Review of Accounting Studies (2023) 28:1263–1315
https://doi.org/10.1007/s11142-023-09791-8

Analyst ability and research effort: non-EPS forecast provision as a research quality signal

Peter F. Pope[1] · **Tong Wang**[2]

Accepted: 18 June 2023 / Published online: 11 September 2023

Abstract
The range of non-EPS forecast types provided by individual analysts to I/B/E/S has increased dramatically over time but varies considerably across firms. We propose that in providing a broader range of forecast types, analysts can signal superior research ability and research effort. Consistent with this hypothesis, we document positive associations between the number of forecast types (NFT) an analyst provides and common proxies for research quality, including earnings forecast accuracy, price target accuracy, stock recommendation profitability, market reactions to stock recommendation revisions, and analyst career outcomes. The effects of NFT are incremental to other quality proxies used in the literature and are distinct from the issuance of specific non-EPS forecast types studied previously (e.g., cash flow forecasts). We demonstrate the information value of NFT to investors by examining the out-of-sample performance of portfolios formed conditional on NFT and exploiting revisions in consensus earnings forecasts and individual analysts' stock recommendations. We conclude that the number of forecast types an analyst provides for a firm is a readily available proxy for the quality of her research.

Keywords Analysts · Earnings forecasts · Forecasting ability · I/B/E/S · Price target forecasts · Stock recommendations · Capital market reactions · Analyst career outcomes

JEL Classification G24 · M41

✉ Peter F. Pope
peter.pope@unibocconi.it

1 Department of Accounting and the Baffi Centre on Economics, Finance and Regulation, Bocconi University, Milan, Italy

2 Barclays Bank plc, London, UK

1 Introduction

The last two decades have seen a dramatic increase in the frequency with which sell-side analysts disclose forecasts of financial statement items other than earnings per share (EPS). The Institutional Brokers Estimate System (I/B/E/S) currently captures 23 forecast types (e.g., cash flow, dividends, book value, capital expenditure, gross margin, operating income, pre-tax income) other than EPS forecasts, price target forecasts, and stock recommendations (Beyer et al. 2010). However, the provision of forecasts is not uniform. Individual analysts provide different numbers of non-EPS forecast types across the firms they cover and for the same firm over time. Beyer et al. (2010) comment that the "decision to forecast only a subset of firms' fundamentals is even more puzzling given the recent findings that analysts rely on some of these measures (e.g., earnings forecasts) to forecast others (e.g., recommendations)" (p.330), and they call for more research on this issue. Recently, Hand et al. (2021) and Bilinski (2020) show that non-EPS forecast surprises are reflected in earnings announcement returns, suggesting that non-EPS forecasts are incrementally useful to investors. In this paper, we present evidence that the extent of provision of non-EPS forecast types captures information about the quality of analysts' research at the firm-year level. When we examine a range of analyst research outcomes studied in the prior literature, we find consistent evidence that research quality is higher when analysts provide more non-EPS forecast types for a firm. We demonstrate the potential implications of our findings for investors by examining the out-of-sample performance of calendar time portfolios formed using analysts' stock recommendations and exploiting information on the number of forecast types provided by analysts.

Our empirical analysis uses a comprehensive set of individual analysts' forecasts of 23 different financial statement line items listed in Data Appendix A at different horizons ("forecast types") in the I/B/E/S Detail US database for the years 2001–2020. We first document substantial within-firm-year variation in the number of forecast types provided by individual analysts to I/B/E/S. This variation is not fully explained by individual analysts having consistent forecasting styles that produce the same set of forecast types across the firms they cover. Rather, analysts provide different numbers of forecast types for different firms and for the same firm in different years. For each analyst and each firm-year, we count the number of forecast types provided in the three-month period before an earnings or price target forecast or a recommendation is issued and then scale the count relative to other analysts covering the same firm during the same period. Our measure is designed to explain variation in analysts' research quality within each firm-year.

We first assess the extent to which the scaled number of forecast types is associated with other proxies for analyst ability and effort studied in the prior literature. These proxies include the ability of an analyst to move consensus earnings forecasts towards actual earnings realizations, a proxy for innate analyst ability (Chen and Jiang 2006; Ertimur et al. 2011); earnings forecasting frequency, a measure of analyst effort (Jacob et al. 1999); and earnings forecasting timeliness, a robust identifier of lead analysts (Cooper et al. 2001; Shroff et al. 2014). Interestingly, we find that less experienced analysts, for whom a long history of research outcomes is unavailable, provide more forecast types than their more experienced peers. However, regressions including only broker or analyst or analyst-firm fixed effects explain substantially more variation in the number of forecast types than these proxies for ability and effort. We interpret these results as indicating that the number of forecast types captures analyst traits (such as innate ability or education and training) (Clement et al. 2007) or analyst effort choices, either of which can contribute to analyst research quality.

Consistent with the number of forecast types reflecting analyst ability and research effort, we find that earnings (price target) forecasts accompanied by more forecast types are 4.90% (1.90%) more accurate than those accompanied by fewer forecast types. Similarly, the market reaction to revisions in analysts' stock recommendations is 34% higher when analysts have provided more non-EPS forecasts in the period before a recommendation revision. We also quantify the value of expanded forecast provision to investors by examining the ex post profitability of stock recommendations. In sample, recommendations are more profitable when the number of forecast types supplied by an analyst is relatively high, earning incremental annualized excess returns of 1.97%. Finally, consistent with the notion that greater non-EPS forecast provision reflects research quality, we find that analysts' career outcomes are positively related to the propensity of an analyst to provide more forecasts.

Out-of-sample portfolio tests avoiding possible look-ahead biases and making conservative assumptions about the timing of trades confirm the potential investment value of restricting earnings consensus revisions and stock recommendations to analysts who provide more non-EPS forecast types for a firm. A hedge portfolio exploiting revisions in EPS forecasts accompanied by a high number of forecast types earns an annualized excess return of 10.75%, compared to 0.66% when earnings forecasts are accompanied by low numbers of forecast types. Similarly, a hedge portfolio formed using individual analysts' stock recommendations accompanied by high numbers of forecast types earns average excess returns of 22.15% on an annualized basis, compared to 13.41% for the recommendations accompanied by fewer forecast types. Further analysis reveals that the superior out-of-sample performance of signals using inputs associated with high numbers of forecast types is robust to controlling for risk using the Fama and French (2015) model.

Our findings are robust to a comprehensive set of control variables using a conservative research design that controls for analyst characteristics and brokerage house fixed effects and focuses on variation in non-EPS forecast provision within firm-years. Importantly, the results are still valid after controlling for firm-analyst fixed effects, suggesting that analysts provide more forecast types when they have exerted higher effort, or that they disclose strategically when they are more confident in the quality of their research.[1]

Our paper extends the recent literature on analysts' non-EPS forecasts in several directions. One strand of the prior literature focuses on analyst provision of cash flow

[1] An analyst provides an item to I/B/E/S if and only if she 1) forecasts the item and 2) chooses to report the item to I/B/E/S. It is not possible to distinguish between the "forecasting" and "reporting" activities of analysts because the disclosure from sell-side financial analysts is voluntary. An analyst can always selectively report forecast types unless her brokerage has specific disclosure policies, which are not observable to outsiders. In other words, although an analyst can privately forecast an item, she might not report it to I/B/E/S. To avoid overgeneralizing the conclusions of this paper, we do not seek to disentangle "forecasting" from "reporting."

forecasts. For example, Call et al. (2009) document that analysts' earnings forecasts are more accurate when accompanied by cash flow forecasts, interpreting their results as suggesting that analysts better understand the dynamics of earnings and adopt a more disciplined approach to forecasting earnings as a result of forecasting cash flow. Hashim and Strong (2018) make similar arguments to explain their finding that analysts' price target forecasts are more accurate when accompanied by cash flow forecasts. We contribute to this stream of literature by establishing the association between the accuracy of earnings forecasts and price target forecasts and analysts' provision of a wider set forecasts beyond cash flow forecasts.

Our findings also contribute to a better understanding of the determinants of stock recommendations' profitability—a relatively underresearched area (Bradshaw et al. 2013). Loh and Mian (2006) and Ertimur et al. (2007) document that the profitability of analysts' stock recommendations is positively linked to the accuracy of their earnings forecasts. In our tests, the superior forecasting ability of analysts who provide more forecast types for EPS and price targets forecasts is mirrored in their stock recommendations. When accompanied by more forecast types, stock recommendations earn higher annualized market-adjusted returns of 1.97% over holding periods of up to 180 days. Furthermore, we show that role of expanded forecast provision in predicting stock recommendation profitability is incremental to the accuracy of earnings forecasts, the most important input to analysts' valuation models (Bradshaw 2004; Ertimur et al. 2007; Loh and Mian 2006).

If recommendations accompanied by more forecast types are perceived by the market to be of higher quality, the market reaction to analysts' recommendation revisions should be more pronounced when the analyst provides more forecast types.[2] Jung et al. (2012) find that stock markets react less strongly to stock recommendation revisions accompanied by cash flow forecasts, a result that is inconsistent with the notion that cash flow forecasts signal better forecasting ability. However, they also find higher market reaction to stock recommendation revisions when the revisions are accompanied by long term growth forecasts. We find stronger stock market reaction to recommendation revisions when an analyst reports more forecast types. Our results suggest that analysts who provide more forecast types have higher stock market credibility.

Finally, focusing on the labor market, we find that analysts who provide more forecast types than their peers are significantly less likely to be fired and more likely to be promoted from smaller to larger brokerage houses. These results are also consistent with the number of non-EPS forecast types capturing information about the innate ability and research effort of analysts. Our results therefore extend the results of Ertimur et al. (2011) and Jung et al. (2012), who find that analysts' provision of, respectively, revenue and long-term earnings growth forecasts is positively related to career outcomes.

Our results also contribute to the literature considering the motivation of analysts in providing non-EPS forecasts. For example, Keung (2010) argues that analysts are

[2] In prior research, for example, Stickel (1995) documents that investors' reactions to recommendation revisions are a function of the reputation of the analyst, the size of the brokerage house, the size of the focal firm, and the accompanying earnings forecast revisions. Loh and Stulz (2011) document that recommendation revisions are more likely to be influential if they are from leader, star, or previously influential analysts, issued away from consensus, and accompanied by earnings forecasts.

more likely to issue earnings forecasts and sales forecasts simultaneously when they are better informed; Jung et al. (2012) argue that the provision of long-term earnings growth forecasts signals analysts' ability and effort to analyze firms' long-term prospects; and Ertimur et al. (2011) provide evidence that analysts issue sales forecasts as a means of establishing reputation. Our results are consistent with two non-mutually exclusive interpretations. First, expanded forecast provision could signal higher quality of an analyst's research, reflecting both her research ability and her research effort. As such, our results can be viewed as generalizing Ertimur et al. (2011)'s conclusion that provision of revenue forecasts can help establish reputation. Alternatively, if an analyst devotes the same effort in researching all of the companies she follows, she could selectively provide expanded forecasts to strategically signal her confidence in her research outputs, reports, and recommendations.

From the perspective of users of sell-side research, our results suggest that the number of forecast types is an informative yet parsimonious ex ante measure of an analyst's research quality and forecasting ability that can be applied to most firms and analysts in the I/B/E/S universe.[3] Estimation does not require a long time series of data, unlike other quality proxies such as historical forecast accuracy, forecasting frequency, and forecasting timeliness. Therefore, it could be particularly useful for understanding the research quality of new analysts without a long track record. Our results also suggest that using our expanded forecast type metric, investors can identify and put more weight on more accurate forecasts and recommendations and realize higher investment returns. The measure also has potential in developing better expectations proxies from individual analysts' forecasts (Clement 1999).

The remainder of the paper is organized as follows. Section 2 discusses the prior literature on non-EPS forecasts and develops the hypotheses; Section 3 introduces the research design for the main tests; Section 4 describes the sample and descriptive statistics for the variable of interest—the number of forecast types; Section 5 investigates the associations between the number of forecast types and analyst characteristics; Section 6 reports and discusses the main empirical results; Section 7 reports and discusses the results of the out-of-sample analysis; Section 8 reports the results of additional tests; and Section 9 concludes.

2 Prior research and hypotheses

Earnings forecasts have been studied widely by researchers and are a key input to the accounting-based equity valuation models used by investment professionals (Bradshaw 2004; Ertimur et al. 2007; Loh and Mian 2006). Forecasting fundamentals other than earnings can help improve the quality of research because the forecasting process can enhance an analyst's understanding of a firm's operating, investing, and financing activities (Lundholm and Sloan 2004). Recent empirical evidence also confirms

[3] In contrast, if the issuance of cash flow forecasts is used to identify superior analysts, then analysts of firms without cash flow forecast coverage will be ignored, even if they are providing a range of other non-EPS forecasts.

that a broader set of non-EPS forecasts are value-relevant and used by investors.[4] For example, Bilinski (2020) and Hand et al. (2021) report that non-EPS forecast surprises have incremental explanatory power beyond EPS forecast surprises in explaining earnings announcement returns. These results suggest that sell-side analysts face investor demand for non-EPS forecasts.[5]

Meeting investor demand by providing forecasts of fundamental outcomes may benefit analysts, but it is also costly for at least two reasons. First, forecasting additional line items increases analysts' task complexity and necessitates costly research effort. These costs will be higher for analysts with lower innate ability. Second, reporting more forecast types exposes analysts to greater ex post scrutiny of the quality of their work by investors and brokerage firms' line managers. If forecasts by less capable and less diligent analysts are relatively inaccurate ex post, voluntary forecast provision carries a risk of reputation loss when the inaccuracy is revealed.[6] If the expected costs of non-EPS forecast disclosure outweigh the benefits for less capable and less diligent analysts, we expect lower non-EPS forecast provision by these analysts. Hence, we predict that by providing more non-EPS forecast types, analysts signal higher ability, effort, and quality of research.

We test our predictions by examining a comprehensive set of outcomes used in the prior literature to assess the quality of analysts' research. These outcomes include earnings forecast accuracy, price target forecast accuracy, profitability of stock recommendations, stock market reactions to stock recommendation revisions, and analyst career outcomes. Earnings forecasts are important intermediate products that are generated in pursuit of the ultimate product of an analyst's research effort—stock recommendations (Bradshaw 2011; Ertimur et al. 2007; Loh and Mian 2006; Schipper 1991). Consistent with earnings forecast accuracy being relevant in evaluating analysts' research performance, it is associated with analyst turnover (Mikhail et al. 1999) and is a metric used in Starmine analyst ratings. Price target forecasts quantify the valuations underlying an analyst's stock recommendations based on estimated input forecasts and indicate, at a granular level, analysts' ability to predict expected returns (Asquith et al. 2005; Brav and Lehavy 2003). The profitability of stock recommendations is a direct estimate of the ex post value of analysts' research for investors and is known to play a significant role in analysts' compensation (Brown et al. 2015; Groysberg et al. 2011). Stock market reactions to revisions in analysts' stock

4 Corporate managers may also learn from analysts' non-EPS forecasts. See Choi et al. (2020) for capital expenditure forecasts and Bratten et al. (2017) for pre-tax earnings forecasts.

5 Analysts' compensation partially depends on their ability to generate trading commissions (Brown et al. 2015; Groysberg et al. 2011), so analysts also have incentives to provide an expanded set of forecast types if they stimulate investor trading, independent of their value-relevance. Consistent with this view, a recent paper by Harford et al. (2019) also suggests that analysts produce higher-quality earnings forecasts and recommendations for firms that are more economically important to their brokerage firms and hence for their careers. While we do not rule out firm-specific incentives also affecting analyst behavior, our research design controls for inter-firm variation driven by firm characteristics and focuses on within-firm variation in the quality of research and research effort.

6 Although analysts' forecasting ability can grow with experience (Clement et al. 2007; Mikhail et al. 1997; 2003), a common assumption is that forecasting ability is an innate characteristic. Ability is not the only factor affecting analysts' forecasting performance. There is substantial variation in forecasting behavior by individual analysts over time and across firms, suggesting that their research effort varies.

recommendations provide an ex ante indication of whether the market views stock market recommendations as informative. Finally, if analysts ultimately pursue a successful career, terminations involving moves to smaller brokerage houses (or promotions involving moves to larger brokerage houses) can be interpreted as negative (positive) indicators of individual analysts' research quality, under the assumption that such outcomes, on average, are merit based and reflect analysts' research ability and effort (Call et al. 2009).

3 Research design

To test whether the extent of non-EPS forecast provision captures research quality, we examine whether earnings forecast accuracy, price target forecast accuracy, stock recommendation profitability, market reactions to recommendation revisions, and career outcomes depend on the number of forecast types provided by an analyst, denoted by "NFT."

3.1 Measuring NFT and its components

NFT is based on a count of the number of financial statement line item forecasts over multiple horizons that an analyst provides to the annual detail I/B/E/S database (maximum 23).[7] Besides financial statement line items, analysts also provide stock recommendations and price targets. We exclude stock recommendations and price targets in the calculation of NFT because they are the ultimate products of an analyst's research (Schipper 1991) and two of the proxies for analyst forecasting performance that we test in this paper. To capture the time-varying and slow-moving properties of NFT, we calculate NFT on a three-month rolling window at monthly frequency. We choose a three-month window because analysts usually do not release every forecast type simultaneously.[8]

Specifically, we define NFT as the total number of unique forecast types that analyst j has provided to I/B/E/S for firm i during the three calendar months *prior to* month t. Formally $\mathrm{NFT}_{\mathrm{ijt}} = \sum_{\mathrm{h}=0}^{5} \sum_{\mathrm{k}=1}^{23} \mathrm{I\_AF}_{\mathrm{ijt}}^{\mathrm{k,h}}$, where $\mathrm{I\_AF}_{\mathrm{ijt}}^{\mathrm{k,h}}$ is an indicator variable that equals one if analyst j has provided at least one h-year-ahead type k forecast for firm i during the three-month window prior to month t, zero otherwise. We denote long-term growth forecast types as h = 0. Except for long-term earnings growth forecasts, we ignore any forecasts with a forecast period exceeding five years because they are scarce in I/B/E/S.[9] Note that we ignore the second and later forecasts of the same type for the same horizon when measuring NFT. We consider only forecasts issued prior to month t in calculating NFT, to mitigate concerns that non-EPS forecast provision could affect analyst research performance directly due to information conveyed by

7 I/B/E/S also collects key performance indicator (KPI) forecasts and forecasts at the geographic, product, and segment levels from analysts. Future research could consider these other forecasts.

8 We obtain qualitatively similar results when using a six-month or one-year window.

9 Including quarterly forecasts, or forecasts of horizons over five years, does not qualitatively change our results.

the additional forecasts.[10] We construct NFT for each analyst on a firm-specific basis because analysts provide different numbers of forecast types for different firms that they cover and in different years.

