growth extrapolative absolute error, measured by EMAFE2_{it+1} = abs[COIpershare_{it} * FACTOR_{it} - COIpershare_{it+1}] / PRICE_{it}. EMAPE2 is with-growth extrapolative absolute percentage error, measured by EMAPE2_{it+1} = abs[(COIpershare_{it} * FACTOR_{it} - COIpershare_{it+1}) / COIpershare_{it+1}].

ASUP1-ASUP3 are proxies of analyst superiority based on the ratio of industry-adjusted extrapolative and analyst forecast errors. ASUP1 = Ln(EMESP1^2 / AMESP1^2), ASUP2 = Ln(EMAFE1^2 / AMAFE1^2), and ASUP3 = Ln(EMAPE1^2 / AMAPE1^2).

RD is R&D expense, AD is advertising expense, and BI is recognized intangibles on the balance sheet, all deflated by equity market value as of the end of year t.

LMV is (the log of) firm size measured as equity market value at the end of year t, DISP is forecast dispersion which is measured as the cross-sectional standard deviation of upcoming annual forecasts deflated by the absolute value of the mean forecast across analysts, DAYS is the number of days between the last fiscal year end date and the forecast date, STDE is the standard deviation of profitability measured as ROA over the years preceding year t+1, LOSS is a dummy variable for loss measured as a negative income before extraordinary items at the end of year t, MTB is market to book value of equity at the end of year t, COV is analyst coverage measured as the number of estimates underlying the I/B/E/S median forecast, and MKT is the annual rate of return of the Dow Jones 5000 index.

All variables are deviations from the three-digit SIC medians to adjust for industry effects, except MKT which is a market variable.
Table 3: Multivariate Multiple Regression for Comparing Analyst Forecasts with No-Change Extrapolative Forecasts

\[
\begin{align*}
\text{AnalystError}_{it+1} & = \alpha_0 + \alpha_1 \text{RD}_{it} + \alpha_2 \text{AD}_{it} + \alpha_3 \text{BI}_{it} + \alpha_4 \text{LMV}_{it} + \alpha_5 \text{DISP}_{it} + \alpha_6 \text{STDE}_{it} + \alpha_7 \text{LOSS}_{it} + \alpha_8 \text{MTB}_{it} + \alpha_9 \text{COV}_{it} + \alpha_{10} \text{MKT}_{t} + \sum \gamma_k YR_k + \varepsilon_{it} \\
\text{ExtrapolativeError}_{it+1} & = \eta_0 + \eta_1 \text{RD}_{it} + \eta_2 \text{AD}_{it} + \eta_3 \text{BI}_{it} + \eta_4 \text{LMV}_{it} + \eta_5 \text{DISP}_{it} + \eta_6 \text{STDE}_{it} + \eta_7 \text{LOSS}_{it} + \eta_8 \text{MTB}_{it} + \eta_9 \text{COV}_{it} + \eta_{10} \text{MKT}_{t} + \sum \gamma_k YR_k + \varepsilon_{it}
\end{align*}
\]

Panel A: Regression Results

<table>
<thead>
<tr>
<th>Expected Sign†</th>
<th>Analyst</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AMESP</td>
<td>Extrapolative</td>
<td>AMESP</td>
<td>Extrapolative</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.0584</td>
<td>0.0713</td>
<td>0.4302</td>
<td>0.5986</td>
</tr>
<tr>
<td>RD</td>
<td>+</td>
<td>0.1611***</td>
<td>0.5548***</td>
<td>0.5369***</td>
</tr>
<tr>
<td>AD</td>
<td>+</td>
<td>0.1117***</td>
<td>0.2993***</td>
<td>0.3966***</td>
</tr>
<tr>
<td>BI</td>
<td>+</td>
<td>0.0295***</td>
<td>0.0804***</td>
<td>0.0438***</td>
</tr>
<tr>
<td>LMV</td>
<td>-</td>
<td>-0.0077***</td>
<td>-0.0163***</td>
<td>0.0013</td>
</tr>
<tr>
<td>DISP</td>
<td>+</td>
<td>0.0081***</td>
<td>0.0304***</td>
<td>0.0710***</td>
</tr>
<tr>
<td>DAYS</td>
<td>-</td>
<td>-0.0000*</td>
<td>-0.0003</td>
<td>-0.0004</td>
</tr>
<tr>
<td>STDE</td>
<td>+</td>
<td>-0.0001</td>
<td>0.0009</td>
<td>-0.0011</td>
</tr>
<tr>
<td>LOSS</td>
<td>+</td>
<td>0.0092***</td>
<td>0.0462***</td>
<td>0.0500***</td>
</tr>
<tr>
<td>MTB</td>
<td>+/-</td>
<td>-0.0004***</td>
<td>-0.0009*</td>
<td>-0.0091***</td>
</tr>
<tr>
<td>COV</td>
<td>+/-</td>
<td>0.0003***</td>
<td>0.0012***</td>
<td>-0.0030***</td>
</tr>
<tr>
<td>MKT</td>
<td>+/-</td>
<td>-1.4324</td>
<td>-2.3816</td>
<td>-8.1361</td>
</tr>
<tr>
<td>Year Dummies</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>24488</td>
<td>24488</td>
<td>24488</td>
<td>24488</td>
</tr>
<tr>
<td>Adj R-square</td>
<td>0.1043</td>
<td>0.0458</td>
<td>0.0336</td>
<td>0.0641</td>
</tr>
<tr>
<td>Pr &gt; Model F</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Adj R-square