In additional tests reported in Section 8.2, we seek a better understanding of the sources of information in NFT by decomposing NFT into a component capturing the breadth of forecast type provision and a component capturing the horizon over which an analyst is forecasting. We define the breadth component as the number of unique one-year-ahead forecast types provided by analyst j for firm i during the three months prior to the month in which the dependent variable is measured, and denote this variable NFT1. The forecast type with the maximum horizon length is typically EPS. We define the horizon component of NFT as the number of unique EPS forecast horizons, including long-term earnings growth (LTG), issued by analyst j for firm i during the three months prior to the month in which the dependent variable is measured. We denote this variable NFT_HZN.[11]

3.2 Scaling NFT and other characteristics

To facilitate interpretation of regression coefficients across models, we scale NFT and all other non-indicator control variables using the range transformation introduced by Clement and Tse (2003, 2005) and defined as follows:

$$\text{Characteristic\_R}_{\text{ijt}} = \frac{\text{Characteristic}_{\text{ijt}} - \min_{\text{it}}(\text{Characteristic})}{\max_{\text{it}}(\text{Characteristic}) - \min_{\text{it}}(\text{Characteristic})}, \quad (1)$$

where Characteristic_R is the scaled value of the original raw variable Characteristic within each firm-year; subscripts i, j, and t denote, respectively, firm, analyst, and year; and min(·) (max(·)) is the minimum (maximum) value of the raw variable for firm i in year t. For a given firm-year, the value of each scaled variable lies between zero and one. A high value of $\text{Characteristic\_R}_{\text{ijt}}$ indicates that an analyst scores high on that variable relative to other analyst peers who follow the same firm in the same year. Applying the scaling in Eq. 1 to NFT and non-indicator control variables helps mitigate possible endogeneity problems related to firm–analyst matching on firm characteristics

[10] To further mitigate this concern, in tests of market reactions, we control for four contemporary non-EPS forecast revisions. Nevertheless, even if the effects of non-EPS forecast provision on analysts' forecasting performance are in part attributable to information contained in other recent forecasts, this does not rule out that more capable and diligent analysts provide more forecast types to signal the superior information they possess.

[11] For example, if an analyst issues one-year-ahead and two-year-ahead earnings forecasts for one of the firms she follows, her NFT_HZN for that firm is two. We cannot use simple decomposition of NFT into different components based on forecast horizons (e.g., NFT1 and NFT–NFT1) because the numbers of unique forecast types are correlated across forecast horizons. For example, consider two analysts following the same firm. Analyst A issues five types of one-year-ahead forecasts and, for each one-year-ahead forecast type, also issues a two-year-ahead forecast, but no forecast beyond FY2. NFT2 (= NFT − NFT1) for analyst A is $10 - 5 = 5$. In contrast, analyst B only issues one forecast type but issues forecasts for each of FY1–FY4 plus a long-term growth forecast (LTG). In this case the horizon proxy NFT–NFT1 for analyst B is $5 - 1 = 4$, which is lower than for analyst A, who in reality has a shorter horizon. Because of this problem, instead of decoposing NFT into forecast horizon components, we use the unique number of earnings forecast horizons in NFT as a proxy for the horizon component of NFT.

correlated with research outcomes (DeFond and Hung 2003; Givoly et al. 2019; Mauler 2019).[12] Scaling according to Eq. 1 also facilitates economic interpretation because it eliminates between-firm-year variation in NFT and other analyst characteristics. The regression coefficients on NFT_R therefore capture the expected value of the within-firm-year difference in the dependent variable between the analyst with the minimum number of forecast types (NFT_R=0) and the analyst with the maximum number of forecast types (NFT_R=1). Note, however, that regression R^2 statistics are generally lower as a consequence of scaling because between-firm-year variation in independent variables is eliminated.

3.3 NFT and earnings forecast accuracy tests

If analysts who provide more forecast types have higher ability and exert more effort in forecasting earnings, we should observe a negative association between earnings forecast error and NFT scaled within firm-years. Specifically, we expect a negative coefficient β_1 in the following regression equation:

$$\begin{aligned}100\times \mathrm{FE\_MA}_{ijt} = {} & \beta_0 + \beta_1 \mathrm{NFT\_R}_{ijt} + \beta_2 \mathrm{FE\_R}_{ijt-1} + \beta_3 \mathrm{FEXP\_R}_{ijt} + \beta_4 \mathrm{FREQ\_R}_{ijt} \\ & + \beta_5 \mathrm{GEXP\_R}_{ijt} + \beta_6 \mathrm{HOR\_R}_{ijt} + \beta_7 \mathrm{LFR\_R}_{ijt} + \beta_8 \mathrm{NFRM\_R}_{ijt} \\ & + \beta_9 \mathrm{NIND\_R}_{ijt} + \beta_{10} \mathrm{SIZE\_R}_{ijt} + \beta_{11} \mathrm{WKDN\_R}_{ijt} \\ & + \text{Fixed Effects} + \epsilon_{ijt}. \end{aligned} \tag{2}$$

The dependent variable, mean-adjusted earnings forecast error (FE_MA), is defined as the forecast error for the analyst's latest one-year-ahead EPS forecast in a given year at least 90 days prior to the earnings announcement, scaled by the average earnings forecast error for all analysts following the same firm in that year. The construction of FE_MA controls for heterogeneous forecasting difficulty across firms and years.[13] All independent variables in Eq. 2 are scaled using Eq. 1 to fall between zero and one. The variable of interest, NFT_R, is the scaled number of forecast types provided by the analyst for the firm during the three-month window ending prior to the month when the earnings forecast is issued. We control for scaled lagged earnings forecast error ($\mathrm{FE\_R}_{t-1}$) since the prior literature finds that past earnings forecast accuracy is a determinant of current earnings forecast accuracy (Brown 2001). We control for scaled analyst j's general experience (GEXP_R) and firm-specific forecasting experience (FEPX_R) because Clement (1999) and Jacob et al. (1999) document that they are positively related to analysts' earnings forecast accuracy.[14] We control for scaled earnings forecasting frequency (FREQ_R) since Jacob et al. (1999) use FREQ as a measure of analysts' effort and attentiveness to forecast earnings. We control for earnings scaled forecasting horizon (HOR_R) because earnings forecasts are more

[12] We obtain qualitatively similar main results when we use raw variables together with firm-year fixed effects.

[13] Using scaled earnings forecast error by Eq. 1 leads to qualitatively similar results.

[14] We use a sample starting from 1982 to calculate GEXP and FEXP, to mitigate potential measurement errors arising from truncation from above. The sample period starts in 1982 because EPS forecasts have been prevalent in I/B/E/S Academics (via Wharton's WRDS database) since 1982.

accurate when they are issued closer to the earnings announcement. We control for scaled leader–follower ratio (LFR_R) to capture the timeliness of an analyst's earnings forecasts, which has been found to be a robust indicator of analyst ability (Cooper et al. 2001; Shroff et al. 2014). Following Clement (1999) and Jacob et al. (1999), we control for the number of firms and the number of industries covered by the analyst to capture the analyst's portfolio complexity. We control for brokerage size (SIZE) because Stickel (1995), Clement (1999), and Jacob et al. (1999) find that an analyst's forecasting performance is positively associated with the size of her employer. We control for the analyst's walk-down bias (WKDN) because Ke and Yu (2006) empirically show that in the pre-Regulation Fair Disclosure (REGFD) period, analysts initially issue upward-biased earnings forecasts to curry favor with management and gain access to private information, then issue pessimistic forecasts that induce managers to meet and beat forecast targets. To mitigate concerns that the effects of NFT_R are solely driven by factors at the brokerage level, we control for brokerage fixed effects. To examine whether analysts strategically provide more forecast types or selectively conduct higher-quality research on firms they follow in different years, we include firm-analyst fixed effects in some model specifications. Detailed variable definitions are provided in Data Appendix B.

3.4 NFT and price target forecast accuracy tests

If analysts who provide more forecast types have higher ability or exert higher effort to forecast firms' price targets, we should observe a negative association between price target forecast error and NFT. Specifically, we expect a negative coefficient β_1 for the following regression equation:

$$
\begin{aligned}
100 \times \mathrm{PTE\_MA}_{ijt} = {} & \beta_0 + \beta_1 \mathrm{NFT\_R}_{ijt} + \beta_2 \mathrm{FE\_R}_{ijt} + \beta_3 \mathrm{BOLD\_R}_{ijt} + \beta_4 \mathrm{FEXP\_R}_{ijt} \\
& + \beta_5 \mathrm{FREQ\_R}_{ijt} + \beta_6 \mathrm{FREQ\_PT\_R}_{ijt} + \beta_7 \mathrm{GEXP\_R}_{ijt} \\
& + \beta_8 \mathrm{LFR\_R}_{ijt} + \beta_9 \mathrm{NFRM\_R}_{ijt} + \beta_{10} \mathrm{NIND\_R}_{ijt} + \beta_{11} \mathrm{SIZE\_R}_{ijt} \\
& + \beta_{12} \mathrm{WKDN\_R}_{ijt} + \text{Fixed Effects} + \epsilon_{ijt}. \qquad (3)
\end{aligned}
$$

The dependent variable, mean-adjusted price target forecast error (PTE_MA), is defined as the analyst's price target forecast error scaled by the average forecast error across all price target forecasts for the firm in that year. Price target forecast error is the absolute difference between the one-year-ahead price target and the actual stock price in 360 days, divided by the stock price on the day when the price target forecast is issued. PTE_MA eliminates the effects of variation in forecasting difficulty across firms and years. The variable of interest, NFT_R, is the number of forecast types provided by the analyst for the firm during the three-month window ending prior to the month when the price target forecast is issued. Since earnings forecasts are an important input in analysts' price target forecasting models (Da et al. 2016; Dechow and You 2020; Demirakos et al. 2010; Gleason et al. 2013), we control for concurrent earnings forecast error (FE_R) to examine whether NFT_R captures analysts' ability and effort to forecast price targets beyond their ability and effort to forecast earnings.[15] We control for analysts' price target forecasting frequency (FREQ_PT) to capture analysts' effort and attentiveness to forecast price targets.[16] Other variables are defined in Section 3.3 and Data Appendix B.[17] As before, we include brokerage house or firm-analyst fixed effects for some model specifications.

3.5 NFT and stock recommendation profitability tests

If analysts who provide more forecast types have superior ability or exert greater research effort in evaluating the firms they follow, we should observe a positive association between the profitability of stock recommendations and NFT. Specifically, we expect a positive coefficient β_1 in the following regression equation:

$$\begin{aligned} 100 \times \mathrm{RET\_REC}_{ijt} = {} & \beta_0 + \beta_1 \mathrm{NFT\_R}_{ijt} + \beta_2 \mathrm{FE\_R}_{ijt} + \beta_3 \mathrm{BOLD\_R}_{ijt} + \beta_4 \mathrm{FEXP\_R}_{ijt} \\ & + \beta_5 \mathrm{FREQ\_R}_{ijt} + \beta_6 \mathrm{FREQ\_REC\_R}_{ijt} + \beta_7 \mathrm{GEXP\_R}_{ijt} \\ & + \beta_8 \mathrm{LFR\_R}_{ijt} + \beta_9 \mathrm{NFRM\_R}_{ijt} + \beta_{10} \mathrm{NIND\_R}_{ijt} + \beta_{11} \mathrm{SIZE\_R}_{ijt} \\ & + \beta_{12} \mathrm{WKDN\_R}_{ijt} + \text{Fixed Effects} + \epsilon_{ijt}. \end{aligned} \tag{4}$$

Compared to the portfolio approach presented later, multivariate regressions allow us to control for a broad set of analyst characteristics (Ertimur et al. 2007). We define the dependent variable, RET_REC, to be long in stocks with a strong buy or buy recommendation (coded as 1 and 2 by I/B/E/S) and short in stocks with a Hold, Sell, or strong sell recommendation (coded as 3, 4, and 5 by I/B/E/S). The return accumulation period runs from the day before the recommendation until the earlier of 180 days or the day before the recommendation is revised or reiterated. We are short in stocks with a hold recommendation (coded as 3 by I/B/E/S) to correct for the well-documented optimism bias in analysts' stock recommendations.[18] The returns accumulation window is chosen because prior research (e.g., Womack 1996) shows that although the initial market reaction to recommendations is high, there is also significant post-recommendation drift over a period of up to 180 days. Hence, returns over this window capture the potential total investment value of recommendations to investors with early (i.e., pre–I/B/E/S publication) access to brokerage house recommendations.[19] We control for the concurrent earnings forecast error (FE_R), as Loh

[15] Controlling for lagged earnings forecast error delivers qualitatively similar results but reduces the sample size aggressively.

[16] We still control for analysts' earnings forecasting frequency (FREQ) because earnings forecasts are an important input in generating the ultimate products of an analyst's research, like price targets and stock recommendations. Earnings forecasting frequency reflects an analyst's overall forecasting effort and attentiveness.

[17] Although Hashim and Strong (2018) document that price target forecasts accompanied by cash flow forecasts are more accurate, cash flow forecast provision is not significantly correlated with price target forecast accuracy in our sample under our research design. For simplicity, we do not include analysts' cash flow forecast provision in the regression models. Our main results are insensitive to the inclusion of analysts' cash flow forecast provision.

[18] The main results are qualitatively similar when we exclude hold recommendations from the analysis.

[19] We obtain qualitatively similar results when using a return accumulation period of up to 30 days.

and Mian (2006) and Ertimur et al. (2007) document that earnings forecast accuracy is positively related to analysts' stock recommendation profitability. We control for analysts' stock recommendation frequency (FREQ_REC_R) to capture analysts' effort and attentiveness in recommending stocks. Other control variables are the same as those used in Eqs. 2 and 3.

3.6 NFT and recommendation revision market reaction tests

If the number of forecast types captures incremental information about research ability and effort and the market places more weight on recommendations associated with higher-quality research, we should observe higher market reactions to recommendation revisions when NFT_R is higher. We test this prediction using the following equation:

$$\begin{aligned}100 \times \mathrm{CAR3}_{\mathrm{ijt}} = {} & \beta_0 + \beta_1 \mathrm{RECrev}_{\mathrm{ijt}} + \beta_2 \mathrm{NFT\_R}_{\mathrm{ijt}} + \beta_3 \mathrm{RECrev}_{\mathrm{ijt}} \times \mathrm{NFT\_R}_{\mathrm{ijt}} \\ & + \gamma \, \mathit{OtherForecastRevisions} + \delta(\mathrm{RECrev} \times \mathrm{I\_}\mathit{OtherForecastRevisions}) \\ & + \theta \mathit{AnalystCharacteristics\_R} + \eta(\mathrm{RECrev} \times \mathit{AnalystCharacteristics\_R}) \\ & + \text{Firm-Year Fixed Effects} + \epsilon_{\mathrm{ijt}}. \end{aligned} \tag{5}$$

We predict a negative coefficient β_3 in Eq. 5, given that strong sell (sell) recommendations are coded 5 (4) and strong buy (buy) recommendations are coded 1 (2), and the stock recommendation revision RECrev is equal to the change in recommendation. *OtherForecastRevision* = {EPSrev, EBIrev, PRErev, SALrev} are, respectively, revisions in an analyst's forecasts of earnings per share, earnings before interest and taxes, pre-tax earnings, and sales, or zero if there is no revision. We scale EPS forecast revisions by the stock price two days prior to the forecast revision. We scale all other forecast revisions by the absolute value of the previously outstanding corresponding forecast. The variables I_*Other Forecast Revision* equal one if the analyst revises the related forecast and the stock recommendation on the same day, zero otherwise. As before, *AnalystCharacteristics* = {FE_R$_{t-1}$, BOLD_R, FEXP_R, FREQ_R, FREQ_REC_R, GEXP_R, LFR_R, NFRM_R, NIND_R, SIZE_R, WKDN_R}.

We control for contemporaneous earnings forecast revisions by the same analyst because Stickel (1995) finds stronger market reactions to recommendation revisions when analysts issue a confirming earnings forecast revision on the same day. We control for contemporaneous sales forecast revisions (SALrev) since Keung (2010) documents that investors react more strongly to earnings forecast revisions accompanied by sales forecasts when the additional information contained in sales forecast revisions is controlled for. Although Keung (2010) studies the market reaction to earnings forecast revisions, we empirically show that his finding extends to stock recommendation revisions.[20] For the same reason, we control for contemporaneous other forecast revisions and the interactions between the indicators of those forecast revisions and stock recommendation revisions.

[20] Unreported results show that investors react more strongly to earnings forecast revisions when NFT is higher.

3.7 NFT and analyst career outcome tests

Assuming that successful career outcomes are what analysts ultimately pursue, analysts who have a higher ability or exert higher effort should end up with better career outcomes (CAREER OUTCOME), defined as TERMINATION, PROMOTION or DEMOTION. Specifically, we expect that analysts who provide more forecast types are less likely to be terminated (TERMINATION) or demoted (DEMOTION) by their employers and are more likely to be promoted (PROMOTION) from smaller brokerage houses to larger ones.[21] We expect a negative (positive) coefficient β_1 when TERMINATION or DEMOTION (PROMOTION) is the dependent variable for the following conditional logistic model matched at the brokerage house level[22]:

$$\begin{aligned}\text{CAREER OUTCOME}_{jt+1} = {} & \beta_1\text{NFT\_M}_{\text{jt}} + \beta_2\text{I\_AFLTG}_{\text{jt}} + \beta_3\text{I\_AFSAL}_{\text{jt}} + \beta_4\text{FE\_M}_{\text{jt}} \\ & + \beta_5\text{BOLD\_M}_{\text{jt}} + \beta_6\text{COMP\_M}_{\text{jt}} + \beta_7\text{FREQ\_M}_{\text{jt}} \\ & + \beta_8\text{GEXP\_M}_{\text{jt}} + \beta_9\text{LFR\_M}_{\text{jt}} + \beta_{10}\text{NFRM\_M}_{\text{jt}} \\ & + \beta_{11}\text{NIND\_M}_{\text{jt}} + \beta_{12}\text{SIZE\_M}_{\text{jt}} + \beta_{13}\text{WKDN\_M}_{\text{jt}} \\ & + \text{Brokerage Fixed Effects} + \epsilon_{\text{jt+1}}. \end{aligned} \tag{6}$$

The variable of interest, NFT_M, is the average rank of the number of forecast types provided by the analyst across all firms followed by the analyst in the year. Specifically, we calculate the number of forecast types provided by the analyst for any firm she follows, scale the variable within each firm she covers, and take the average of the scaled variable across all the firms she follows in the year. Other non-indicator regressors are defined similarly and denoted by the suffix _M. We control for the analyst's provision of long-term earnings growth forecasts (I_AFLTG), since Jung et al. (2012) document that analysts who provide long-term earnings growth forecasts are less likely to be fired or demoted. We further control for the analyst's provision of sales forecasts (I_AFSAL), since Ertimur et al. (2011) find that analysts who provide sales forecasts are more (less) likely to be promoted (demoted or terminated). We control for earnings forecast accuracy (FE_M) because prior studies find that the likelihood of analyst turnover is higher when earnings forecasts are relatively less accurate (Mikhail et al. 1999). We control for boldness (BOLD_M) because Hong et al. (2000) document that inexperienced analysts are more likely to be terminated for bold forecasts. We also control for the peer competition facing analysts (COMP_M) and for walk-down bias (WKDN_M).