52
### Panel B: Tests Comparing Coefficients of Analyst versus No-Change Extrapolative Regressions

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>F test Comparing RD</td>
<td>138.80*** (&lt;=.0001)</td>
<td>124.96*** (&lt;.0001)</td>
<td>27.45*** (&lt;.0001)</td>
</tr>
<tr>
<td>F test Comparing AD</td>
<td>12.00*** (0.0005)</td>
<td>18.25*** (&lt;.0001)</td>
<td>20.05*** (&lt;.0001)</td>
</tr>
<tr>
<td>F test Comparing IB</td>
<td>66.24*** (&lt;.0001)</td>
<td>99.57*** (&lt;.0001)</td>
<td>11.08*** (0.0009)</td>
</tr>
</tbody>
</table>

P-values are shown at two-tailed values. ***Significant at <.01; ** Significant at <.05; * Significant at <.1.

**AnalystError** is measured via AMESP, AMAFE, and AMAPE. AMESP is analyst earnings forecast error with respect to earnings of year \( t+1 \). It is defined as the absolute difference between median I/B/E/S analysts’ forecast of earnings per share for year \( t+1 \) and I/B/E/S actual earnings per share of that year. I/B/E/S forecast is issued six months after the end of fiscal year \( t \). AMESP is deflated by stock price as of the quarter before the release of analysts’ earnings forecast. The timing of stock price falls between the release of published reports for year \( t \) and the release of analyst forecast for year \( t+1 \). AMESP\(_{t+1} = \text{abs}[\text{Forecast IBES-EPS}_{t+1} - \text{Actual IBES-EPS}_{t+1}]/\text{PRICE}_t$. AMAFE is analyst absolute error, measured by AMAFE\(_{t+1} = \text{abs}[\text{ForecastIBES-EPS}_{t+1} - \text{ActualIBES-EPS}_{t+1}]$. AMAPE is analyst absolute percentage error, measured by \( \text{abs}[(\text{Forecast IBES-EPS}_{t+1} - \text{Actual IBES-EPS}_{t+1})/\text{Actual IBES-EPS}_{t+1}]$.  

**ExtrapolativeError** is measured via no-change extrapolative time-series models (namely EMESP1, EMAFE1, and EMAPE1). EMESP1 is extrapolative models’ earnings forecast error with respect to earnings of year \( t+1 \), defined as the absolute difference between actual earnings per share of year \( t \) and year \( t+1 \), both as reported in financial statements under income from continuing operations (COI), deflated by the stock price as of the quarter before the release of analysts’ earnings forecast. EMESP1\(_{t+1} = \text{abs}[\text{COIpershare}_{t} - \text{COIpershare}_{t+1}]/\text{PRICE}_t$. EMAFE1 is no-growth extrapolative absolute error, measured by \( \text{abs}[\text{COIpershare}_{t} - \text{COIpershare}_{t+1}]$. EMAPE1 is no-growth extrapolative absolute percentage error, measured by \( \text{abs}[(\text{COIpershare}_{t} - \text{COIpershare}_{t+1})/\text{COIpershare}_{t+1}]$.  

RD is R&D expense, AD is advertising expense, and BI is recognized intangibles on the balance sheet, all deflated by equity market value as of the end of year \( t \).

LMV is (the log of) firm size measured as equity market value at the end of year \( t \), DISP is forecast dispersion which is measured as the cross-sectional standard deviation of upcoming annual forecasts deflated by the absolute value of the mean forecast across analysts, DAYS is the number of days between the last fiscal year end date and the forecast date, STDE is the standard deviation of profitability measured as ROA over
the years preceding year t+1, LOSS is a dummy variable for loss measured as a negative income before extraordinary items at the end of year t, MTB is market to book value of equity at the end of year t, COV is analyst coverage measured as the number of estimates underlying the I/B/E/S median forecast, MKT is the annual rate of return of the Dow Jones 5000 index, and YRk are year dummies. All variables are deviations from the three-digit SIC medians to adjust for industry effects, except MKT and YRk.

† These expected signs are expressed for the analyst regressions. In the extrapolative regressions, all variables have the same expected signs as for the analyst regressions, except DAYS which is not related to extrapolative processes and has no sign expectation.
Table 4: Multivariate Multiple Regression for Comparing Analyst Forecasts with Extrapolative Forecasts Incorporating Growth

\[
\begin{align*}
\text{AnalystError}_{it} & = \alpha_0 + \alpha_1 \text{RD}_{it} + \alpha_2 \text{AD}_{it} + \alpha_3 \text{BI}_{it} + \alpha_4 \text{LMV}_{it} + \alpha_5 \text{DISP}_{it} + \alpha_6 \text{DAYS}_{it} + \alpha_7 \text{STDE}_{it} + \alpha_8 \text{LOSS}_{it} + \alpha_9 \text{MTB}_{it} + \alpha_{10} \text{COV}_{it} + \\
& + \alpha_{11} \text{MKT}_{it} + \sum \gamma_k \text{YR}_k + \epsilon_{it}
\end{align*}
\]

<table>
<thead>
<tr>
<th>Panel A: Regression Results</th>
<th>Expected Sign</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Analyst AMESP</td>
<td>Extrapolative EMESP</td>
<td>Analyst AMAFE</td>
<td>Extrapolative EMAFE</td>
</tr>
<tr>
<td>Intercept</td>
<td>+</td>
<td>0.6669 (0.2187)</td>
<td>0.5332 (0.6236)</td>
<td>0.4562 (0.2606)</td>
</tr>
<tr>
<td>RD</td>
<td>-</td>
<td>0.1658*** (&lt;.0001)</td>
<td>2.0362*** (&lt;.0001)</td>
<td>0.5396*** (&lt;.0001)</td>
</tr>
<tr>
<td>AD</td>
<td>+</td>
<td>0.0592*** (&lt;.0001)</td>
<td>0.9960*** (0.0002)</td>
<td>0.2575*** (0.0069)</td>
</tr>
<tr>
<td>BI</td>
<td>+</td>
<td>0.0124*** (&lt;.0001)</td>
<td>0.2590*** (&lt;.0001)</td>
<td>0.0698*** (&lt;.0001)</td>
</tr>
<tr>
<td>LMV</td>
<td>-</td>
<td>-0.0064*** (&lt;.0001)</td>
<td>-0.0405*** (&lt;.0001)</td>
<td>0.0039 (0.1735)</td>
</tr>
<tr>
<td>DISP</td>
<td>+</td>
<td>0.0129*** (&lt;.0001)</td>
<td>0.2909*** (&lt;.0001)</td>
<td>0.1030*** (&lt;.0001)</td>
</tr>
<tr>
<td>DAYS</td>
<td>-</td>
<td>-0.0000* (0.5944)</td>
<td>-0.0005 (0.9101)</td>
<td>-0.0008** (&lt;.0001)</td>
</tr>
<tr>
<td>STDE</td>
<td>+</td>
<td>-0.0000 (0.9022)</td>
<td>0.0151 (0.2176)</td>
<td>-0.0019 (0.6584)</td>
</tr>
<tr>
<td>LOSS</td>
<td>+</td>
<td>0.0146*** (&lt;.0001)</td>
<td>0.5650*** (&lt;.0001)</td>
<td>0.0911*** (&lt;.0001)</td>
</tr>
<tr>
<td>MTB</td>
<td>+/-</td>
<td>-0.0004*** (0.8839)</td>
<td>-0.0004 (0.9000)</td>
<td>-0.0098*** (&lt;.0001)</td>
</tr>
<tr>
<td>COV</td>
<td>+/-</td>
<td>0.0001 (0.1076)</td>
<td>0.0002 (0.9000)</td>
<td>-0.0033*** (&lt;.0001)</td>
</tr>
<tr>
<td>MKT</td>
<td>+/-</td>
<td>-1.6207 (0.3339)</td>
<td>-19.2052 (0.5850)</td>
<td>-8.8723 (0.4779)</td>
</tr>
<tr>
<td>Year Dummies</td>
<td>Yes Yes</td>
<td>Yes Yes Yes Yes</td>
<td>Yes Yes Yes Yes</td>
<td>Yes Yes Yes Yes</td>
</tr>
<tr>
<td>N</td>
<td>19452</td>
<td>19452</td>
<td>19452</td>
<td>19452</td>
</tr>
<tr>
<td>Adj R-square</td>
<td>0.0904</td>
<td>0.0487</td>
<td>0.0386</td>
<td>0.0692</td>
</tr>
<tr>
<td>Pr &gt; Model F</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
Panel B: Tests Comparing Coefficients of Analyst versus No-Change Extrapolative Regressions

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>F test Comparing RD</td>
<td>97.27*** (&gt;.0001)</td>
<td>168.70*** (&lt;.0001)</td>
<td>68.41*** (&lt;.0001)</td>
</tr>
<tr>
<td>F test Comparing AD</td>
<td>12.66*** (&lt;.0008)</td>
<td>1.72 (0.1903)</td>
<td>13.29*** (0.0034)</td>
</tr>
<tr>
<td>F test Comparing IB</td>
<td>33.36*** (&lt;.0001)</td>
<td>40.48*** (&lt;.0001)</td>
<td>20.86*** (&lt;.0001)</td>
</tr>
</tbody>
</table>

P-values are shown at two-tailed values. ***Significant at <0.01; ** Significant at <.05; * Significant at <0.1.