4 Sample and descriptive statistics

4.1 Sample selection

The initial sample comprises all one-year-ahead to five-year-ahead forecasts and long-term-growth forecasts in the I/B/E/S US detail file from the period 2001–2020. The

[21] Note that in the test of PROMOTION (DEMOTION), the test samples are conditioned on non-termination in year t and the analyst's having worked at a small (large) brokerage house in year t.

[22] Conditional logistic regression models are also known as fixed-effects logit models for panel data. See McFadden (1973) for details.

sample period for our regression analysis is 2002–2019 due to inclusion of leading and lagged variables. The sample begins from 2002 because most I/B/E/S forecast types have been available since that year.[23] To rule out the possibility that analysts signal higher ability and effort by providing more forecast types only in some sample years, in unreported tests we split the sample into pre- (including 2010) and post-2010 subsamples and re-estimate Eqs. 2–6 on the two subsamples. All our main results hold in both the pre- and post-2010 periods.

Forecasts other than long-term-growth forecasts with a horizon over five years are excluded from our sample because they are sparse. To avoid noise caused by possible currency conversion errors, we only consider US dollar-denominated forecasts and reported realizations. We require a firm-year to have at least two analysts following it to be included. We require an analyst to provide at least two one-year-ahead EPS forecasts for each firm she follows in a given year. The analyst's first forecast included for a firm-year must be issued after the previous fiscal year-end and no later than 90 days before the annual earnings announcement date for year t, and the final forecast for a firm-year must be issued before the annual earnings announcement date. These requirements ensure that walk-down bias (WKDN) is well-defined for each analyst-firm-year combination. The units of analysis in our various tests are at different levels of granularity (e.g., analyst-firm-year, analyst-year, analyst-price target, analyst-recommendation, analyst-recommendation revision) with dependent variables available for different data subsets. Since the observation units differ across tests, we provide descriptive statistics for the sample used in each set of tests in the corresponding sections that follow.

4.2 Descriptive statistics for NFT and its components

Table 1 first shows the numbers of all analysts, brokerage houses, and firms covered in the I/B/E/S Detail database for the United States over the period 2002–2019. While there is some growth in the number of analysts, brokerage houses and firms covered, at least until 2016, the average number of analysts per firm is quite stable over time at around 16.50. However, the NFT column in Table 1 reports the growth over the sample period in the average (unscaled) number of (unique) forecast types (NFT) and the breadth component of NFT based on one-year-ahead forecast types, NFT1. It shows that the average number of forecast types at the firm-year level increased from 7.6 in 2002 to 25.9 in 2019. One-year-ahead forecasts have increased from an average of 3.2 in 2002 to 8.6 in 2019.[24] The ratio between average NFT and average NFT1 forecasts has increased from 2.4 (= 7.59/3.20) in 2002 to 3.0 (= 25.91/8.64) in 2019; and, consistent with this, the average value of NFT_HZN has also increased over the sample period, from 2.3 in 2002 to over 2.9 in 2019. If the horizon in EPS forecast provision is correlated with the horizon of other non-EPS forecast types, this

[23] Using a post-2002 sample also mitigates the concern that the introduction of Reg FD might distort the results.

[24] A more granular analysis of the trends in non-EPS line-item forecast frequencies categorized by financial statement, reported in the Online Appendix, shows that by 2019 76% of forecast types are for income statement items, consistent with revenue forecasts and non-GAAP earnings numbers increasing in importance.

Table 1 Analyst coverage and forecast types in the I/B/E/S detail file

Year	# analysts	# brokerage houses	# firms	Avg. # analysts per firm	Avg. # forecast types per firm	NFT	NFT1	NFT_HZN
2002	4,748	260	4,838	16.08	16.19	7.59	3.20	2.33
2003	4,613	363	4,705	16.44	23.35	10.14	4.07	2.51
2004	4,154	386	4,955	15.60	27.51	12.47	4.84	2.67
2005	4,142	383	5,127	15.10	29.61	13.41	5.10	2.71
2006	4,128	350	5,256	14.72	35.03	15.62	5.87	2.67
2007	4,215	333	5,318	14.63	37.11	16.52	6.15	2.69
2008	4,032	323	4,945	14.42	38.72	17.68	6.58	2.71
2009	3,876	357	4,764	15.24	40.04	19.36	7.02	2.79
2010	4,184	390	4,832	16.54	46.40	22.57	8.14	2.83
2011	4,371	347	4,817	17.50	49.27	23.81	8.49	2.80
2012	4,298	342	4,729	17.26	50.16	24.35	8.67	2.78
2013	5,096	450	5,059	17.42	47.19	22.98	7.91	2.74
2014	6,220	521	5,335	17.33	51.50	25.32	8.50	2.86
2015	6,212	525	5,348	17.31	52.59	26.33	8.84	2.91
2016	6,056	498	5,079	17.51	52.89	26.51	8.91	2.90
2017	5,656	465	4,911	17.16	52.75	26.62	8.88	2.94
2018	5,220	443	4,910	16.60	52.49	26.55	8.85	2.96
2019	4,878	417	4,898	16.05	51.80	25.91	8.64	2.92

This table is based on the I/B/E/S US detail file and summarizes trends over time in the number of analysts, number of brokerage houses, number of firms covered, average analyst coverage per firm, average number of unique forecast types per firm, average number of forecast types per firm-analyst pair (NFT), average number of one-year-ahead forecast types per firm-analyst pair (NFT1), and average number of analyst earnings forecast horizons (FY1–FY5 and long-tern earnings growth) per firm-analyst pair (NFT_HZN). None of the filters used to construct the samples studied in this paper are applied in constructing this table suggests that the trend in NFT reflects a trend in horizon and that NFT_HZN is a good proxy for the overall horizon component of NFT.

We cannot determine the extent to which the observed growth in forecast type provision reflects changes in analyst forecasting behavior, in analysts' forecast reporting practices, or in I/B/E/S data collection methodologies. However, the trends we find in the data emphasize the importance of controlling for differences, across firm-years, in patterns of forecast provision, using the firm-year scaling process from Eq. 1 (as described earlier).

In Fig. 1, Panel A shows graphically the trends in average NFT, NFT1, and NFT_HZN, and Panel B shows the trends in the average within-firm dispersions of the same variables (consistent with Eq. 1 measured by the range, i.e., maximum–minimum). The average within-firm range for NFT grew from 18.6 in 2002 to 43.59 in 2019. There is a more modest but still material growth in dispersion of NFT1, our breadth measure. The within-firm-year dispersion we observe for NFT has empirical advantages, compared to alternative proxies for research quality or effort based on single forecast types (e.g., revenue or cash flow forecasts). First, the granularity of forecast type underlying NFT allows for more precise ranking of analysts' research quality. Second, NFT applies to a larger set of firms than any single forecast type. For example, in 2020, only 56% of firms in the I/B/E/S universe had cash flow forecast coverage. The research quality of the analysts following the remaining 44% of the firms is ambiguous. Also, within-firm-year variation in the provision of some forecast types has been decreasing over time. For example, in 2020, within the firms with sales revenue forecast coverage, 92.2% of the following analysts provided at least one sales forecast. Ertimur et al. (2011) point out that the decreasing cross-sectional variation in forecasts of earnings components "may weaken the potency of disaggregation as a means for establishing reputation."

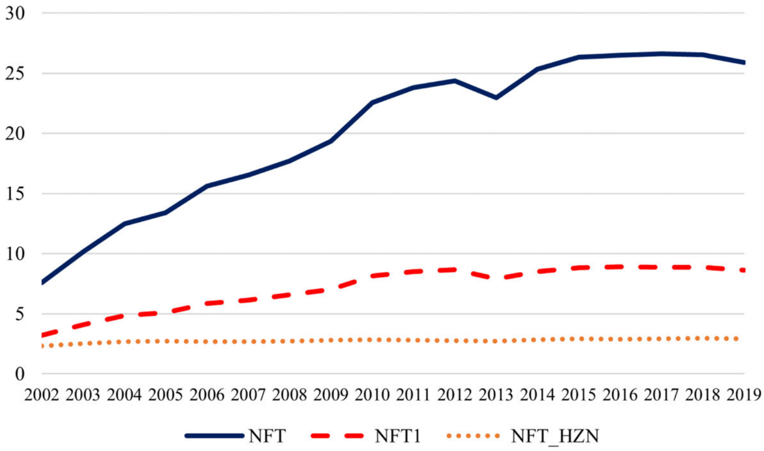
5 What does NFT capture?

5.1 Association between NFT and other analyst characteristics

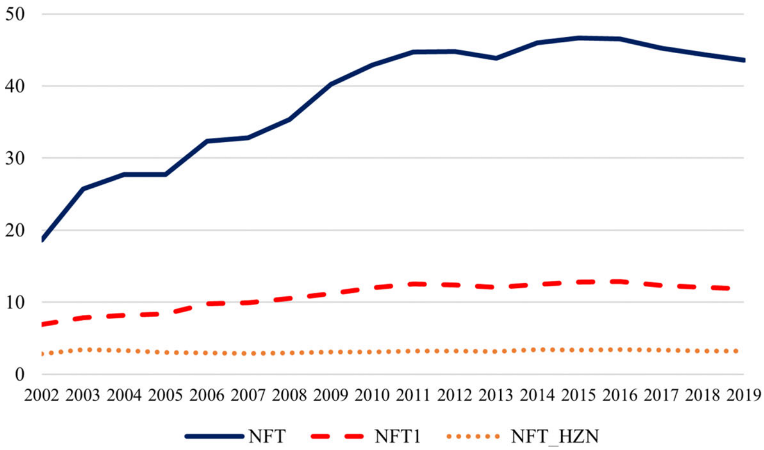
Before testing the relation between NFT and research outcomes, we first investigate how NFT relates to the characteristics of analysts' forecasting behavior and analyst quality studied in the prior literature. Specifically, we estimate the following regression equation:

$$\begin{aligned}100 \times \mathrm{NFT\_R_{ijt}} = {} & \beta_0 + \beta_1 \mathrm{ABLT\_R_{ijt}} + \beta_2 \mathrm{BOLD\_R_{ijt}} + \beta_3 \mathrm{CONS\_R_{ijt}} + \beta_4 \mathrm{FEXP\_R_{ijt}} \\ & + \beta_5 \mathrm{FREQ\_R_{ijt}} + \beta_6 \mathrm{GEXP\_R_{ijt}} + \beta_7 \mathrm{LFR\_R_{ijt}} + \beta_8 \mathrm{NFRM\_R_{ijt}} \\ & + \beta_9 \mathrm{NIND\_R_{ijt}} + \beta_{10} \mathrm{SIZE\_R_{ijt}} + \beta_{11} \mathrm{STAR_{ijt}} + \beta_{12} \mathrm{WKDN\_R_{ijt}} + \epsilon_{ijt}.\end{aligned} \tag{7}$$

The dependent variable, NFT_R, is the number of forecast types provided by the analyst for a given firm-year. All variables except all-star status (STAR), which is an indicator variable, are scaled using Eq. 1 to fall between zero and one. As a proxy for innate analyst ability, we define analysts' ability to move the consensus earnings



(a) Avg. within-firm year NFT, NTF1, and NFT_HZN



(b) Avg. within-firm year range of NFT, NTF1, and NFT_HZN

Fig. 1 Panel (a) shows average within-firm-year analysts' number of forecast types (NFT), number of one-year-ahead forecast types (NFT1), and number of earnings forecast horizons (NFT_HZN). Analyst earnings forecast horizons include FY1–FY5 and long-term earnings growth. Panel (b) shows average within-firm-year range of analysts' number of forecast types (NFT), number of one-year-ahead forecast types (NFT1) and number of earnings forecast horizons (NFT_HZN)

forecasts towards the actual earnings (ABLT_R) following Chen and Jiang (2006) and Ertimur et al. (2011). We include forecast consistency (CONS), since Hilary and Hsu (2013) show that the stock market reacts more strongly to earnings forecast revisions of analysts who forecast earnings more consistently. The association between NFT_R and all-star status (STAR) is potentially important because the prior literature documents a positive association between analysts' external reputation and their performance (Stickel 1992, 1995). All other variables are defined in previous sections, and detailed definitions are provided in the Data Appendix B. Unreported correlations indicate no evidence of significant multicollinearity. For Eq. 7, we cluster standard errors at the firm and the analyst level.

Table 2 Panel A reports descriptive statistics for raw values of NFT and analyst characteristics before scaling. The sample contains up to 310,546 firm-analyst-year observations for the period 2002–2015, based on the availability of all-star status data.[25] The distributions of analyst characteristics are broadly in line with prior research. Panel B of Table 2 presents the descriptive statistics for the scaled variables.[26] The descriptive statistics indicate that there is reasonably uniform variation in the distributions of NFT_R and other scaled analyst characteristics, providing reassurance than any significant associations are not driven by extreme values of analyst characteristics.

Table 2 Panel C reports the regression results for Eq. 7 in Column (1). The results show that NFT_R has a strong within-firm-year statistical association with most of the analyst characteristics studied in the prior analyst literature. Analysts who provide more forecast types have greater ability to move the consensus earnings forecasts (ABLT_R), higher forecast boldness (BOLD_R), higher earnings forecasting frequency (FREQ_R), and greater forecasting timelines (LFR_R); follow more firms (NFRM_R); and work in larger brokerage firms (SIZE_R). However, they also display lower forecast consistency (CONS_R), have followed firms for fewer years (FEXP_R), have less general experience (GEXP_R), and follow fewer industries. Note, however, that the adjusted R^2 in Column 1 is only 2%, suggesting that these analyst characteristics explain a only small proportion of the within-firm-year variation in NFT_R.

5.2 Regressing NFT on fixed effects

Columns (2)–(4) of Table 2 Panel C report the adjusted R^2s for the regressions of NFT on brokerage fixed effects, analyst fixed effects, and firm-analyst fixed effects, respectively. In addition to providing analysts with better resources and distribution networks, brokers may also affect analysts' forecast provision through internal policies—they may provide analysts with template forecasting models or guidance on forecasting methodologies and disclosure formats. Column (2) indicates that over 20% of the variation in NFT_R is explained by brokerage house fixed effects. Column (3) shows that nearly 38% of the variation in NFT_R is explained by analyst fixed effects, capturing analysts' average innate ability and average level of research effort across all of the firms they follow. Column (4) indicates that firm-analyst fixed effects explain over 50% of the variation in NFT_R. We expect firm-analyst fixed effects to capture any firm-specific innate ability of analysts (as suggested by Clement et al. (2007)), as well as firm-year-specific differences in research effort expended by individual analysts.

[25] Our sample period for the analysis in Section 5 is 2002–2015. The sample period ends in 2015 because the All-American Research Team nominations from *Institutional Investor* magazine are not available to us after 2014. In our main tests, we exclude all-star status (STAR) and extend the sample period to 2019. We believe the benefit of including more observations from recent years outweighs the cost of dropping STAR from the regressions, because unreported results show that the estimated coefficients for STAR are insignificant in all regressions.

[26] We require an observation to have non-missing analyst characteristics and well-defined NFT to be included in the sample. This requirement applies to all samples studied in this paper.

Table 2 Descriptive statistics and regression results from tests of NFT and analyst characteristics or fixed effects

Panel A: Descriptive statistics of raw NFT and analyst characteristics

	N	**Mean**	**Sd**	**Q1**	**Median**	**Q3**
NFT	310,546	21.145	13.102	12.000	19.000	28.000
FEXP	310,546	4.813	3.642	2.000	4.000	6.000
FREQ	310,546	5.101	2.600	3.000	5.000	6.000
GEXP	310,546	10.368	6.334	5.000	9.000	14.000
NFRM	310,546	17.297	7.666	13.000	16.000	21.000
NIND	310,546	4.084	2.443	2.000	4.000	5.000
SIZE	310,546	67.593	61.675	22.000	48.000	104.000
STAR	310,546	0.102	0.302	0.000	0.000	0.000

Panel B: Descriptive statistics of scaled NFT and analyst characteristics

	N	**Mean**	**Sd**	**Q1**	**Median**	**Q3**
NFT_R	310,546	0.436	0.337	0.143	0.389	0.692
ABLT_R	310,546	0.518	0.329	0.260	0.532	0.786
BOLD_R	310,546	0.372	0.344	0.071	0.268	0.606
CONS_R	310,546	0.588	0.352	0.293	0.666	0.918
FEXP_R	310,546	0.428	0.358	0.100	0.364	0.714
FREQ_R	310,546	0.452	0.343	0.167	0.429	0.700
GEXP_R	310,546	0.422	0.340	0.125	0.357	0.684
LFR_R	310,546	0.311	0.350	0.028	0.156	0.497
NFRM_R	310,546	0.455	0.332	0.176	0.417	0.714
NIND_R	310,546	0.418	0.351	0.000	0.353	0.667
SIZE_R	310,546	0.349	0.337	0.065	0.238	0.511
WKDN_R	310,546	0.498	0.343	0.200	0.500	0.800

Table 2 continued

Panel C: Results of regression of NFT_R on analyst characteristics or fixed effects				
	(1)	**(2)**	**(3)**	**(4)**
ABLT_R	1.839***			
	(4.090)			
BOLD_R	0.869***			
	(3.062)			
CONS_R	−2.834***			
	(−6.095)			
FEXP_R	−3.699***			
	(−6.969)			
FREQ_R	6.138***			
	(11.770)			
GEXP_R	−1.718*			
	(−1.659)			
LFR_R	2.118***			
	(4.398)			
NFRM_R	6.317***			
	(6.503)			
NIND_R	−2.396***			
	(−2.956)			
SIZE_R	8.344***			
	(7.477)			
STAR	−0.594			
	(−0.512)			
WKDN_R	0.007			
	(0.026)			
Observations	310,546	310,520	310,335	302,679
Adjusted R^2	0.020	0.202	0.377	0.501
Fixed Effects	No	Broker	Analyst	Firm-Analyst

This table reports descriptive statistics of raw (Panel A) and scaled (Panel B) NFT and analysts' characteristics and results from regressing scaled NFT (NFT_R) on analyst characteristics or fixed effects (Panel C). NFT is the number of forecast types provided by analyst j for firm i during year t. In Panel C, Column (1) reports regression results for the following regression equation:

$$\begin{aligned} 100 \times \mathrm{NFT\_R}_{ijt} = {} & \beta_0 + \beta_1 \mathrm{ABLT\_R}_{ijt} + \beta_2 \mathrm{BOLD\_R}_{ijt} + \beta_3 \mathrm{CONS\_R}_{ijt} + \beta_4 \mathrm{FEXP\_R}_{ijt} \\ & + \beta_5 \mathrm{FREQ\_R}_{ijt} + \beta_6 \mathrm{GEXP\_R}_{ijt} + \beta_7 \mathrm{LFR\_R}_{ijt} + \beta_8 \mathrm{NFRM\_R}_{ijt} \\ & + \beta_9 \mathrm{NIND\_R}_{ijt} + \beta_{10} \mathrm{SIZE\_R}_{ijt} + \beta_{11} \mathrm{STAR}_{ijt} + \beta_{12} \mathrm{WKDN\_R}_{ijt} + \epsilon_{ijt}. \end{aligned}$$

Columns (2)–(4) report the adjusted R^2 statistics for the following regression equation:

$$\mathrm{NFT\_R}_{ijt} = \beta_0 + \text{Fixed Effects} + \epsilon_{ijt},$$

where FixedEffects are brokerage house fixed effects, analyst fixed effects, or firm-analyst fixed effects. An intercept is estimated but unreported for Column (1). The sample period is 2002–2015. Standard errors are clustered at the firm and analyst level. *, **, and *** denote significance at the 0.10, 0.05, and 0.01 levels, respectively. Detailed variable definitions are provided in Data Appendix B

This evidence suggests that NFT may be a credible surrogate for analyst innate ability and effort.

6 Main results

6.1 NFT and earnings forecast accuracy

Panel A of Table 3 reports the descriptive statistics for mean-adjusted earnings forecast error (FE_MA) and NFT_R.[27] The distribution of NFT_R confirms that within-firm-year variation is quite uniform, ensuring that our regression results are not driven by extreme forecast type provision by a small number of analysts. Unreported correlations show no significant multicollinearity.[28]

Table 3 Panel B reports the regression results for Eq. 2. The coefficients on the control variables confirm that analyst characteristics that were identified as relevant by the prior literature remain significant determinants of earnings forecast accuracy. However, NFT_R contributes incrementally to the explanatory power in all regressions and is associated with higher forecast accuracy. In Column (1), the coefficient on NFT_R indicates that the expected difference in earnings forecast accuracy when comparing the maximum NFT analyst and the minimum NFT analyst within each firm-year is 4.90%. An additional measure of economic significance is that, for a one standard deviation change in NFT_R reported in Panel A (0.34), the expected marginal change in earnings forecast accuracy is 1.67% ($= 0.34 \times 4.90\%$). When we introduce more conservative fixed effects designs, including brokerage house fixed effects (in Column (2)) and firm-analyst fixed effects (in Column (3)), the economic significance of NFT_R falls slightly, but the results remain robustly statistically significant and are consistent with NFT capturing variation in earnings forecast accuracy that cannot be explained by brokerage house styles and/or analysts' innate abilities. In Section 7.1 we test the value of NFT to investors in out-of-sample earnings forecast revision portfolios.

6.2 NFT and price target forecast accuracy

Table 4 Panel B reports the regression results for Eq. 3. The coefficients on the control variables confirm that several analyst characteristics that were identified as relevant by the prior literature remain significant determinants of price target forecast accuracy. However, NFT_R contributes incrementally to the explanatory power in all regressions and is again associated with higher forecast accuracy. In Column (1), the coefficient on NFT_R indicates that the expected difference in price target forecast accuracy when comparing the maximum NFT analyst and the minimum NFT analyst within

[27] In this and subsequent tables we report only descriptive statistics for the main variables of interest. We omit the descriptive statistics for other independent variables, as their distributions are largely similar to the variables in Panel B of Table 2.

[28] Similar properties of NFT_R and control variables apply in other tests reported below, and therefore we do not discuss them further.

Table 3 Descriptive statistics and regression results from tests of NFT and earnings forecast accuracy

Panel A: Descriptive statistics

	N	**Mean**	**Sd**	**Q1**	**Median**	**Q3**
FE_MA	390,204	0.986	0.799	0.470	0.857	1.234
NFT_R	390,204	0.431	0.341	0.118	0.382	0.692

Panel B: Regression results for tests of NFT and earnings forecast accuracy

	(1)	**(2)**	**(3)**
NFT_R	−4.900***	−4.368***	−2.412***
	(−8.207)	(−8.845)	(−3.933)
FE_R_{t-1}	17.430***	15.920***	−16.740***
	(23.300)	(22.260)	(−31.130)
FEXP_R	−4.100***	−2.593***	8.456***
	(-8.628)	(-5.869)	(11.510)
FREQ_R	−12.430***	−11.160***	−9.851***
	(−22.100)	(−21.830)	(−16.920)
GEXP_R	−4.546***	−1.505**	−0.003
	(−6.093)	(−2.145)	(−0.002)
HOR_R	55.030***	54.130***	50.700***
	(67.480)	(66.020)	(58.080)
LFR_R	−8.203***	−6.922***	−4.075***
	(−19.040)	(−16.510)	(−9.177)
NFRM_R	−5.256***	−4.751***	−6.542***
	(−7.470)	(−7.485)	(−7.891)
NIND_R	−1.139*	−1.134*	0.356
	(−1.777)	(-1.903)	(0.496)
SIZE_R	1.353*	−1.142*	1.793*
	(1.842)	(−1.827)	(1.769)
WKDN_R	5.647***	5.405***	5.303***
	(8.857)	(8.601)	(8.247)
Observations	390,204	390,140	350,386
Adjusted R^2	0.097	0.108	0.181
Fixed Effects	No	Broker	Firm-Analyst

This table reports descriptive statistics (Panel A) and regression results for tests of NFT and analyst earnings forecast accuracy (Panel B). Specifically, Panel B reports the regression results of the following regression equation:

$$\begin{aligned}100 \times FE\_MA_{ijt} = \beta_0 &+ \beta_1 NFT\_R_{ijt} + \beta_2 FE\_R_{ijt-1} + \beta_3 FEXP\_R_{ijt} + \beta_4 FREQ\_R_{ijt} \\ &+ \beta_5 GEXP\_R_{ijt} + \beta_6 HOR\_R_{ijt} + \beta_7 LFR\_R_{ijt} + \beta_8 NFRM\_R_{ijt} \\ &+ \beta_9 NIND\_R_{ijt} + \beta_{10} SIZE\_R_{ijt} + \beta_{11} WKDN\_R_{ijt} + \text{Fixed Effects} + \epsilon_{ijt},\end{aligned}$$

where FE_MA is the mean-adjusted forecast error for the latest one-year-ahead earnings per share (EPS) forecast no later than 90 days prior to the earnings announcement provided by analyst j for firm i in year t. FE_MA is winsorized at the 99th percentile for each cross section. NFT is the number of forecast types provided by analyst j for firm i during the three months prior to the month in which the dependent variable is measured. All non-indicator independent variables are scaled using Eq. 1 and denoted by the suffix "_R." An intercept is estimated but unreported in each column. The sample period is 2002–2019. Standard errors are clustered at the firm and analyst level. *, **, and *** denote significance at the 0.10, 0.05, and 0.01 levels, respectively. Detailed variable definitions are provided in Data Appendix B

Table 4 Descriptive statistics and regression results for tests of NFT and price target forecast accuracy

Panel A: Descriptive statistics

	N	Mean	Sd	Q1	Median	Q3
PTE	841,572	0.322	0.293	0.109	0.241	0.447
PTE_MA	841,572	0.995	0.617	0.525	0.943	1.363
NFT_R	841,572	0.427	0.292	0.200	0.400	0.625
FREQ_PT_R	841,572	0.565	0.338	0.300	0.500	1.000

Panel B: Regression results from tests of NFT and price target forecast accuracy

	(1)	(2)	(3)
NFT_R	−1.901***	−1.225***	−2.315***
	(−4.476)	(−3.331)	(−5.539)
FE_R	1.768***	1.626***	1.159***
	(7.587)	(7.174)	(4.319)
BOLD_R	2.568***	2.362***	1.143***
	(10.000)	(9.259)	(3.998)
FEXP_R	1.861***	1.843***	3.559***
	(6.206)	(6.289)	(7.353)
FREQ_R	0.412	0.229	0.015
	(1.360)	(0.804)	(0.0483)
FREQ_PT_R	−0.801**	−0.752**	0.523
	(−2.271)	(−2.327)	(1.455)
GEXP_R	0.209	0.550	0.764
	(0.519)	(1.456)	(1.177)
LFR_R	−0.414	−0.354	−0.340
	(−1.581)	(−1.390)	(−1.211)
NFRM_R	0.292	0.532	0.086
	(0.713)	(1.341)	(0.168)
NIND_R	0.848**	0.808**	0.225
	(2.252)	(2.290)	(0.511)
SIZE_R	−2.376***	−2.043***	−0.139
	(−6.024)	(−5.445)	(−0.249)
WKDN_R	0.762***	0.649***	0.241
	(2.998)	(2.628)	(0.914)
Observations	841,572	841,525	830,119
Adjusted R^2	0.001	0.003	0.037
Fixed Effects	No	Broker	Firm-Analyst

This table reports descriptive statistics (Panel A) and regression results from tests of NFT and analyst price target forecast accuracy (Panel B). Specifically, Panel B reports the regression results of the following regression equation:

$$\begin{aligned}100 \times \mathrm{PTE\_MA}_{ijt} = {} & \beta_0 + \beta_1 \mathrm{NFT\_R}_{ijt} + \beta_2 \mathrm{FE\_}R_{ijt} + \beta_3 \mathrm{BOLD\_R}_{ijt} \\ & + \beta_4 \mathrm{FEXP\_R}_{ijt} + \beta_5 \mathrm{FREQ\_R}_{ijt} + \beta_6 \mathrm{FREQ\_PT\_R}_{ijt} \\ & + \beta_7 \mathrm{GEXP\_R}_{ijt} + \beta_8 \mathrm{LFR\_R}_{ijt} + \beta_9 \mathrm{NFRM\_R}_{ijt} \\ & + \beta_{10} \mathrm{NIND\_R}_{ijt} + \beta_{11} \mathrm{SIZE\_R}_{ijt} + \beta_{12} \mathrm{WKDN\_R}_{ijt} + \text{Fixed Effects} + \epsilon_{ijt},\end{aligned}$$

where PTE_MA is mean-adjusted price target error, defined as analyst j's absolute price target error scaled by the mean of all absolute price target error for firm i in year t. Absolute price target error is defined as the absolute value of the difference between the one-year-ahead price target and the stock price in 360 days, divided by the current stock price. PTE_MA is winsorized at the 99th percentile for each cross section. NFT is the number of forecast types provided by analyst j for firm i during the three months prior to the month in which the dependent variable is measured. All non-indicator independent variables are scaled using Eq. 1. An intercept is estimated but not reported in each column. The sample period is 2002–2019. Standard errors are clustered at the firm and analyst level. *, **, and *** denote significance at 0.10, 0.05, and 0.01 levels, respectively. Detailed variable definitions are provided in Data Appendix B each firm-year is 1.90%. An additional measure of economic significance is that, for a one standard deviation change in NFT_R reported in Panel A (0.29), the expected marginal change in price target forecast accuracy is 0.56% (= $0.29 \times 1.90\%$). When we introduce fixed effects, the results for NFT_R remain robustly statistically significant. Including brokerage house fixed effects (in Column (2)) reduces the economic significance for price target forecast accuracy slightly, whereas inclusion of firm-analyst fixed effects (in Column (3)) even increases the economic significance, consistent with NFT capturing the quality of analysts' research after controlling for brokerage house styles and the innate ability of analysts.

6.3 NFT and the profitability of stock recommendations

Table 5 Panel B reports the regression results for Eq. 4. The coefficients on the control variables confirm that several analyst characteristics that were identified as relevant by the prior literature remain significant determinants of recommendation profitability (defined as market-adjusted returns). However, NFT_R again contributes incrementally to the explanatory power in all regressions, and it is associated with higher recommendation profitability. In Column (1), the coefficient on NFT_R indicates that the expected difference in recommendation profitability when comparing the maximum NFT analyst and the minimum NFT analyst within each firm-year is 0.99% over a holding period of up to 180 days, equivalent to annualized returns of around 1.97%. An additional measure of economic significance is that, for a one standard deviation change in NFT_R reported in Panel A (0.37), the expected marginal change in recommendation profitability is approximately 37 basis points (= $0.37 \times 0.99\%$), equivalent to a spread in excess of market returns of around 74 basis points on an annualized basis. When we introduce fixed effects, NFT_R remains robustly statistically significant; the marginal effect of NFT_R increases slightly after including brokerage house fixed effects (in Column (2)) but falls after including firm-analyst fixed effects (in Column (3)). Although the economic significance of NFT_R for stock recommendation profitability is quite modest, the results are nevertheless consistent with NFT capturing the quality of analysts' research after controlling for brokerage house styles and the innate ability of analysts. Note, however, that because the scaling of NFT is performed within year t and uses information from later in the year to define the maximum and minimum values of NFT in the year, we cannot interpret the results in this section as a profitable trading strategy. In Section 7.2 we test the value of NFT to investors in out-of-sample stock recommendation portfolios.

6.4 NFT and market reactions to recommendation revisions

Table 6 reports the descriptive statistics (Panel A) and regression results from the tests of NFT and market reactions to analyst stock recommendation revisions (Panel B). Panel A shows that, in our sample, 35.6%, 11.9%, 23.4%, and 28.7% of recommendation revisions are accompanied by concurrent revisions in EPS, earnings before interest and taxes (EBI), pre-tax income (PRE), and sales (SAL), respectively. Unreported correlations show no evidence of severe multicollinearity.

Table 5 Descriptive statistics and regression results from tests of NFT and stock recommendation profitability

Panel A: Descriptive statistics

	N	Mean	Sd	Q1	Median	Q3
RET_REC	208,261	0.033	0.292	−0.099	0.028	0.162
NFT_R	208,261	0.436	0.372	0.059	0.387	0.778
FREQ_REC_R	208,261	0.489	0.424	0.000	0.500	1.000

Panel B: Regression results from tests of NFT and stock recommendation profitability

	(1)	(2)	(3)
NFT_R	0.985***	1.085***	0.583***
	(5.745)	(6.138)	(2.658)
FE_R	−0.293*	−0.254*	−0.459**
	(−1.956)	(−1.708)	(−2.356)
BOLD_R	0.013	−0.002	−0.111
	(0.087)	(−0.013)	(−0.574)
FEXP_R	1.127***	0.820***	0.100
	(6.437)	(4.730)	(0.296)
FREQ_R	0.615***	0.445***	0.155
	(3.936)	(2.880)	(0.678)
FREQ_REC_R	−0.108	0.061	−0.137
	(−0.717)	(0.395)	(−0.681)
GEXP_R	−0.028	−0.114	−0.487
	(−0.140)	(−0.595)	(−1.154)
LFR_R	0.707***	0.512***	−0.096
	(4.783)	(3.462)	(−0.456)
NFIRM_R	−0.058	0.133	0.276
	(−0.294)	(0.657)	(0.868)
NIND_R	0.207	−0.033	−0.522*
	(1.050)	(−0.171)	(−1.678)
SIZE_R	0.188	1.329***	−0.786**
	(1.115)	(5.773)	(−2.454)
WKDN_R	0.353**	0.369**	0.592***
	(2.367)	(2.454)	(2.867)
Observations	208,261	208,188	182,323
Adjusted R^2	0.001	0.006	0.027
Fixed Effects	No	Broker	Firm-Analyst

This table reports descriptive statistics (Panel A) and regression results from tests of NFT and analyst stock recommendation profitability (Panel B). Specifically, Panel B reports the regression results of the following regression equation:

$$\begin{aligned}100 \times \mathrm{RET\_REC}_{\mathrm{ijt}} = \beta_0 \; + \; & \beta_1 \mathrm{NFT\_R}_{\mathrm{ijt}} + \beta_2 \mathrm{FE\_R}_{\mathrm{ijt}} + \beta_3 \mathrm{BOLD\_R}_{\mathrm{ijt}} \\ & + \beta_4 \mathrm{FEXP\_R}_{\mathrm{ijt}} + \beta_5 \mathrm{FREQ\_R}_{\mathrm{ijt}} + \beta_6 \mathrm{FREQ\_REC\_R}_{\mathrm{ijt}} \\ & + \beta_7 \mathrm{GEXP\_R}_{\mathrm{ijt}} + \beta_8 \mathrm{LFR\_R}_{\mathrm{ijt}} + \beta_9 \mathrm{NFRM\_R}_{\mathrm{ijt}} \\ & + \beta_{10} \mathrm{NIND\_R}_{\mathrm{ijt}} + \beta_{11} \mathrm{SIZE\_R}_{\mathrm{ijt}} + \beta_{12} \mathrm{WKDN\_R}_{\mathrm{ijt}} + \text{Fixed Effects} + \epsilon_{\mathrm{ijt}},\end{aligned}$$

where RET_REC is, at most, six months' market-adjusted buy-and-hold return of the stock recommendation issued by analyst j for firm i on day t. It is long in stocks with a strong buy or buy recommendation (coded as 1 and 2 by I/B/E/S) and short in stocks with a hold, sell, or strong sell recommendations (coded as 3, 4, and 5 by I/B/E/S). The return accumulation period runs from the day before the recommendation until the earlier of 180 days or the day before the recommendation is revised or reiterated. Market-adjusted returns are calculated by deducting the value-weighted average return of the market portfolio. We require the stocks to be traded on NYSE, NASDAQ, or AMEX. RET_REC is winsorized at the 1% and 99% levels. NFT is the number of forecast types provided by analyst j for firm i during the three months prior to the month in which the dependent variable is measured. All non-indicator independent variables are scaled using Eq. 1. An intercept is estimated but unreported in each column. The sample period is 2002–2019. Standard errors are clustered at the firm and analyst level. *, **, and *** denote significance at 0.10, 0.05, and 0.01 levels, respectively. Detailed variable definitions are provided in Data Appendix B

Table 6 Descriptive statistics and regression results from tests of NFT and market reactions to analyst stock recommendation revisions

Panel A: Descriptive statistics

	N	Mean	Sd	Q1	Median	Q3
CAR3	95,490	−0.004	0.103	−0.033	−0.001	0.030
NFT_R	95,490	0.458	0.401	0.000	0.400	0.931
RECrev	95,490	0.014	1.315	−1.000	0.000	1.000
EPSrev	95,490	−0.001	0.015	0.000	0.000	0.000
EBIrev	95,490	−0.005	0.125	0.000	0.000	0.000
PRErev	95,490	−0.015	0.233	0.000	0.000	0.000
SALrev	95,490	−0.001	0.030	0.000	0.000	0.000
I_EPSrev	95,490	0.356	0.479	0.000	0.000	1.000
I_EBIrev	95,490	0.119	0.324	0.000	0.000	0.000
I_PRErev	95,490	0.234	0.423	0.000	0.000	0.000
I_SALrev	95,490	0.287	0.452	0.000	0.000	1.000