AnalystError is measured via AMESP, AMAFE, and AMAPE. AMESP is analyst earnings forecast error with respect to earnings of year t +1. It is defined as the absolute difference between median I/B/E/S analysts’ forecast of earnings per share for year t+1 and I/B/E/S actual earnings per share of that year. I/B/E/S forecast is issued six months after the end of fiscal year t. AMESP is deflated by stock price as of the quarter before the release of analysts’ earnings forecast. The timing of stock price falls between the release of published reports for year t and the release of analyst forecast for year t+1. AMESPt+1 = abs[Forecast IBES-EPS_t+1 - Actual IBES-EPS_t+1 ]/PRICE_t. AMAFE is analyst absolute error, measured by AMAFE_t+1 = abs[ForecastIBES-EPS_t+1 - ActualIBES-EPS_t+1]. AMAPE is analyst absolute percentage error, measured by abs [(Forecast IBES-EPS_t+1 - Actual IBES-EPS_t+1)/Actual IBES-EPS_t+1].

ExtrapolativeError is measured via extrapolative time-series models incorporating growth (namely EMESP2, EMAFE2, and EMAPE2). EMESP2 is extrapolative models’ earnings forecast error with respect to earnings of year t+1. The computation is as EMESP1, except that actual earnings per share of year t is multiplied by the earnings growth experienced between year t-1 and year t. FACTOR_t = 1+ (COIpershare_t - COIpershare_t-1)/COIpershare_t-1. EMESP2t+1 = abs[COIpershare_t * FACTOR_t - COIpershare_t+1]/PRICE_t. EMAFE2 is with-growth extrapolative absolute error, measured by EMAFE2t+1 = abs[COIpershare_t * FACTOR_t - COIpershare_t+1]. EMAPE2 is with-growth extrapolative absolute percentage error, measured by EMAPE2t+1 = abs[(COIpershare_t * FACTOR_t - COIpershare_t+1)/COIpershare_t+1].

RD is R&D expense, AD is advertising expense, and BI is recognized intangibles on the balance sheet, all deflated by equity market value as of the end of year t.

LMV is (the log of) firm size measured as equity market value at the end of year t, DISP is forecast dispersion which is measured as the cross-sectional standard deviation of upcoming annual forecasts deflated by the absolute value of the mean forecast across analysts, DAYS is the
number of days between the last fiscal year end date and the forecast date, STDE is the standard deviation of profitability measured as ROA over the years preceding year \( t+1 \), LOSS is a dummy variable for loss measured as a negative income before extraordinary items at the end of year \( t \), MTB is market to book value of equity at the end of year \( t \), COV is analyst coverage measured as the number of estimates underlying the I/B/E/S median forecast, MKT is the annual rate of return of the Dow Jones 5000 index, and \( YR_k \) are year dummies.

All variables are deviations from the three-digit SIC medians to adjust for industry effects, except MKT and \( YR_k \).

†These expected signs are expressed for the analyst regressions. In the extrapolative regressions, all variables have the same expected signs as for the analyst regressions, except DAYS which is not related to extrapolative processes and has no sign expectation.
Table 5: Multiple Regression of Analyst Forecast Superiority

\[ \text{ASUP}_{it} = \alpha_0 + \alpha_1 \text{RD}_{it} + \alpha_2 \text{AD}_{it} + \alpha_3 \text{BI}_{it} + \alpha_4 \text{LMV}_{it} + \alpha_5 \text{DISP}_{it} + \alpha_6 \text{DAYS}_{it} + \alpha_7 \text{STDE}_{it} + \alpha_8 \text{LOSS}_{it} + \alpha_9 \text{MTB}_{it} + \alpha_{10} \text{COV}_{it} + \alpha_{11} \text{MKT}_{t} + \sum \gamma_k Y_k + \epsilon_{it} \]