Panel B: Regression results from tests of NFT and market reactions to recommendation revisions

	(1)	(2)
RECrev	−1.598***	−1.236***
	(−13.450)	(−10.030)
NFT_R	−0.033	−0.012
	(−0.386)	(−0.136)
RECrev × NFT_R	−0.581***	−0.432***
	(−5.400)	(−4.552)
EPSrev		19.180
		(1.520)
EBIrev		1.582***
		(3.279)
PRErev		1.484***
		(4.755)
SALrev		24.710***
		(8.459)
RECrev × I_EPSrev		−0.637***
		(−5.384)
RECrev × I_EBIrev		0.583***
		(4.533)
RECrev × I_PRErev		−0.042
		(−0.338)
RECrev × I_SALrev		−0.465***
		(−3.550)
Observations	95,490	95,490

Table 6 continued

Adjusted R^2	0.217	0.232
Analyst Characteristics	Yes	Yes
Analyst Characteristics × RECrev	Yes	Yes
Fixed Effects	Firm-Year	Firm-Year

This table reports descriptive statistics (Panel A) and regression results (Panel B) from tests of NFT and stock market reactions to analyst recommendation revisions. Specifically, Panel B reports the regression results of the following regression equation:

$$\begin{aligned} 100 \times \mathrm{CAR3}_{\mathrm{ijt}} = {} & \beta_0 + \beta_1 \mathrm{RECrev}_{\mathrm{ijt}} + \beta_2 \mathrm{NFT\_R}_{\mathrm{ijt}} + \beta_3 \mathrm{RECrev}_{\mathrm{ijt}} \times \mathrm{NFT\_R}_{\mathrm{ijt}} \\ & + \gamma \mathit{OtherForecastRevisions} + \delta(\mathrm{RECrev} \times \mathrm{I\_}\mathit{OtherForecastRevisions}) \\ & + \theta \mathit{AnalystCharacteristics\_R} + \eta(\mathrm{RECrev} \times \mathit{AnalystCharacteristics\_R}) \\ & + \text{Firm-Year Fixed Effects} + \epsilon_{ijt}, \end{aligned}$$

where CAR3 is three-day market-adjusted abnormal returns around stock recommendation revisions, RECrev is the level of stock recommendation revisions, NFT is the number of forecast types provided by analyst j for firm i during the three months prior to the month in which the dependent variable is measured, *OtherForecastRevisions* = {EPSrev, EBIrev, PRErev, SALrev}, I_*OtherForecastRevision* = {I_EPSrev, I_EBIrev, I_PRErev, I_SALrev}, and *Analyst Characteristics* = {FE_{t-1}, BOLD, FEXP, FREQ, FREQ_REC, GEXP, LFR, NFRM, NIND, SIZE, WKDN}. All non-indicator independent variables are scaled using Eq. 1 and denoted by the suffix "_R." An intercept is estimated but unreported in each column. Estimated coefficients for analyst characteristics and their interactions with stock recommendation revisions are not reported to conserve space. The sample period is 2002–2019. Standard errors are clustered at the firm and analyst level. t-statistics are reported in the parentheses. *, **, and *** denote significance at 0.10, 0.05, and 0.01 levels, respectively. Detailed variable definitions are provided in Data Appendix B

Table 6 reports the regression results for Eq. 5. The results in Column (1) show that, even after controlling for the analyst characteristics identified in the prior literature and for firm-year fixed effects but without controlling for other forecast revisions, the coefficient on the interaction between RECrev and NFT_R is −0.58 with a t-statistic of −5.40. In terms of economic significance, this indicates that a one category recommendation revision by the highest NFT_R analyst is, on average, associated with a three-day market-adjusted return (CAR3) that is 0.55% (=0.58−0.03) higher compared to the lowest NFT_R analyst, equivalent to a 34.4% (=0.55/1.60) higher market reaction. In Column (2), we further control for other contemporaneous analyst forecast revisions, and the coefficient for RECrev × NFT_R remains statistically and economically significant (Coef.=−0.43, t-stat=−4.55), indicating that NFT_R explains market reactions to stock recommendation revisions incrementally to the information conveyed by analysts' issuance of other forecast revisions.[29]

6.5 NFT and analyst career outcomes

Table 7 Panel A reports descriptive statistics for the variables used in tests of NFT and analyst career outcomes. NFT in this section is the average rank of the number of forecast types issued by the analyst in a given year, as described in Section 3.7. Panel B

[29] Keung (2010) documents that the market reacts more strongly to *earnings forecast* revisions accompanied by sales forecast revisions. Our results extend the findings in Keung (2010) by showing that the market also reacts more strongly to *stock recommendation revisions* accompanied by concurrent sales forecast revisions.

Table 7 Descriptive statistics and univariate analysis of NFT and analyst career outcomes

Panel A: Summary statistics

	N	Mean	Sd	Q1	Median	Q3
TERMINATION	58,246	0.193	0.394	0.000	0.000	0.000
PROMOTION	13,923	0.011	0.103	0.000	0.000	0.000
DEMOTION	15,775	0.012	0.110	0.000	0.000	0.000
NFT_M	58,246	0.431	0.281	0.204	0.416	0.639
I_AFLTG	58,246	0.401	0.490	0.000	0.000	1.000
I_AFSAL	58,246	0.891	0.312	1.000	1.000	1.000
FE_M	58,246	0.344	0.203	0.209	0.308	0.441
BOLD_M	58,246	0.355	0.179	0.240	0.332	0.440
COMP_M	58,246	0.474	0.247	0.289	0.489	0.654
FREQ_M	58,246	0.435	0.226	0.260	0.445	0.592
GEXP_M	58,246	0.331	0.295	0.074	0.252	0.537
LFR_M	58,246	0.274	0.202	0.125	0.236	0.379
NFRM_M	58,246	0.289	0.255	0.058	0.242	0.460
NIND_M	58,246	0.283	0.263	0.000	0.233	0.467
SIZE_M	58,246	0.327	0.288	0.077	0.247	0.520
WKDN_M	58,246	0.494	0.196	0.393	0.495	0.596

Panel B: Average analyst career outcomes conditional on NFT

NFT_Q	[0,25)	[25,50)	[50,75)	[75,100]	High-Low	t-statistic
TERMINATION	0.282	0.191	0.159	0.133	−0.149	−32.123
PROMOTION	0.003	0.012	0.016	0.014	0.010	4.903
DEMOTION	0.012	0.011	0.013	0.013	0.001	0.287

Panel A reports summary statistics for variables used in the tests of NFT and analyst career outcomes. Panel B reports average analyst career outcomes conditional on NFT. NFT_M is the average rank of the number of forecast types issued by analyst j in year t. Specifically, we first use Eq. 1 to scale the number of forecast types issued by analyst j for all firms she follows in year t. We then take the average of the scaled number from the previous step across all firms she follows in year t. I_AFLTG (I_AFSAL) equals one if analyst j provides at least one long-term earnings growth (one-year-ahead sales) forecast for any firm she follows in year t, zero otherwise. All other non-indicator variables are first scaled by Eq. 1 and are then averaged across all firms followed by the analyst in year t. NFT_Q is the cross-sectional percentile rank of NFT. The sample period is 2002–2019. Detailed variable definitions are provided in Data Appendix B

shows the average analyst career outcomes conditional on NFT_Q, the cross-sectional quartiles of NFT_M. The likelihood of termination (promotion) decreases (increases) in NFT_Q, and differences between the top and bottom quartiles are statistically significant. In contrast, analyst demotion is not associated with NFT_Q.

Table 8 Panel B shows the results for the conditional logistic regressions based on Eq. 6. Consistent with the prior literature, several analyst characteristics are related to career outcomes, including the provision of long-term earnings growth forecasts (I_AFLTG),[30] earnings forecast accuracy (FE_M), forecast boldness (BOLD_M),

30 Unreported results show that sales forecasts provision significantly reduces the probability of an analyst's termination without NFT in the regression. NFT subsumes the effects of sales forecasts provision in Eq. 6.

Table 8 Regression results from tests of NFT and analyst career outcomes

CareerOutcome	(1) TERMINATION	(2) PROMOTION	(3) DEMOTION
NFT_M	−0.746***	1.243**	0.186
	(−8.475)	(2.458)	(0.465)
I_AFLTG	−0.301***	−0.365	−0.237
	(−4.255)	(−1.232)	(−1.175)
I_AFSAL	0.026	0.637	0.551
	(0.501)	(1.062)	(1.096)
FE_M	2.176***	−0.354	1.835***
	(23.760)	(−0.488)	(4.083)
BOLD_M	−0.804***	−1.206**	−0.678
	(−7.237)	(−2.382)	(−1.499)
COMP_M	−0.030	−0.512	0.733**
	(−0.533)	(−0.882)	(1.960)
FREQ_M	−2.238***	−0.580	−2.735***
	(−16.710)	(−0.790)	(−4.552)
GEXP_M	0.161**	1.033**	0.471*
	(2.125)	(2.344)	(1.818)
LFR_M	−0.181**	−0.257	−0.451
	(−2.082)	(−0.388)	(−1.380)
NFRM_M	−2.029***	0.416	−0.078
	(−19.830)	(0.803)	(−0.188)
NIND_M	−0.316***	−0.864*	−0.188
	(−4.042)	(−1.843)	(−0.455)
SIZE_M	0.068	−0.816	−0.689
	(0.432)	(−0.439)	(−1.064)
WKDN_M	−0.044	0.607	−0.681
	(−0.739)	(1.362)	(−1.640)
Observations	57,783	3,964	15,545
Fixed Effects	Broker	Broker	Broker

Table 8 continued

CareerOutcome	(1) TERMINATION	(2) PROMOTION	(3) DEMOTION
Pseudo R^2	0.158	0.040	0.062
Prob > χ^2	0.000	0.009	0.000

This table reports the regression results for the following equation estimated by the conditional logistic model matched at the brokerage house level:

$$\begin{aligned}\text{CAREER OUTCOME}_{jt+1} = {} & \beta_1\text{NFT\_M}_{jt} + \beta_2\text{I\_AFLTG\_M}_{jt} + \beta_3\text{I\_AFSAL\_M}_{jt} \\ & + \beta_4\text{AFE\_M}_{jt} + \beta_5\text{BOLD\_M}_{jt} \\ & + \beta_6\text{COMP\_M}_{jt} + \beta_7\text{FREQ\_M}_{jt} + \beta_8\text{GEXP\_M}_{jt} \\ & + \beta_9\text{LFR\_M}_{jt} + \beta_{10}\text{NFRM\_M}_{jt} \\ & + \beta_{11}\text{NIND\_M}_{jt} + \beta_{12}\text{SIZE\_M}_{jt} + \beta_{13}\text{WKDN\_M}_{jt} \\ & + \text{Brokerage Fixed Effects} + \epsilon_{jt+1},\end{aligned}$$

where CAREER OUTCOME $\in$ {TERMINATION, PROMOTION, DEMOTION}. NFT_M is the average rank of the number of forecast types issued by analyst j in year t. Specifically, we first use Eq. 1 to scale the number of forecast types issued by analyst j for all firms she follows in year t. We then take the average of the scaled number from the previous step across all firms she follows in year t. I_AFLTG (I_AFSAL) equals one if analyst j provides at least one long-term earnings growth (one-year-ahead sales) forecast for any firm she follows in year t, zero otherwise. All other non-indicator variables are first scaled by Eq. 1 and are then averaged across all firms followed by the analyst in year t. The sample period is 2002–2019. Z-scores are reported in the parentheses. Standard errors are clustered at the brokerage level. *, **, and *** denote significance at 0.10, 0.05, and 0.01 levels, respectively. Detailed variable definitions are provided in Data Appendix B

earnings forecast frequency (FREQ_M), earnings forecast timeliness (LFR_M), and the number of firms (NFRM_M) and industries (NIND_M) covered. However, Table 8 Panel B also reveals that, after controlling for these analyst characteristics, NFT_M is incrementally negatively related to the likelihood of termination (Column (1)) and positively related to the likelihood of promotion (Column (2)). Consistent with the Table 7 Panel B results, Column (3) indicates that NFT_M fails to explain the likelihood of demotion.

7 Out-of-sample tests

7.1 The value of NFT-based consensus forecast revisions

The in-sample ability of NFT to explain analysts' forecast accuracy begs the question of whether investors can exploit the higher quality of high-NFT research in their investment decisions. Our first test addressing this question compares the ability of revisions in consensus earnings forecasts formed using high-NFT forecasts or low-NFT forecasts to predict returns.

At the end of each month, we extract high-NFT consensus and low-NFT consensus earnings per share forecasts for each of FY1 and FY2 from the I/B/E/S detail file. The high-NFT (low-NFT) consensus FY1 and FY2 forecasts are the medians of the FY1 and FY2 forecasts classified as high-NFT (low-NFT) in the firm-month. We define a forecast as high-NFT (low-NFT) if its associated NFT value is strictly above (below) the median NFT in a given firm-month, where NFT is the total number of forecast types provided by the analyst during the three-month period up to the end of the current month. If an analyst does not revise a previous forecast during a month, the old forecast is brought forward and the level of NFT is recomputed as at the end of the current month. High-NFT (low-NFT) consensus revisions are computed for FY1 and FY2 as the difference between the high-NFT (low-NFT) consensus at the end of a month and the high-NFT (low-NFT) consensus at the end of the previous month scaled by end-of-month stock price.

Next, we compute a composite forecast revision signal for each firm-month for the high-NFT and low-NFT consensuses separately. A composite revision blends the corresponding revisions in corresponding FY1 and FY2 consensus EPS forecasts as a weighted average where the weights depend on the forecast horizon of FYI relative to the fiscal year end, as follows: REV_COMP12 $= (\mathrm{m} \times \mathrm{REV1} + (12 - \mathrm{m}) \times \mathrm{REV2})$, where m is the number of months remaining before the end of FY1, REV1 is the revision in the FY1 consensus EPS forecast, and REV2 is the revision in the FY2 consensus EPS forecast. The horizon-dependent weights help mitigate the negative expected revisions resulting from the walk-down in EPS forecast optimism predicted by prior research (Richardson et al. 2004).

Finally, we consider the predictive ability of high-NFT revisions and low-NFT composite revisions for returns over the next month. We form portfolios monthly based on REV_COMP12 for each group. First, we partition revisions into three portfolios based on the sign of REV_COMP12: negative, zero, and positive. Second, we sort non-zero revisions for both negative revisions and positive revisions into quintile portfolios. For each portfolio approach, we report the annualized portfolio excess returns, Sharpe ratios, and estimated portfolio alphas from the Fama and French (2015) model.

Table 9 contains the results of our analysis. Panel A contains results for high-NFT revisions. The three-way portfolio sort reveals that annualized one-month-ahead returns are 2.96% higher for positive revisions than for negative revisions. The corresponding portfolio alphas of −1.27% for the negative revisions portfolio and 2.20% for the hedge portfolio long in positive revisions and short in negative revisions are significant at better than the 5% level. The decile portfolios indicate that sorting on the size of revisions also sorts on excess returns, although not quite monotonically in the case of negative revisions. The hedge portfolio that is long in the most positive revision stocks and short in the most negative revision stocks delivers an annualized excess return of 10.75%. While the individual portfolio alphas are insignificant except for the most negative (NEG5) and most positive (POS5) portfolios, the hedge portfolio alpha of 10.44% indicates that the hedge portfolio excess returns are due to risk differences across portfolios.

Panel B contains comparable results for low-NFT consensus revisions. The excess returns on the three-way sorted portfolios are less extreme, and the hedge portfolio excess return (equal to 1.35%) is only 45% of the corresponding high-NFT return. Portfolio alphas are uniformly insignificant, indicating no abnormal performance after controlling for portfolio risk. Only one of the 10 portfolio alphas is significant, and the POS5–NEG5 hedge portfolio has a statistically insignificant alpha of 2.23%, around one-fifth of the corresponding high-NFT hedge portfolio return. The difference between the returns on the high-NFT POS5–NEG5 hedge portfolio and the low-NFT POS5–NEG5 hedge portfolio, which is equal to an 8.21% annualized return, is statistically significant at better than the 1% level.

Table 9 Annualized performance of value-weighted portfolios sorted on composite EPS revisions

	(1) # Stocks	(2) Ret-Rf	(3) Sd	(4) t-stat	(5) Sharpe Ratio	(6) FF5 α	(7) t(FF5 α)
Panel A: High-NFT portfolios							
NEG	718	6.073	15.393	1.493	0.395	−1.265	−2.082
ZERO	658	7.605	14.890	1.832	0.511	−0.340	−0.380
POS	832	9.030	13.611	2.559	0.663	0.937	1.471
POS-NEG		2.958	6.184	2.322	0.478	2.202	1.994
NEG5	143	1.929	27.944	0.242	0.069	−8.368	−2.355
NEG4	144	5.435	20.448	0.988	0.266	−2.686	−1.382
NEG3	144	7.118	17.792	1.549	0.400	−0.133	−0.087
NEG2	144	5.609	15.736	1.360	0.356	−1.584	−1.135
NEG1	144	7.402	13.385	2.153	0.553	0.525	0.492
POS1	166	8.109	12.899	2.612	0.629	0.210	0.183
POS2	166	8.226	14.751	1.980	0.558	0.161	0.128
POS3	166	10.043	14.392	2.870	0.698	2.279	1.379
POS4	166	11.026	15.423	2.752	0.715	3.186	2.000
POS5	167	12.676	21.206	2.308	0.598	2.076	1.055
POS5-NEG5		10.747	17.852	2.756	0.602	10.444	3.185
Panel B: Low-NFT portfolios							
NEG	632	6.887	15.353	1.640	0.449	−0.921	−1.407
ZERO	731	7.829	14.311	2.090	0.547	−0.291	−0.302
POS	757	8.239	13.834	2.294	0.596	0.510	0.768
POS-NEG		1.352	5.537	0.934	0.244	1.431	1.175
NEG5	126	5.913	26.291	0.827	0.225	−5.091	−1.585
NEG4	126	3.418	20.457	0.622	0.167	−5.344	−2.544
NEG3	126	8.744	17.526	1.844	0.499	0.670	0.402
NEG2	126	6.978	15.922	1.675	0.438	0.072	0.057
NEG1	127	8.113	13.090	2.378	0.620	0.739	0.756
POS1	151	7.507	13.296	2.185	0.565	−0.519	−0.436
POS2	151	7.498	14.284	2.188	0.525	0.558	0.419
POS3	151	9.374	15.116	2.403	0.620	1.543	1.029
POS4	151	8.407	16.256	1.990	0.517	0.183	0.131
POS5	152	6.571	20.210	1.211	0.325	−2.858	−1.367
POS5-NEG5		0.658	15.319	0.172	0.043	2.233	0.632

Table 9 continued

Panel C: Tests for Differences Between Alphas

	POS-NEG	**POS5-NEG5**
Alpha_H-Alpha_L	0.771	8.211***
t-statistic	0.702	2.708

Panels A and B report annualized performance of portfolios sorted on composite consensus EPS revisions, for high- and low-NFT consensus revisions, respectively. Columns (1)–(7) in Panels A and B report, respectively, the average number of stocks, annualized average excess return (%/yr), annualized standard deviation of excess return (%/yr), t-statistic of excess returns, annualized Sharpe ratio, annualized Fama-French five-factor alpha (%/yr), and t-statistic of Fama-French five-factor alphas, for each portfolio. Panel C reports the difference in Fama-French five-factor alphas between High- and Low-NFT POS-NEG and POS5-NEG5 portfolios and the t-statistics of the difference between Fama-French five-factor alphas. Composite EPS revision is calculated as follows: REV_COMP12 $= (m \times \text{REV1} + (12 - m) \times \text{REV2})$, where REV1 and REV2 are revisions in FY1 and FY2 consensus EPS forecasts, respectively. A revision in consensus forecast is defined as current month consensus forecast minus previous month consensus forecast scaled by current month stock price. A high-NFT (low-NFT) consensus forecast is the median of outstanding high-NFT (low-NFT) forecasts. A forecast is denoted as high-NFT (low-NFT) if the associated NFT value is strictly above (below) the median NFT within a given firm-month. NFT is the number of forecast types provided by the analyst for the firm in a given month during the three-month period to the end of the current month. At the end of each month, stocks are sorted into NEG, ZERO, and POS portfolios if REV_COMP12 is negative, zero, and positive, respectively. Within the NEG (POS) portfolio, stocks are further sorted into quintile portfolios NEG5–NEG1 (POS1–POS5) on REV_COMP12, where NEG5 is the most negative quintile and POS5 is the most positive quintile. Stocks with a price below $5 are excluded and portfolios are value-weighted. Newey-West standard errors are used in t-statistics to adjust for serial correlation (maximum lag = 6 m). *, **, and *** denote significance at 0.10, 0.05, and 0.01 levels, respectively

Overall, these results illustrate the potential usefulness of NFT to investors wishing to exploit analysts' forecasts in investment decisions. A revisions-based investment strategy exploiting information in NFT can deliver economically significant investment returns on a gross basis before consideration of transaction costs. The tests are conservative in assuming that investors wait to trade on consensus forecast revisions until the end of each month. It is possible that higher-frequency updating of consensus revisions and more frequent portfolio updating further enhance portfolio performance. Our purpose here is not to identify an optimal strategy, but to illustrate the potential value, to investors, of identifying higher-quality analyst research using NFT.

7.2 Dynamic recommendation portfolios conditioned on NFT

Since earnings forecasts are an important input to stock recommendations (Bradshaw 2011), the out-of-sample superiority of high-NFT earnings forecasts suggests that high-NFT recommendations will also be more profitable than low-NFT recommendations. The second test of the potential usefulness of NFT to investors compares the out-of-sample profitability of dynamic high-NFT and low-NFT recommendation portfolios. The portfolios are dynamic because they are updated daily using stock recommendations as they are issued by individual analysts.

We classify recommendations as high-NFT or low-NFT using only information available at the time a recommendation is issued.[31] A recommendation for firm i by analyst j at time t is associated with NFT_{ijt} set equal to the number of forecast types issued for firm i by analyst j in the three months before the recommendation month. Specifically, we define $NFT\_P_{ijt} = (NFT_{ijt} - \min(NFT_{it}))/(\max(NFT_{it}) - \min(NFT_{it}))$, where $\min(NFT_{it})$ is the minimum (maximum) value of NFT and $\max(NFT_{it})$ is the maximum value of NFT across all recommendations issued for firm i over the six months prior to recommendation month t. The high-NFT portfolio includes recommendations with $NFT\_P \geq 0.7$, and the low-NFT portfolio includes recommendations with $NFT\_P \leq 0.3$.[32]

We include a stock in a buy (long) portfolio from the first day after an analyst issues a buy recommendation or strong buy recommendation for that stock. Similarly, we include a stock in a sell (short) portfolio from the first day after an analyst issues a sell or strong sell recommendation. A stock is held in a portfolio for a maximum of 30 days (or, alternatively, 180 days) unless the same analyst issues a new recommendation for the same stock. If the new recommendation confirms the previous portfolio assignment and the analyst is still in the same (high or low) NFT category, then the recommendation is treated as if it were new, and the maximum holding period is reset to 30 (or 180) days. If the new recommendation or its NFT category indicates a different portfolio assignment, then the previous portfolio position is closed and a new portfolio position is opened. Following Barber et al. (2007), a portfolio will include more than one position in the same stock when other analysts with the same level of NFT issue similar recommendations for the same stock.

To examine the profitability of portfolios, we use a calendar time portfolio methodology similar to Barber et al. (2007), Cohen et al. (2010), and Coleman et al. (2022) to compute daily portfolio returns. The daily return on a portfolio is the value-weighted return on the individual stock recommendation in the portfolio each day, defined as

$$R_{pt} = \frac{\sum_{i=1}^{n_t} w_{it} R_{it}}{\sum_{i=1}^{n_t} w_{it}},$$

where n_{it} is the number of stocks in the portfolio on day t, and the weights on each recommendation position are the values of each recommendation position $w_{it} = w_{it-1}(1 + R_{it-1})$. We set the initial investment weight on the day a new recommendation enters a portfolio as $w_{it} = 1$. We assess the potential value of NFT to investors who follow recommendations by comparing the performance of the buy (B), sell (S), and long-short (Hedge) portfolios for both high-NFT (H) and low-NFT (L) recommendations. Hence, the portfolios of interest are labelled BH, BL, SH, SL, Hedge_H, and Hedge_L.

Table 10 reports the results from this analysis. To facilitate economic interpretation, daily excess returns, standard deviations, and Sharpe ratios are annualized. Panel A

[31] The standardization of NFT used in the regression results reported earlier introduces a potential look-ahead bias because NFT_R is scaled using information that is only fully known by the end of a year (i.e., the maximum and minimum NFT within the firm-year).

[32] Robustness checks confirm our results are qualitatively similar when we form portfolios using a cutoff of NFT_P equal to 0.5, hence including all recommendations for which NFT_P is available in one of the portfolios.

Table 10 Annualized performance of calendar time recommendation portfolios

Panel A: Maximum holding window = [t+1d, t+30d]							
	Market	**BH**	**BL**	**SH**	**SL**	**Hedge_H**	**Hedge_L**
Avg. # Stocks		162	180	41	43		
Avg. Exc. ret, %	8.644	18.850	16.934	−3.301	3.528	22.151	13.406
Sd, %	18.573	21.907	22.224	25.082	26.828	11.271	13.652
t-statistic	2.189	3.818	3.389	−0.557	0.549	8.226	3.914
Sharpe Ratio	0.465	0.860	0.762	−0.132	0.132	1.965	0.982
Panel B: Maximum holding window = [t+1d, t+180d]							
	Market	**BH**	**BL**	**SH**	**SL**	**Hedge_H**	**Hedge_L**
Avg. # Stocks		911	1,019	227	233		
Avg. Exc. ret, %	8.644	11.164	11.063	5.922	8.366	5.242	2.671
Sd, %	18.573	21.748	21.907	23.336	23.336	5.874	6.350
t-statistic	2.189	2.274	2.251	1.086	1.529	3.457	1.620
Sharpe Ratio	0.465	0.513	0.505	0.254	0.359	0.892	0.421

Panels A and B show the sizes and performance of the following calendar time portfolios when the maximum holding windows are [t+1d, t+30d] and [t+1d, t+180d], respectively: the market portfolio (Market), high-NFT buy recommendation portfolio (BH), low-NFT buy recommendation portfolio (BL), high-NFT sell recommendation portfolio (SH), low-NFT sell recommendation portfolio (SL), high-NFT hedge portfolio (Hedge_H, BH minus SH) and low-NFT hedge portfolio (Hedge_L, BL minus SL). High- and low-NFT portfolios are updated daily based on NFT_P. Specifically, for each recommendation, we define $NFT\_P_{ijt} = (NFT_{ijt} - \min(NFT_{it}))/(\max(NFT_{it}) - \min(NFT_{it})$, where NFT_{ijt} is the number of forecast types issued for firm i by analyst j during the three months prior to the recommendation announcement month, and $\min(NFT_{it})$ $(\max(NFT_{it}))$ is the minimum (maximum) NFT across all analysts who have issued at least one recommendation for firm i during the six months prior to the recommendation announcement month. The high-NFT (low-NFT) portfolio consists of recommendations with an NFT_P greater (smaller) than or equal to 0.7 (0.3). The recommendation is included in the related portfolio one day after the announcement of the recommendation. The recommendation exits the related portfolio when either 1) the holding period attains 30 (180) days, or 2) a new recommendation of a different direction is announced. The sample period is 2002–2019

reports the results for the shorter maximum holding period of 30 days, while Panel B contains the corresponding results for the longer maximum holding period of 180 days. Consistent with the prior literature, buy recommendations are more frequent than sell recommendations. In Panel A, the high-NFT buy portfolio (BH) has an annualized return of 18.85% and a Sharpe ratio of 0.86, compared to 16.93% and 0.76 for the low-NFT buy portfolio (BL).[33] The superior performance of the high-NFT portfolio is even stronger for the sell portfolio. The high–NFT sell portfolio has a return of −3.30% and a Sharpe ratio of −0.13, compared to 3.53% and 0.13 for the low-NFT sell portfolio. This difference of nearly 7% in the performance of high-NFT sell recommendations compared to low-NFT sell recommendations is substantial. The differences in performance between the high-NFT and low-NFT portfolios are emphasized when one compares the return of 22.51% on the high-NFT hedge portfolio (Hedge_H) with the return of 13.41% on the low-NFT hedge portfolio (Hedge_L); the Sharpe ratio of the Hedge_H portfolio is twice that of the Hedge_L portfolio.

[33] Both high-NFT and low-NFT buy portfolios have higher Sharpe ratios than the market index.

Panel B reveals that when the maximum holding period is increased to 180 days, returns on buy portfolios are lower and those on sell portfolios are higher than for the corresponding portfolios in Panel A. The results suggest that, over the longer-term holding period from day +31 to day +180, there is considerably more noise in portfolio returns and possibly some reversal in the initial 30-day holding period returns patterns. Nevertheless, Panel B is still consistent with high-NFT recommendations having higher value to investors than low-NFT recommendations. The returns on the sell and hedge portfolios are considerable better for high-NFT recommendations than for low-NFT recommendations, and the Sharpe ratio for the high-NFT recommendation hedge portfolio is still more than twice that for the low-NFT recommendation hedge portfolio.

The results in Table 10 show that high-NFT recommendations are more profitable than low-NFT recommendations. But it is possible that differences in returns across portfolios result from differences in risk across portfolios. To address this concern, we estimate portfolio alphas from time-series regressions of daily portfolio returns using the Fama and French (2015) five-factor asset-pricing model. To test for differences in alphas, we regress the differences between comparable high-NFT and low-NFT portfolio returns (i.e., BH–BL, SH–SL and Hedge_H–Hedge_L) on the Fama and French (2015) factors and examine the significance of the constants.[34] Table 11 presents the results. Panel A reports the results for the shorter maximum holding period of 30 days, while Panel B contains the corresponding results for the longer maximum holding period of 180 days.

Panel A indicates that all buy and sell portfolios load positively on SMB (i.e., they are biased towards smaller stocks); all except the high-NFT buy portfolio load positively on HML (i.e., they are biased towards value stocks); all load negatively on RMW (i.e., they are biased towards weaker profitability stocks); but only the low-NFT buy portfolio loads on the CMA factor (i.e., is biased towards low investment growth stocks). The two hedge portfolios load negatively on both SMB and HML (i.e., they are tilted towards large growth stocks) and positively on RMW (i.e., they are tilted towards more profitable stocks), while Hedge_L loads negatively at the 10% level on CMA (i.e., it is tilted towards more conservative investment growth stocks). Factor loadings for the longer 180-day maximum holding period are qualitatively quite similar.

After controlling for portfolio exposures to the Fama and French (2015) factors, all portfolio alphas for the 30-day maximum holding period (Panel A) are statistically significant. Comparisons between high-NFT and low-NFT portfolios reveal that the alpha for the high-NFT buy portfolio is higher and for the sell portfolio is lower than the alphas of the corresponding low-NFT portfolios, and the differences are significant at better than the 10% level. This indicates superior performance of high-NFT recommendations. Consequently, the high-NFT hedge portfolio alpha for the 30-day maximum holding period is significantly higher than the low-NFT hedge portfolio alpha, and the difference is significant at less than the 5% level. Over 75% of the outperformance of the high-NFT portfolio is attributable to the higher performance of the sell portfolio. Results are qualitatively similar for the 180-day maximum holding

[34] The inferences are identical if we estimate portfolio alphas in seemingly unrelated regression systems using the Stata sureg command and then test whether estimated alphas are equal.

Table 11 Regressions of calendar time portfolio returns on Fama-French 5 factors

Panel A: Maximum holding window = [t+1d, t+30d]

	(1)	(2)	(3)	(4)	(5)	(6)
	High-NFT			Low-NFT		
	Buy	Sell	Hedge	Buy	Sell	Hedge
1000 × Alpha	0.382***	−0.489***	0.870***	0.299***	−0.231*	0.529***
	(8.370)	(−4.896)	(8.653)	(7.040)	(−1.739)	(4.067)
Mkt-Rf	1.048***	1.046***	0.003	1.069***	1.068***	0.001
	(123.150)	(81.400)	(0.189)	(133.977)	(69.093)	(0.078)
SMB	0.512***	0.561***	−0.049*	0.478***	0.562***	−0.083**
	(34.786)	(21.948)	(−1.701)	(28.185)	(16.207)	(−2.297)
HML	−0.014	0.260***	−0.274***	0.058***	0.318***	−0.260***
	(−0.788)	(9.479)	(−8.333)	(3.999)	(7.290)	(−5.932)
RMW	−0.002***	−0.003***	0.001*	−0.002***	−0.003***	0.001**
	(−13.148)	(−6.155)	(1.731)	(−11.926)	(−6.148)	(2.477)
CMA	−0.000	−0.001	0.001	−0.001***	−0.000	−0.001*
	(−0.437)	(−1.181)	(0.942)	(−4.324)	(−0.230)	(−1.732)
Observations	4,531	4,531	4,531	4,531	4,531	4,531
R^2	0.951	0.815	0.067	0.955	0.750	0.055
Tests for differences between High- and Low-NFT Alphas						
				Buy	Sell	Hedge
1000 × (Alpha_H-Alpha_L)				0.083	−0.258	0.341
t-statistic				1.866*	−1.883*	2.420**

Panel B: Maximum holding window = [t+1d, t+180d]

	(1)	(2)	(3)	(4)	(5)	(6)
	High-NFT			Low-NFT		
	Buy	Sell	Hedge	Buy	Sell	Hedge
1000 × Alpha	0.076***	−0.133**	0.210***	0.060**	−0.041	0.101*
	(2.747)	(−2.269)	(3.878)	(2.132)	(−0.652)	(1.719)
Mkt-Rf	1.056***	1.058***	−0.002	1.070***	1.058***	0.012
	(249.888)	(139.319)	(−0.218)	(197.116)	(118.089)	(1.077)
SMB	0.516***	0.559***	−0.043***	0.490***	0.539***	−0.049***
	(59.812)	(36.960)	(−2.804)	(43.214)	(32.456)	(−2.727)
HML	−0.020*	0.188***	−0.208***	0.020	0.217***	−0.197***
	(−1.780)	(13.276)	(−11.955)	(1.637)	(12.539)	(−9.540)
RMW	−0.002***	−0.003***	0.001**	−0.002***	−0.003***	0.001***
	(−20.605)	(−10.527)	(2.547)	(−16.871)	(−9.830)	(3.906)
CMA	−0.000	0.001*	−0.001***	−0.000**	0.001***	−0.001***
	(−0.730)	(1.844)	(−2.591)	(−2.324)	(3.174)	(−4.580)
Observations	4,531	4,531	4,531	4,531	4,531	4,531
R^2	0.980	0.938	0.173	0.979	0.928	0.158

Table 11 continued

Tests for differences between High- and Low-NFT Alphas			
	Buy	**Sell**	**Hedge**
1000 × (Alpha_H − Alpha_L)	0.016	−0.092	0.109
t-statistic	0.880	−1.921*	2.117**

Panels A and B show the results from time-series regressions of calendar time portfolio returns on Fama-French five factors when the maximum holding windows are [t+1d t+30d] and [t+1d t+180d], respectively. The calendar time portfolios through columns (1) to (6) are the high-NFT buy recommendation portfolio (BH), high-NFT sell recommendation portfolio (SH), high-NFT hedge portfolio (Hedge_H, BH minus SH), low-NFT buy recommendation portfolio (BL), low-NFT sell recommendation portfolio (SL), and low-NFT hedge portfolio (Hedge_L, BL minus SL). High- and low-NFT portfolios are updated daily based on NFT_P. Specifically, for each recommendation, we define $NFT\_P_{ijt} = (NFT_{ijt} - \min(NFT_{it}))/(\max(NFT_{it}) - \min(NFT_{it})$, where NFT_{ijt} is the number of forecast types issued for firm i by analyst j during the three months prior to the recommendation announcement month, and $\min(NFT_{it})$ ($\max(NFT_{it})$) is the minimum (maximum) NFT across all analysts who have issued at least one recommendation for firm i during the past six months prior to the recommendation announcement month. The high-NFT (low-NFT) portfolio consists of recommendations with an NFT_P greater (smaller) than or equal to 0.7 (0.3). The recommendation is included in the related portfolio one day after the announcement of the recommendation. The recommendation exits the related portfolio when either 1) the holding period attains 30 (180) days, or 2) a new recommendation of a different direction is announced. Alpha is in percentage. The sample period is 2002–2019. t-statistics are reported in parentheses. Newey-West standard errors (maximum lag = 5) are used to adjust for serial correlation. *, **, and *** denote significance at the 0.10, 0.05, and 0.01 levels, respectively

period portfolios, but the alpha values are lower. This suggests that the investment value of the recommendations is higher in the first 30 days. Moreover, the high-NFT and low-NFT buy portfolio alphas are not statistically different over the longer maximum holding period. Nevertheless, because the high-NFT sell portfolio outperforms the low-NFT sell portfolio, the high-NFT hedge portfolio also outperforms over the 180-day maximum holding period.