<table>
<thead>
<tr>
<th>Variable</th>
<th>Expected Sign</th>
<th>Model 1 ASUP1</th>
<th>Model 2 ASUP2</th>
<th>Model 3 ASUP3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td>1.0089 (0.7165)</td>
<td>-3.1069 (0.7133)</td>
<td>-2.0215 (0.4868)</td>
</tr>
<tr>
<td>RD</td>
<td>+</td>
<td>1.0178*** (0.0059)</td>
<td>-0.4842 (0.6674)</td>
<td>0.3944 (0.3102)</td>
</tr>
<tr>
<td>AD</td>
<td>+</td>
<td>-0.3023 (0.6143)</td>
<td>-2.3304 (0.2014)</td>
<td>0.1046 (0.8676)</td>
</tr>
<tr>
<td>BI</td>
<td>+</td>
<td>0.2480*** (0.0003)</td>
<td>0.4032* (0.0555)</td>
<td>0.1434** (0.0477)</td>
</tr>
<tr>
<td>LMV</td>
<td>+</td>
<td>0.0999*** (&lt;0.001)</td>
<td>0.1580*** (0.0047)</td>
<td>0.1104*** (&lt;0.0001)</td>
</tr>
<tr>
<td>DISP</td>
<td>-</td>
<td>0.0294 (0.5146)</td>
<td>-0.1810 (0.1886)</td>
<td>-0.1367*** (0.0040)</td>
</tr>
<tr>
<td>DAYS</td>
<td>+</td>
<td>0.0039 (0.1209)</td>
<td>-0.0067 (0.3761)</td>
<td>0.0055** (0.0335)</td>
</tr>
<tr>
<td>STDE</td>
<td>+</td>
<td>0.0421*** (0.0081)</td>
<td>0.0118 (0.7470)</td>
<td>0.0090 (0.4735)</td>
</tr>
<tr>
<td>LOSS</td>
<td>+</td>
<td>0.9408*** (&lt;0.001)</td>
<td>1.1420*** (&lt;0.0001)</td>
<td>0.8717*** (&lt;0.0001)</td>
</tr>
<tr>
<td>MTB</td>
<td>+/-</td>
<td>-0.0131** (0.0310)</td>
<td>0.0045 (0.8055)</td>
<td>-0.0055 (0.3828)</td>
</tr>
<tr>
<td>COV</td>
<td>+/-</td>
<td>-0.0021 (0.6120)</td>
<td>-0.0082 (0.5138)</td>
<td>0.0071 (0.1019)</td>
</tr>
<tr>
<td>MKT</td>
<td>+/-</td>
<td>-56.8070 (0.5067)</td>
<td>109.0458 (0.6754)</td>
<td>52.8406 (0.5550)</td>
</tr>
<tr>
<td>Year Dummies</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td></td>
<td>24349</td>
<td>24433</td>
<td>24406</td>
</tr>
<tr>
<td>Adj R-square</td>
<td>0.0212</td>
<td>0.0040</td>
<td>0.0168</td>
<td></td>
</tr>
<tr>
<td>Pr &gt; Model F</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
</tr>
</tbody>
</table>

P-values are shown at two-tailed values. ***Significant at <0.01; ** Significant at <.05; * Significant at <0.1.

ASUP1-ASUP3 are proxies of analyst superiority based on the ratio of industry-adjusted extrapolative and analyst forecast errors. ASUP1 = Ln(EMESP1^2/AMESP^2), ASUP2 = Ln(EMAFE1^2/AMAFE^2), and ASUP3 = Ln(EMAPE1^2/AMAPE^2).
Extrapolative forecast errors are defined as follows. EMESP1 is extrapolative models’ earnings forecast error with respect to earnings of year \( t+1 \), defined as the absolute difference between actual earnings per share of year \( t \) and year \( t+1 \), both as reported in financial statements under income from continuing operations (COI), deflated by the stock price as of the quarter before the release of analysts’ earnings forecast. EMESP1\(_{it+1}\) = \( \text{abs}\left[\frac{\text{COIpershare}_{it} - \text{COIpershare}_{it+1}}{\text{PRICE}_{it}}\right] \). EMAFE1 is no-growth extrapolative absolute error, measured by abs[COIpershare\(_{it}\) - COIpershare\(_{it+1}\)]. EMAPE1 is no-growth extrapolative absolute percentage error, measured by abs[(COIpershare\(_{it}\) - COIpershare\(_{it+1}\))/COIpershare\(_{it+1}\)].

Analyst forecast errors are defined as follows. AMESP is analyst earnings forecast error with respect to earnings of year \( t+1 \). It is defined as the absolute difference between median I/B/E/S analysts’ forecast of earnings per share for year \( t+1 \) and I/B/E/S actual earnings per share of that year. I/B/E/S forecast is issued six months after the end of fiscal year \( t \). AMESP is deflated by stock price as of the quarter before the release of analysts’ earnings forecast. The timing of stock price falls between the release of published reports for year \( t \) and the release of analyst forecast for year \( t+1 \). AMESP\(_{it+1}\) = \( \text{abs}\left[\frac{\text{Forecast IBES-EPS}_{it+1} - \text{Actual IBES-EPS}_{it+1}}{\text{PRICE}_{it}}\right] \). AMAFE is analyst absolute error, measured by AMAFE\(_{it+1}\) = abs[ForecastIBES-EPS\(_{it+1}\) - ActualIBES-EPS\(_{it+1}\)]. AMAPE is analyst absolute percentage error, measured by abs [(Forecast IBES-EPS\(_{it+1}\) - Actual IBES-EPS\(_{it+1}\))/Actual IBES-EPS\(_{it+1}\)].