To assess the economic significance of our results, we focus on the 30-day maximum holding period portfolios. The high-NFT hedge portfolio Hedge_H has an annualized Sharpe ratio of 1.97 and an estimated alpha of 8.70 bps per day (t-statistic = 8.65), equivalent to an annualized abnormal return of 21.92%. This accounts for nearly all the excess return reported in Table 10. In contrast, the low-NFT hedge portfolio Hedge_L has an annualized Sharpe ratio of 0.98 and an estimated alpha of 5.29 bps per day (t-statistic = 4.07), equivalent to an annualized return of 13.33%. The difference in alphas of 0.341 is equivalent to an annualized alpha of 8.6%. The potential contribution, to investment performance, of information on NFT is therefore considerable. Figure 2 compares the cumulative hedge portfolio returns over the 30-day maximum holding period of high-NFT recommendations and low-NFT recommendations.

These results come with some caveats. First, the results do not allow for transaction costs. Second, our results assume that investors can trade on recommendations at the close of the day a recommendation is announced (the first day returns are earned is day +1). If investors cannot trade on new recommendations until later, then returns will be lower if the market reacts on day +1. In unreported tests, we find that returns to strategies implemented one day later are lower than those reported but still econom-

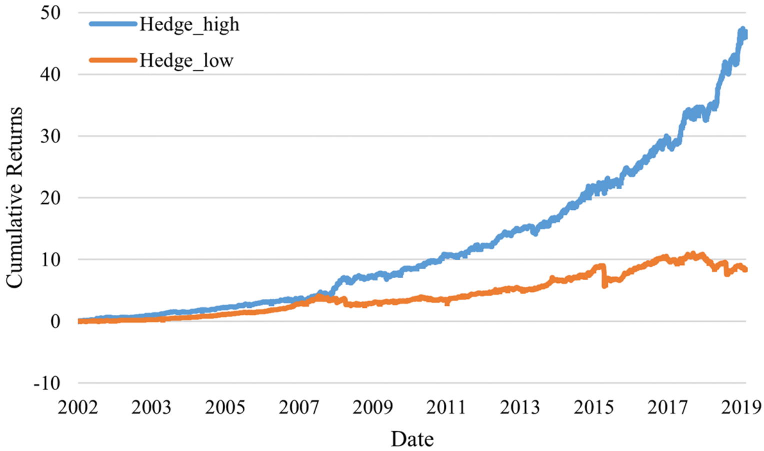


Fig. 2 Cumulative returns for hedge portfolios long in high-NFT (low-NFT) buy recommendations and short in high-NFT_P (low-NFT) sell recommendations. High- and low-NFT portfolios are updated daily based on NFT_P. Specifically, for each recommendation we define $NFT\_P_{ijt} = (NFT_{ijt} - \min(NFT_{it}))/(\max(NFT_{it}) - \min(NFT_{it}))$, where NFT_{ijt} is the number of forecast types issued for firm i by analyst j during the three months prior to the recommendation announcement month, and $\min(NFT_{it})$ $(\max(NFT_{it}))$ is the minimum (maximum) NFT across all analysts who have issued at least one recommendation for firm i during the six months prior to the recommendation announcement month. The high-NFT (low-NFT) portfolio consists of recommendations with an NFT_P greater (smaller) than or equal to 0.7 (0.3). The recommendation is included in the related portfolio one day after the announcement of the recommendation. The recommendation exits the related portfolio when either 1) the holding period attains 30 days, or 2) a new recommendation of a different direction is announced

ically and statistically significant. Moreover, high-NFT recommendations continue to outperform. On the other hand, Irvine et al. (2007), Juergens and Lindsey (2009), and Christophe et al. (2010) suggest that sell-side analysts share information privately with preferred clientele before they publicly release their forecasts or recommendations. In unreported results, we also find that investors would realize significantly higher performance than we report if they were able to trade on recommendations one or two days earlier than the public release dates.[35]

8 Additional tests

8.1 Horse-race tests: NFT v. individual forecast types

Prior studies have shown that individual non-EPS forecast items, including cash flow, long-term earnings growth, and sales forecasts, can signal research ability and effort (Call et al. 2009; Ertimur et al. 2011; Jung et al. 2012; Keung 2010). To mitigate possible concerns that the associations between NFT and analyst research quality documented above are driven by specific forecast types included in NFT, we implement

[35] Results of these unreported tests are available from the authors on request.

Table 12 Regression results for tests of NFT and analyst research quality controlling for the issuance of individual forecast types

Panel A: Means of analysts' issuance of individual forecast types			
	(1)	**(2)**	**(3)**
Dependent Variable	**100×FE_MA**	**100×PTE_MA**	**100×RET_REC**
I_AFCPS	0.171	0.183	0.131
I_AFLTG	0.096	0.109	0.111
I_AFSAL	0.752	0.770	0.639
Panel B: Regression results			
NFT_R	−4.596***	−0.785*	1.016***
	(−8.691)	(−1.845)	(4.484)
I_AFCPS	0.906	−0.265	−0.690***
	(1.511)	(−0.653)	(−2.852)
I_AFLTG	−0.355	0.308	−0.060
	(−0.633)	(0.938)	(−0.307)
I_AFSAL	0.213	−0.632**	0.294
	(0.449)	(−2.045)	(1.580)
Observations	390,140	841,525	208,188
Adjusted R^2	0.108	0.003	0.007
Control Variables	Yes	Yes	Yes
Fixed Effects	Broker	Broker	Broker

This table reports the results from the analysis of earnings forecast accuracy (Column (1), based on Eq. 2), price target forecast accuracy (Column (2), based on Eq. 3), and stock recommendation profitability (Column (3), based on Eq. 4), after controlling for analysts' issuance of one-year-ahead cash flow forecasts, long-term earnings-growth forecasts, and one-year-ahead sales forecasts. Specifically, I_AFCPS (I_AFLTG, I_AFSAL) is an indicator variable set equal to one if analyst j issued at least one cash flow per share (long-term earnings-growth, sales) forecast during the three months prior to the month in which the dependent variable is measured, zero otherwise. Panel A reports the sample averages of I_AFCPS, I_AFLTG, and I_AFSAL in Columns (1)–(3), respectively. Panel B reports the regression results. The sample period is 2002–2019. t-statistics are reported in parentheses. *, **, and *** denote significance at the 0.10, 0.05, and 0.01 levels, respectively. Detailed variable definitions are provided in Data Appendix B

"horse-race" tests to establish the incremental contribution of NFT. We re-estimate Eqs. 2 –7 after including additional indicator variables capturing when an analyst issues forecasts of cash flow, long-term earnings growth, or sales in the previous three months, the period over which NFT is calculated.[36]

Panel A of Table 12 reports the mean values of the indicator variables for analysts' issuance of cash flow, long-term earnings growth, and sales forecast types in the samples employed in tests of earnings forecast accuracy (Column (1)), price target forecast accuracy (Column (2)), and recommendation profitability (Column (3)). In each test sample, fewer than 20% of observations (ranging from 13.1% to 18.3%) are

[36] In the test of stock market reactions to recommendation revisions (Eq. 5), we control for other contemporaneous forecast revisions. In the test of analyst career outcomes (Eq. 6), we control for analysts' issuance of long-term earnings growth forecasts and sales forecasts, which have been shown by prior studies to predict career outcomes.

accompanied by cash flow forecasts in the previous three months. Long-term earnings growth forecasts (I_AFLTG) are even less common (between 9.6% and 11.1% of sample observations). On the other hand, sales forecasts are more common (between 63.9% and 77% of sample observations), consistent with trends documented in the Online Appendix.

Panel B of Table 12 reports the regression results, which are designed to be comparable to the second columns in the main tests reported in Tables 3–5 (i.e., after including brokerage house fixed effects and clustering standard errors by analyst and firm). In the tests of earnings forecast accuracy and stock recommendation profitability, the coefficients on NFT_R remain robustly significant at better than the 1% level after including the individual forecast indicators. In the test of price target accuracy, the coefficient on NFT_R remains significant at better than the 10% level (p-value = 0.065).[37] Moreover, the R^2 statistics are almost identical to the corresponding values in Tables 3–5. Overall, these results suggest that NFT captures incremental information explaining research quality proxies beyond analysts' issuance of cash flow forecasts, long-term earnings growth forecasts, or sales forecasts. Interestingly, only NFT_R is consistently significant across the three research outcomes. These results suggest that NFT is effective as a general indicator of research quality, whereas the provision of cash flow forecasts, LTG, or sales forecasts does not consistently capture differences in research quality across analysts.

8.2 Horizon and breadth effects in NFT

By construction, our NFT measure captures both the breadth of forecast types—that is, the set of unique line items an analyst forecasts (the k-dimension in the definition of NFT in Section 3.1)—and the number of horizons over which an analyst forecasts earnings (the h-dimension). To separately capture the potential effects of the breadth component and the horizon component, we re-estimate Eqs. 2–5 replacing NFT_R as the variable of interest with NFT1_R and/or NFT_HZN. In Table 13 Panel A, we report the descriptive statistics for each test sample of NFT1 and NFT_HZN. The corresponding descriptive statistics for the scaled variables NFT1_R and NFT_HZN_R are reported in Panel B. NFT1 and NFT_HZN each display considerable variation within firm-years and analyst-years.

The regression results are reported in Panel C of Table 13. In models that include both the NFT1_R and NFT_HZN_R components, the coefficients for NFT1_R are statistically significant at the 1% level except in the case of analyst career demotion. These results suggest that the breadth component of NFT is important and that the explanatory power of NFT_R in our main results is not driven by the horizon component. In contrast, the coefficients on NFT_HZN_R are at best weakly significant after controlling for NFT1_R. However, on its own, the horizon component NFT_HZN_R

[37] The statistical significance is sensitive to the method of clustering standard errors. We use two-way clustering at the firm level and the analyst level. This is empirically a more conservative approach to testing statistical significance (Gow et al. 2010). Unreported results using less conservative approaches where we cluster by firm or firm-analyst pair yield t-statistics on NFT_R that are well above 2.0. In addition, unreported results on a subsample where there is no variation in analyst issuance of sales forecasts (I_AFSAL=1) show that NFT_R remains economically and statistically significant.

Table 13 Descriptive statistics and regression results from tests of analyst horizon versus breadth of forecast types

Panel A: Descriptive statistics of raw NFT1 and NFT_HZN

Dep. Variable	100×FE_MA		100×PTE_MA		100×RET_REC		100×CAR3		TERMINATION	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	NFT1	NFT_HZN	NFT1	NFT_HZN	NFT1	NFT_HZN	NFT1	NFT_HZN	NFT1	NFT_HZN
N	374,859	374,859	836,643	836,643	200,361	200,361	86,715	86,715	58,246	58,246
Mean	6.375	2.183	6.697	2.277	5.102	1.970	5.072	2.054	9.854	3.645
Sd	3.885	1.151	3.935	1.192	3.995	1.243	3.862	1.184	4.243	1.080
Q1	3.000	2.000	4.000	2.000	1.000	1.000	1.000	1.000	8.000	3.000
Median	7.000	2.000	7.000	2.000	5.000	2.000	5.000	2.000	10.000	3.000
Q3	9.000	3.000	9.000	3.000	8.000	3.000	8.000	3.000	13.000	4.000

Panel B: Descriptive statistics of scaled NFT1 and NFT_HZN

Dep. Variable	100×FE_MA		100×PTE_MA		100×RET_REC		100×CAR3		TERMINATION	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	NFT1_R	NFT_HZN_R	NFT1_R	NFT_HZN_R	NFT1_R	NFT_HZN_R	NFT1_R	NFT_HZN_R	NFT1_M	NFT_HZN_M
N	374,859	374,859	836,643	836,643	200,361	200,361	86,715	86,715	58,246	58,246
Mean	0.487	0.485	0.507	0.527	0.481	0.524	0.491	0.527	0.495	0.418
Sd	0.344	0.365	0.302	0.298	0.380	0.382	0.405	0.413	0.323	0.351
Q1	0.167	0.000	0.286	0.333	0.083	0.000	0.000	0.000	0.222	0.000
Median	0.500	0.500	0.529	0.500	0.500	0.500	0.500	0.500	0.500	0.333
Q3	0.769	0.750	0.733	0.750	0.846	1.000	1.000	1.000	0.750	0.667

Table 13 continued

Panel C: Regression results from tests of analyst horizon versus breadth of forecast types											
Dep. Variable	**100×FE_MA**		**100×PTE_MA**		**100×RET_REC**		**100×CAR3**		**TERM**	**PROM**	**DEMO**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**
NFT1_R		−4.348***		−1.293***		0.908***					
		(−8.189)		(−3.401)		(4.544)					
NFT_HZN_R	−0.534	0.567	−1.081***	−0.577	0.472***	0.060					
	(−1.320)	(1.357)	(−3.349)	(−1.610)	(2.879)	(0.322)					
RECrev × NFT1_R								−0.369***			
								(−3.681)			
RECrev × NFT_HZN_R							−0.281***	−0.152*			
							(−3.262)	(−1.776)			
NFT1_M									−0.787***	1.171**	0.440
									(−8.635)	(2.338)	(0.985)
NFT_HZN_M									−0.117*	0.376	0.058
									(−1.917)	(0.908)	(0.200)

Table 13 continued

Observations	374,793	374,793	836,595	836,595	200,291	200,291	86,175	86,175	57,783	4,039	15,535
Adjusted/Pseudo R^2	0.110	0.110	0.003	0.003	0.006	0.006	0.232	0.233	0.158	0.042	0.063
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Broker	Broker	Broker	Broker	Broker	Broker	Firm-Year	Firm-Year	Broker	Broker	Broker

This table reports descriptive statistics (Panel A and Panel B) and regression results (Panel B) from the tests of analyst horizon versus breadth of forecast types. Specifically, Panel A shows the number of one-year-ahead forecast types (Columns (1), (3), (5), (7), and (9)) and the number of earnings forecast horizons (Columns (2), (4), (6), (8), and (10)) at the firm-analyst-year level. Columns (1)–(8) of Panel B show the variables in Columns (1)–(8) of Panel A scaled by Eq. 1. The variables in Columns (9) and (10) of Panel B are scaled along the lines of Section 3.7. Panel C shows the regression results from the tests of earnings forecast accuracy (Columns (1) and (2), based on Eq. 2), price target forecast accuracy (Columns (3) and (4), based on Eq. 3), stock recommendation profitability (Columns (5) and (6), based on Eq. 4), stock market reactions to recommendation revisions (Columns (7) and (8), based on Eq. 5), and analyst career outcomes (Column (9) for termination, Column (10) for promotion and Column (11) for demotion, based on Eq. 6), using the number of one-year-ahead forecast types (NFT1) and the number of earnings forecast horizons (NFT_HZN). Specifically, NFT1 is the number of one-year-ahead forecast types issued by analyst j for firm i during the three months prior to the month in which the dependent variable is measured. NFT_HZN is the number of unique earnings forecast horizons, including FY1–FY5 and long-term earnings growth forecasts, provided by analyst j for firm i during the three months prior to the month in which the dependent variable is measured. The sample period is 2002–2019. t-statistics are reported in parentheses. *, **, and *** denote significance at the 0.10, 0.05, and 0.01 levels, respectively. Detailed variable definitions are provided in Data Appendix B is important for the stock price-related outcomes, including price target accuracy, stock recommendation profitability, and market reactions to recommendation revisions. These outcomes are expected to be associated with longer-term information captured by the horizon component.

9 Conclusion

In this study, we first document substantial variation in the number of forecast types provided by sell-side analysts to I/B/E/S for the firms they cover. The provision of forecasts differs between analysts, across the firms analysts cover, and over time, even for the same analyst and firm. We hypothesize that analysts provide additional forecast types selectively as a means of signaling their superior ability and research effort in forecasting firm fundamentals and producing their stock recommendations.

Consistent with our hypothesis, we find that analysts who provide more forecast types have more accurate earnings forecasts and price target forecasts, more profitable stock recommendations, more influential stock recommendation revisions, and better career outcomes. We demonstrate the potential investment value of our results by forming portfolios conditioned on the number of forecast types and exploiting revisions in consensus earnings forecasts and individual analysts' stock recommendations in out-of-sample tests. Our results indicate that the earnings forecast revisions and the recommendations by analysts who provide higher numbers of forecast types are more successful in predicting future stock returns.

Our study extends our understanding of the information in analysts' voluntary forecast provision. Our findings suggest that analysts who provide a broader set of forecast types, on average, are more capable or more diligent and deliver superior research performance. The effects we document extend beyond specific forecast types such as cash flow, long-term earnings growth, and sales.

The results we report suggest an ex ante measure of analyst forecasting performance and research quality. The number of forecast types provided by an analyst is readily observable and applies to all analysts following most firms in the I/B/E/S universe. It does not require a long time series of data to estimate, so it could be particularly useful in identifying higher-quality, less-experienced analysts who do not have long track records.

We also contribute by adding new evidence on the determinants of stock recommendation profitability and price target accuracy—a relatively under-researched area—and we offer new insights on the determinants of stock market reactions to stock recommendation revisions. Finally, our study improves our understanding of how analyst outputs are valued in the labor market for sell-side financial analysts.

Our paper has several limitations. First, we only study forecasts of financial statement line items. We note that I/B/E/S and FactSet (along with other data vendors) collect many more types of forecast, including non-financial KPIs (Givoly et al. 2019; Hand et al. 2021). Future research might consider the roles of those other forecast types and seek to extend the findings in this paper. Second, our paper only considers analysts' forecasts captured in the I/B/E/S database. Future research could compare analysts' provision to I/B/E/S with the forecasts provided in research reports.

Appendix A: Categorization of I/B/E/S forecast types by financial statement sections

Balance sheet forecast types

BPS	book value per share
ENT	enterprise value (non per share)
NAV	net asset value (non per share)
NDT	net debt (non per share)

Cash flow statement forecast types

CPS	cash flow per share
CSH	cash earnings per share
CPX	capital expenditure (non per share)
DPS	dividend per share
FFO	funds from operations per share

Income statement forecast types

EBI	earnings before interest and taxes (non per share)
EBG	earnings per share – before goodwill
EBS	earnings before interest, taxes, depreciation, and amortization per share
EBT	earnings before interest, taxes, depreciation, and amortization (non per share)
EPS	earnings per share
EPX	earnings per share – alternate
GPS	GAAP earnings per share
GRM	gross margin (percent)
NET	net income (non per share)
OPR	operating profit (non per share)
PRE	pre-tax profit (non per share)
ROA	return on asset (percent)
ROE	return on equity (percent)
SAL	revenue/sales (non per share)

Appendix B: Variable definitions

We define below the raw variables used in the study. With the exception of the Career Outcome tests, in all other regression tests all independent variables except for indicator variables are scaled within firm-years using Eq. 1.

Raw variables in tests of NFT and analysts' characteristics before scaling

NFT	the number of forecast types provided by analyst j for firm i in year t.
NFT1	the number of one-year-ahead forecast types provided by analyst j for firm i in year t.