RD is R&D expense, AD is advertising expense, and BI is recognized intangibles on the balance sheet, all deflated by equity market value as of the end of year \( t \).

LMV is (the log of) firm size measured as equity market value at the end of year \( t \), DISP is forecast dispersion which is measured as the cross-sectional standard deviation of upcoming annual forecasts deflated by the absolute value of the mean forecast across analysts, DAYS is the number of days between the last fiscal year end date and the forecast date, STDE is the standard deviation of profitability measured as ROA over the years preceding year \( t+1 \), LOSS is a dummy variable for loss measured as a negative income before extraordinary items at the end of year \( t \), MTB is market to book value of equity at the end of year \( t \), COV is analyst coverage measured as the number of estimates underlying the I/B/E/S median forecast, MKT is the annual rate of return of the Dow Jones 5000 index, and YR\(_k\) are year dummies.

All variables are deviations from the three-digit SIC medians to adjust for industry effects, except MKT and YR\(_k\).
Table 6: Multiple Regression of Analyst Forecast Superiority in Interaction with Time and Complexity

Panel A: Interaction with Time

\[
\text{ASUP}_{it} = \alpha_0 + \alpha_1 \text{RD}_{it} + \alpha_2 \text{AD}_{it} + \alpha_3 \text{BI}_{it} + \alpha_4 \text{LMV}_{it} + \alpha_5 \text{DISP}_{it} + \alpha_6 \text{DAYS}_{it} + \alpha_7 \text{STDE}_{it} + \alpha_8 \text{LOSS}_{it} + \alpha_9 \text{MTB}_{it} + \alpha_{10} \text{COV}_{it} + \alpha_{11} \text{MKT}_{it} + \alpha_{12} \text{REC}_{it} + \alpha_{13} \text{REC}_{it} \times \text{RD}_{it} + \alpha_{14} \text{REC}_{it} \times \text{AD}_{it} + \alpha_{15} \text{REC}_{it} \times \text{BI}_{it} + \epsilon_{it}
\]

<table>
<thead>
<tr>
<th>Variable</th>
<th>Expected Sign</th>
<th>Model 1 ASUP1</th>
<th>Model 2 ASUP2</th>
<th>Model 3 ASUP3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td>-0.3282*** (&lt;.0001)</td>
<td>0.5690*** (0.0024)</td>
<td>-0.3606*** (&lt;.0001)</td>
</tr>
<tr>
<td>RD</td>
<td>+</td>
<td>2.7590 (0.3556)</td>
<td>0.5786 (0.9492)</td>
<td>1.7424 (0.5772)</td>
</tr>
<tr>
<td>AD</td>
<td>+</td>
<td>-0.3315 (0.6261)</td>
<td>-1.8029 (0.3824)</td>
<td>0.0571 (0.9361)</td>
</tr>
<tr>
<td>BI</td>
<td>+</td>
<td>0.2174*** (0.0033)</td>
<td>0.1236 (0.5819)</td>
<td>0.0925 (0.2310)</td>
</tr>
<tr>
<td>LMV</td>
<td>+</td>
<td>0.1133*** (&lt;.0001)</td>
<td>0.1658*** (0.0027)</td>
<td>0.1196*** (&lt;.0001)</td>
</tr>
<tr>
<td>DISP</td>
<td>-</td>
<td>0.0209 (0.6433)</td>
<td>-0.1935 (0.1593)</td>
<td>-0.1394*** (0.0033)</td>
</tr>
<tr>
<td>DAYS</td>
<td>+</td>
<td>-0.0008 (0.6789)</td>
<td>-0.0109* (0.0619)</td>
<td>0.0043** (0.0306)</td>
</tr>
<tr>
<td>STDE</td>
<td>+</td>
<td>0.0435*** (0.0062)</td>
<td>0.0145 (0.6915)</td>
<td>0.0093 (0.4588)</td>
</tr>
<tr>
<td>LOSS</td>
<td>+</td>
<td>0.9600*** (&lt;.0001)</td>
<td>1.1609*** (&lt;.0001)</td>
<td>0.8912*** (&lt;.0001)</td>
</tr>
<tr>
<td>MTB</td>
<td>+/-</td>
<td>-0.0126** (0.0376)</td>
<td>0.0050 (0.7666)</td>
<td>-0.0047 (0.4577)</td>
</tr>
<tr>
<td>COV</td>
<td>+/-</td>
<td>-0.0060 (0.1434)</td>
<td>-0.0106 (0.3925)</td>
<td>0.0043 (0.3084)</td>
</tr>
<tr>
<td>MKT</td>
<td>+/-</td>
<td>0.5074** (0.0408)</td>
<td>0.3606 (0.6315)</td>
<td>0.7312*** (0.0047)</td>
</tr>
<tr>
<td>REC</td>
<td>+</td>
<td>0.1261* (0.0587)</td>
<td>0.4723** (0.0197)</td>
<td>0.0608 (0.3831)</td>
</tr>
<tr>
<td>REC*RD</td>
<td>+</td>
<td>-1.9315 (0.5244)</td>
<td>-1.9969 (0.8287)</td>
<td>-1.6231 (0.6092)</td>
</tr>
<tr>
<td>REC*AD</td>
<td>+</td>
<td>-0.4804 (0.7319)</td>
<td>-3.9712 (0.3515)</td>
<td>-0.0722 (0.9607)</td>
</tr>
<tr>
<td>REC*BI</td>
<td>+</td>
<td>0.4063** (0.0312)</td>
<td>2.3643*** (&lt;.0001)</td>
<td>0.5277*** (0.0074)</td>
</tr>
<tr>
<td>N</td>
<td></td>
<td>24,349</td>
<td>24,433</td>
<td>24,406</td>
</tr>
<tr>
<td>Adj R-square</td>
<td></td>
<td>0.0184</td>
<td>0.0037</td>
<td>0.0152</td>
</tr>
<tr>
<td>Pr &gt; Model F</td>
<td></td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
Panel B: Interaction with Intensity of Intangible Activity

\[ \text{ASUP}_{it} = \alpha_0 + \alpha_1 \text{RD}_{it} + \alpha_2 \text{AD}_{it} + \alpha_3 \text{BI}_{it} + \alpha_4 \text{LMV}_{it} + \alpha_5 \text{DISP}_{it} + \alpha_6 \text{DAYS}_{it} + \alpha_7 \text{STDE}_{it} + \alpha_8 \text{LOSS}_{it} + \alpha_9 \text{MTB}_{it} + \alpha_{10} \text{COV}_{it} + \alpha_{11} \text{MKT}_{t} + \alpha_{12} \text{COM}_{it} + \alpha_{13} \text{COM}_{it}^* \text{RD}_{it} + \alpha_{14} \text{COM}_{it}^* \text{AD}_{it} + \alpha_{15} \text{COM}_{it}^* \text{BI}_{it} + \Sigma \gamma_k \text{YR}_k + \varepsilon_{it} \]

<table>
<thead>
<tr>
<th>Variable</th>
<th>Expected Sign</th>
<th>Model 1 ASUP1</th>
<th>Model 2 ASUP2</th>
<th>Model 3 ASUP3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td>1.0149 (0.7149)</td>
<td>-3.1340 (0.7109)</td>
<td>-2.0506 (0.4805)</td>
</tr>
<tr>
<td>RD</td>
<td>+</td>
<td>2.0688 (0.4965)</td>
<td>-3.0176 (0.7446)</td>
<td>1.1833 (0.7102)</td>
</tr>
<tr>
<td>AD</td>
<td>+</td>
<td>-0.3066 (0.6540)</td>
<td>-1.6321 (0.4322)</td>
<td>0.1077 (0.8805)</td>
</tr>
<tr>
<td>BI</td>
<td>+</td>
<td><strong>0.1921</strong>* (0.0099)</td>
<td>0.1005 (0.6570)</td>
<td>0.0836 (0.2831)</td>
</tr>
<tr>
<td>LMV</td>
<td>+</td>
<td><strong>0.0967</strong>* (&lt;.0001)</td>
<td>0.1387** (0.0138)</td>
<td>0.1049*** (&lt;.0001)</td>
</tr>
<tr>
<td>DISP</td>
<td>-</td>
<td>0.0300 (0.5070)</td>
<td>-0.1783 (0.1954)</td>
<td>-0.1372*** (0.0039)</td>
</tr>
<tr>
<td>DAYS</td>
<td>+</td>
<td>0.0037 (0.1417)</td>
<td>-0.0070 (0.3587)</td>
<td>0.0053** (0.0407)</td>
</tr>
<tr>
<td>STDE</td>
<td>+</td>
<td><strong>0.0422</strong>* (0.0080)</td>
<td>0.0114 (0.7566)</td>
<td>0.0089 (0.4785)</td>
</tr>
<tr>
<td>LOSS</td>
<td>+</td>
<td><strong>0.9411</strong>* (&lt;.0001)</td>
<td>1.1262*** (&lt;.0001)</td>
<td>0.8665*** (&lt;.0001)</td>
</tr>
<tr>
<td>MTB</td>
<td>+/-</td>
<td>-0.0125** (0.0421)</td>
<td>0.0061 (0.7409)</td>
<td>-0.0058 (0.3668)</td>
</tr>
<tr>
<td>COV</td>
<td>+/-</td>
<td>-0.0017 (0.6748)</td>
<td>-0.0062 (0.6238)</td>
<td>0.0075* (0.0833)</td>
</tr>
<tr>
<td>MKT</td>
<td>+/-</td>
<td>-56.9078 (0.5060)</td>
<td>108.8258 (0.6759)</td>
<td>52.8748 (0.5547)</td>
</tr>
<tr>
<td>TENS</td>
<td>+</td>
<td>-0.0005 (0.9925)</td>
<td>0.0937 (0.5644)</td>
<td>0.0714 (0.2021)</td>
</tr>
<tr>
<td>TENS*RD</td>
<td>+</td>
<td>-1.2843 (0.6754)</td>
<td>1.1905 (0.8985)</td>
<td>-1.3527 (0.6735)</td>
</tr>
<tr>
<td>TENS*AD</td>
<td>+</td>
<td>-0.1816 (0.8984)</td>
<td>-3.9635 (0.3599)</td>
<td>-0.2283 (0.8780)</td>
</tr>
<tr>
<td>TENS*BI</td>
<td>+</td>
<td><strong>0.4017</strong>* (0.0388)</td>
<td><strong>2.2456</strong>* (0.0001)</td>
<td><strong>0.4511</strong>* (0.0263)</td>
</tr>
</tbody>
</table>

Year Dummies | Yes |
N | 24349 |
Adj R-square | 0.0225 |
Pr > Model F | <.0001 |
P-values are shown at two-tailed values. ***Significant at <0.01; ** Significant at <.05; * Significant at <0.1.

ASUP1-ASUP3 are proxies of analyst superiority based on the ratio of industry-adjusted extrapolative and analyst forecast errors. ASUP1 = Ln(EMESP1²/AMESP²), ASUP2 = Ln(EMAFE1²/AMAFE²), and ASUP3 = Ln(EMAPE1²/AMAPE²).

Extrapolative forecast errors are defined as follows. EMESP1 is extrapolative models’ earnings forecast error with respect to earnings of year t+1, defined as the absolute difference between actual earnings per share of year t and year t+1, both as reported in financial statements under income from continuing operations (COI), deflated by the stock price as of the quarter before the release of analysts’ earnings forecast. EMESP1\_{t+1} = abs [COIpershare\_t - COIpershare\_t+1]/PRICE\_t. EMAFE1 is no-growth extrapolative absolute error, measured by abs[COIpershare\_t - COIpershare\_t+1]. EMAPE1 is no-growth extrapolative absolute percentage error, measured by abs[(COIpershare\_t - COIpershare\_t+1)/COIpershare\_t+1].

Analyst forecast errors are defined as follows. AMESP is analyst earnings forecast error with respect to earnings of year t+1. It is defined as the absolute difference between median I/B/E/S analysts’ forecast of earnings per share for year t+1 and I/B/E/S actual earnings per share of that year. I/B/E/S forecast is issued six months after the end of fiscal year t. AMESP is deflated by stock price as of the quarter before the release of analysts’ earnings forecast. The timing of stock price falls between the release of published reports for year t and the release of analyst forecast for year t+1. AMESP\_{t+1} = abs [Forecast IBES-EPS\_t+1 - Actual IBES-EPS\_t+1]/PRICE\_t. AMAFE is analyst absolute error, measured by AMAFE\_{t+1} = abs[ForecastIBES-EPS\_t+1 - ActualIBES-EPS\_t+1]. AMAPE is analyst absolute percentage error, measured by abs [(Forecast IBES-EPS\_t+1 - Actual IBES-EPS\_t+1)/ Actual IBES-EPS\_t+1].

RD is R&D expense, AD is advertising expense, and BI is recognized intangibles on the balance sheet, all deflated by equity market value as of the end of year t.

LMV is (the log of) firm size measured as equity market value at the end of year t, DISP is forecast dispersion which is measured as the cross-sectional standard deviation of upcoming annual forecasts deflated by the absolute value of the mean forecast across analysts, DAYS is the number of days between the last fiscal year end date and the forecast date, STDE is the standard deviation of profitability measured as ROA over the years preceding year t+1, LOSS is a dummy variable for loss measured as a negative income before extraordinary items at the end of year t, MTB is market to book value of equity at the end of year t, COV is analyst coverage measured as the number of estimates underlying the I/B/E/S median forecast, MKT is the annual rate of return of the Dow Jones 5000 index, and YR\_k are year dummies. REC is a dummy variable equal 1 if the year is later than 1998, and 0 otherwise. TENS is a dummy variable equal 1 if industry-adjusted RD is larger than the industry median, and 0 otherwise.

All variables are deviations from the three-digit SIC medians to adjust for industry effects, except MKT, YR\_k, REC, TENS, and the interaction terms.