NFT_HZN	the number of unique earnings forecast horizons, including FY1–FY5 and long-term earnings growth forecasts, provided by analyst j for firm i in year t.
ABLT	the average of sign indicators for all earnings forecasts made by analyst j for firm i multiplied by negative one. The sign indicator equals 1, 0, or -1 if the product of the deviation of analyst j's EPS forecast from the consensus and the difference between the consensus EPS forecast and the actual EPS has a positive, zero, or negative sign. The consensus forecast is calculated for each forecast as the average of the latest five outstanding forecasts.
BOLD	the absolute deviation of analyst j's first one-year-ahead EPS forecasts for firm i immediately after year t-1's fiscal year-end from the average of those issued by all other analysts.
CONS	the standard deviation of analyst j's quarterly earnings forecasts for firm i throughout the analyst's professional career multiplied by negative one.
FEXP	the number of years analyst j has provided at least one EPS forecast for firm i through year t.
FREQ	the number of EPS forecasts provided by analyst j for firm i in year t.
GEXP	the number of years analyst j has provided at least one EPS forecast to I/B/E/S through year t.
LFR	the cumulative number of days by which the preceding two one-year-ahead EPS forecasts lead the focal forecast issued by analyst j for firm i during fiscal year t divided by the cumulative number of days by which the subsequent two forecasts follow that forecast.
NFRM	the number of firms for which analyst j provides at least one EPS forecast in year t.
NIND	the number of two-digit SIC industries analyst j covers in year t.
SIZE	the number of analysts employed by analyst j's brokerage house in year t.
STAR	an indicator variable set equal to one if analyst j was nominated as an All-American All-Star analyst by *Institutional Investor* magazine in year t-1, zero otherwise.
WKDN	the signed difference between analyst j's first and last one-year-ahead EPS forecast for firm i in year t, divided by the absolute value of the actual. The first forecast is no earlier than year t-1's fiscal year-end and no later than 90 days prior to year t's earnings announcement, and the last forecast is no later than year t's earnings announcement.

Variables in tests of earnings forecast accuracy

FE_MA	the mean-adjusted earnings forecast error equal to the absolute forecast error, FE, scaled by the average absolute forecast error of one-year-ahead EPS forecasts issued by all analysts who follow firm i in year t. We only consider the latest one-year-ahead EPS forecast provided by analyst j for firm i in year t no later than 90 days prior to the firm's earnings announcement date and no earlier than the previous fiscal year-end. We exclude earnings forecasts where the corresponding actual earnings are released more than 90 days after the fiscal year-end.
NFT	the number of forecast types provided by analyst j for firm i during the past three months prior to the month when the earnings forecast is issued.
HOR	the number of days between analyst j's earnings forecast for firm i outstanding on June 30 in year t and the earnings announcement.

Variables in tests of price target forecast accuracy

PTE_MA	the mean-adjusted price target forecast error, defined as the absolute value of the difference between analyst j's one-year-ahead price target issued on day d and firm i's realized stock price on day d+360, divided by firm i's stock price on day d.
NFT	the number of forecast types provided by analyst j for firm i during the past three months prior to the price target forecast issuance month.
FREQ_PT	the number of analyst j's price target forecasts for firm i in year t.

Variables in tests of the profitability of stock recommendations

RET_REC	the market-adjusted buy-and-hold return of the stock recommendation issued by analyst j for firm i on day d of year t measured over at most six months. The buy-and-hold return is calculated as the return on long positions in stocks with a strong buy or buy recommendation (coded as 1 and 2 by I/B/E/S), and short positions in stocks with hold, sell, or strong sell recommendations (coded as 3, 4, and 5 by I/B/E/S). The return accumulation period runs from the day before the recommendation until the earlier of 180 days or the day before the recommendation is revised or reiterated. Market-adjusted returns are calculated by deducting the value-weighted average return of the market portfolio.
NFT	the number of forecast types provided by analyst j for firm i during the past three months prior to the recommendation issuance month.
FREQ_REC	the number of analyst j's stock recommendations for firm i in year t.

Variables in tests of stock market reactions

CAR3	the three-day market-adjusted abnormal return from day t-1 to t+1 centered on stock recommendation revision day t.
RECrev	the stock recommendation revision by analyst j for firm i on day d in year t. RECrev is defined as the difference between the I/B/E/S recommendation code for the newly issued recommendation and that for the previous recommendation for the same firm by the same analyst. For example, RECrev equals 1 if the analyst downgrades the stock by one level, e.g., from buy (I/B/E/S recommendation code = 2) to hold (I/B/E/S recommendation code = 3). Revisions with a distance between two consecutive recommendations greater than 365 calendar days are deleted.
NFT	the number of forecast types provided by analyst j for firm i during the past three months prior to the month of the day on which the recommendation is revised.
EPSrev	the one-year-ahead earnings forecast revision by analyst j for firm i on day t. Specifically, EPSrev equals the EPS forecast issued on day d minus the previously outstanding EPS forecast divided by the stock price two days prior to the forecast revision.
EBIrev	the one-year-ahead forecast revision of earnings before interest and taxes by analyst j for firm i on day t. Specifically, EBIrev equals the EBI forecast issued on day d minus the previously outstanding EBI forecast divided by the absolute value of the previously outstanding EBI forecast.
PRErev	the one-year-ahead forecast revision of pre-tax earnings (PRE) by analyst j for firm i on day t. Specifically, PRErev equals the PRE forecast issued on day d minus the previously outstanding PRE forecast divided by the absolute value of the previously outstanding PRE forecast.
SALrev	the one-year-ahead sales (SAL) forecast revision by analyst j for firm i on day t. Specifically, SALrev equals the SAL forecast issued on day t minus the previously outstanding SAL forecast divided by the absolute value of the previously outstanding SAL forecast.

Variables in tests of analyst career outcomes (continued on the following page)

TERMINATION	An indicator variable set equal to one if analyst j disappears from I/B/E/S in year t+1, zero otherwise.
PROMOTION	An indicator variable set equal to one if analyst j worked at a small brokerage house in year t but works at a large brokerage house in year t+1, zero otherwise. A brokerage house is categorized as a large (small) if its number of employees is above (below) the second (first) tercile.

DEMOTION	An indicator variable set equal to one if analyst j worked at a big brokerage house in year t but works at a small brokerage house in year t+1, zero otherwise.

Variables in tests of analyst career outcomes (continued)

NFT_M	the average rank of the number of forecast types issued by analyst j in year t. Specifically, we first use Eq. 1 to normalize the number of forecast types issued by analyst j for all firms she follows within each firm she follows in year t. We then take average of the normalized number of forecast types across all firms followed by analyst j in year t.
I_AFLTG	An indicator variable set equal to one if analyst j issued at least one long-term earnings growth forecast in year t, zero otherwise.
I_AFSAL	An indicator variable set equal to one if analyst j issued at least one one-year-ahead sales forecast in year t, zero otherwise.
BOLD_M	the absolute deviation of analyst j's first one-year-ahead EPS forecasts for firm i immediately after year t-1's fiscal year-end from the average of those issued by all other analysts, transformed following the average absolute deviation for firm i in year t.
COMP_M	the average number of analysts following the firms covered by analyst j in year t.

Variables in out-of-sample tests

REV_COMP12	the composite EPS revision for firm i at the end of month t. Specifically, REV_COMP12 = (m × REV1 + (12 − m) × REV2), where m is the number of months remaining before the end of FY1, and REV1 (REV2) is the revision in the FY1 (FY2) consensus EPS forecast, defined as the difference between the consensus at the end of month t and the consensus at the end of the previous month scaled by end-of-month stock price. The consensus is separately calculated for high-NFT and low-NFT EPS forecasts. The high-NFT (low-NFT) consensus FY1 and FY2 forecasts are the medians of the FY1 and FY2 forecasts classified as high-NFT (low-NFT) in the firm-month. A forecast is classified as high-NFT (low-NFT) if its associated NFT value is strictly above (below) the median NFT in a given firm-month, where NFT is the total number of forecast types provided by the analyst during the three-month period up to the end of the current month.
NFT_P	the rank-transformed NFT measure used in the calendar time portfolio analysis. Specifically, for the recommendation for firm i issued by analyst j on day t, $NFT\_P_{ijt} = (NFT_{ijt} - \min(NFT_{it}))/(\max(NFT_{it}) - \min(NFT_{it})$, where NFT_{ijt} is the number of forecast types issued for firm i by analyst j during

the three months prior to the recommendation announcement month, and $\min(NFT_{it})$ ($\max(NFT_{it})$) is the minimum (maximum) NFT across all analysts who have issued at least one recommendation for firm i during the six months prior to the recommendation announcement month.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11142-023-09791-8.

Acknowledgements The authors appreciate many helpful comments received from an anonymous reviewer, Scott Richardson (the editor), Dawn Matsumoto (the discussant), participants at the Review of Accounting Studies Conference, Stanford, December 2022, workshop participants at Bocconi University, BI Norwegian Business School, Pompeu Fabra University, Durham University, Jupiter Asset Management London, the Inquire UK Practitioner Seminar, George Foster, Yingjia Ning, Xiaoxi Wu, and Wanli Zhao. This paper was largely completed when Tong Wang was a Ph.D. student at Bocconi University and hence is unrelated to his current employer. Contact details: peter.pope@unibocconi.it (Pope); wang.tong@unibocconi.it (Wang).

Funding Open access funding provided by Università Commerciale Luigi Bocconi within the CRUI-CARE Agreement.

References

Asquith, Paul, Michael B. Mikhail, and Au.S. Andrea. 2005. Information content of equity analyst reports. *Journal of Financial Economics* 75 (2): 245–282.

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

Beyer, Anne, Daniel A. Cohen, Thomas Z. Lys, and Beverly R. Walther. 2010. The financial reporting environment: review of the recent literature. *Journal of Accounting and Economics* 50 (2–3): 296–343.

Bilinski, Pawel 2020. The Relevance of 'other surprises' in Explaining Earnings Announcement Returns. Available at SSRN 3539358

Bradshaw, Mark T. 2004. How do analysts use their earnings forecasts in generating stock recommendations? *The Accounting Review* 79 (1): 25–50.

Bradshaw, Mark T. 2011. Analysts' forecasts: what do we know after decades of work? Available at SSRN 1880339

Bradshaw, Mark T., Lawrence D. Brown, and Kelly Huang. 2013. Do sell-side analysts exhibit differential target price forecasting ability? *Review of Accounting Studies* 18 (4): 930–955.

Bratten, Brian, Cristi A. Gleason, Stephannie A. Larocque, and Lilian F. Mills. 2017. Forecasting taxes: new evidence from analysts. *The Accounting Review* 92 (3): 1–29.

Brav, Alon, and Reuven Lehavy. 2003. An empirical analysis of analysts' target prices: Short-term informativeness and long-term dynamics. *The Journal of Finance* 58 (5): 1933–1967.

Brown, Lawrence D. 2001. How important is past analyst forecast accuracy? *Financial Analysts Journal* 57 (6): 44–49.

Brown, Lawrence D., Andrew C. Call, Michael B. Clement, and Nathan Y. Sharp. 2015. Inside the "black box" of sell-side financial analysts. *Journal of Accounting Research* 53 (1): 1–47.
Call, Andrew C., Shuping Chen, and Yen H. Tong. 2009. Are analysts' earnings forecasts more accurate when accompanied by cash flow forecasts? *Review of Accounting Studies* 14 (2–3): 358–391.
Chen, Qi., and Wei Jiang. 2006. Analysts' weighting of private and public information. *The Review of Financial Studies* 19 (1): 319–355.
Choi, Jin Kyung, Rebecca N. Hann, Musa Subasi, and Yue Zheng. 2020. An empirical analysis of analysts' capital expenditure forecasts: evidence from corporate investment efficiency. *Contemporary Accounting Research* 37 (4): 2615–2648.
Christophe, Stephen E., Michael G. Ferri, and Jim Hsieh. 2010. Informed trading before analyst downgrades: evidence from short sellers. *Journal of Financial Economics* 95 (1): 85–106.
Clement, Michael B. 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* 27 (3): 285–303.
Clement, Michael B., Lisa Koonce, and Thomas J. Lopez. 2007. The roles of task-specific forecasting experience and innate ability in understanding analyst forecasting performance. *Journal of Accounting and Economics* 44 (3): 378–398.
Clement, Michael B., and Senyo Y. Tse. 2003. Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters? *The Accounting Review* 78 (1): 227–249.
Clement, Michael B., and Senyo Y. Tse. 2005. Financial analyst characteristics and herding behavior in forecasting. *The Journal of Finance* 60 (1): 307–341.
Cohen, Lauren, Andrea Frazzini, and Christopher Malloy. 2010. Sell-side school ties. *The Journal of Finance* 65 (4): 1409–1437.
Coleman, Braiden, Kenneth Merkley, and Joseph Pacelli. 2022. Human versus machine: a comparison of robo-analyst and traditional research analyst investment recommendations. *The Accounting Review* 97 (5): 221–244.
Cooper, Rick A., Theodore E. Day, and Craig M. Lewis. 2001. Following the leader: a study of individual analysts' earnings forecasts. *Journal of Financial Economics* 61 (3): 383–416.
Da, Zhi, Keejae P. Hong, and Sangwoo Lee. 2016. What drives target price forecasts and their investment value? *Journal of Business Finance & Accounting* 43 (3–4): 487–510.
Dechow, Patricia M., and Haifeng You. 2020. Understanding the determinants of analyst target price implied returns. *The Accounting Review* 95 (6): 125–149.
DeFond, Mark L., and Mingyi Hung. 2003. An empirical analysis of analysts' cash flow forecasts. *Journal of Accounting and Economics* 35 (1): 73–100.
Demirakos, Efthimios G., Norman C. Strong, and Martin Walker. 2010. Does valuation model choice affect target price accuracy? *European Accounting Review* 19 (1): 35–72.
Ertimur, Yonca, William J. Mayew, and Stephen R. Stubben. 2011. Analyst reputation and the issuance of disaggregated earnings forecasts to I/B/E/S. *Review of Accounting Studies* 16 (1): 29–58.
Ertimur, Yoncan, Jayanthi Sunder, and Shyam V. Sunder. 2007. Measure for measure: The relation between forecast accuracy and recommendation profitability of analysts. *Journal of Accounting Research* 45 (3): 567–606.
Fama, Eugene F., and Kenneth R. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116 (1): 1–22.
Givoly, Dan, Yifan Li, Ben Lourie, and Alexander Nekrasov. 2019. Key performance indicators as supplements to earnings: Incremental informativeness, demand factors, measurement issues, and properties of their forecasts. *Review of Accounting Studies* 24 (4): 1147–1183.
Gleason, Cristi A., Bruce W. Johnson, and Haidan Li. 2013. Valuation model use and the price target performance of sell-side equity analysts. *Contemporary Accounting Research* 30 (1): 80–115.
Gow, Ian D., Gaizka Ormazabal, and Daniel J. Taylor. 2010. Correcting for cross-sectional and time-series dependence in accounting research. *The Accounting Review* 85 (2): 483–512.
Groysberg, Boris, Paul M. Healy, and David A. Maber. 2011. What drives sell-side analyst compensation at high-status investment banks? *Journal of Accounting Research* 49 (4): 969–1000.
Hand, John R., Henry Laurion, Alastair Lawrence, and Nicholas Martin. 2021. Explaining firms' earnings announcement stock returns using FactSet and I/B/E/S data feeds. Review of Accounting Studies 1–32
Harford, Jarrad, Feng Jiang, Rong Wang, and Fei Xie. 2019. Analyst career concerns, effort allocation, and firms' information environment. *The Review of Financial Studies* 32 (6): 2179–2224.
Hashim, Noor A., and Norman C. Strong. 2018. Do analysts' cash flow forecasts improve their target price accuracy? *Contemporary Accounting Research* 35 (4): 1816–1842.

Hilary, Gilles, and Charles Hsu. 2013. Analyst Forecast Consistency. *The Journal of Finance* 68 (1): 271–297.

Hong, Harrison, Jeffrey D. Kubik, and Amit Solomon. 2000. Security analysts' career concerns and herding of earnings forecasts. The Rand Journal of Economics 121–144

Irvine, Paul, Marc Lipson, and Andy Puckett. 2007. Tipping. *The Review of Financial Studies* 20 (3): 741–768.

Jacob, John, Thomas Z. Lys, and Margaret A. Neale. 1999. Expertise in forecasting performance of security analysts. *Journal of Accounting and Economics* 28 (1): 51–82.

Juergens, Jennifer L., and Laura Lindsey. 2009. Getting out early: An analysis of market making activity at the recommending analyst's firm. *The Journal of Finance* 64 (5): 2327–2359.

Jung, Boochun, Philip B. Shane, and Yanhua Sunny Yang. 2012. Do financial analysts' long-term growth forecasts matter? Evidence from stock recommendations and career outcomes. *Journal of Accounting and Economics* 53 (1–2): 55–76.

Ke, Bin, and Yu. Yong. 2006. The effect of issuing biased earnings forecasts on analysts' access to management and survival. *Journal of Accounting Research* 44 (5): 965–999.

Keung, Edmund C. 2010. Do supplementary sales forecasts increase the credibility of financial analysts' earnings forecasts? *The Accounting Review* 85 (6): 2047–2074.

Loh, Roger K., and Mujtaba G. Mian. 2006. Do accurate earnings forecasts facilitate superior investment recommendations? *Journal of Financial Economics* 80 (2): 455–483.

Loh, Roger K., and Ren M. Stulz. 2011. When are analyst recommendation changes influential? *The Review of Financial Studies* 24 (2): 593–627.

Lundholm, Russel and Richard Sloan. 2004. Equity Valuation and Analysis

Mauler, Landon M. 2019. The Effect of analysts' disaggregated forecasts on investors and managers: evidence using pre-tax forecasts. *The Accounting Review* 94 (3): 279–302.

McFadden, Daniel 1973. Conditional logit analysis of qualitative choice behavior. Frontiers in Econometrics 105–142

Mikhail, Michael B., Beverly R. Walther, and Richard H. Willis. 1997. Do security analysts improve their performance with experience? *Journal of Accounting Research* 35: 131–157.

Mikhail, Michael B., Beverly R. Walther, and Richard H. Willis. 1999. Does forecast accuracy matter to security analysts? *The Accounting Review* 74 (2): 185–200.

Mikhail, Michael B., Beverly R. Walther, and Richard H. Willis. 2003. The effect of experience on security analyst underreaction. *Journal of Accounting and Economics* 35 (1): 101–116.

Richardson, Scott, Siew Hong Teoh, and Peter D. Wysocki. 2004. The walk-down to beatable analyst forecasts: The role of equity issuance and insider trading incentives. *Contemporary Accounting Research* 21 (4): 885–924.

Schipper, Katherine. 1991. Analysts' forecasts. *Accounting Horizons* 5 (4): 105.

Shroff, Pervin K., Ramgopal Venkataraman, and Baohua Xin. 2014. Timeliness of analysts' forecasts: The information content of delayed forecasts. *Contemporary Accounting Research* 31 (1): 202–229.

Stickel, Scott E. 1992. Reputation and performance among security analysts. *The Journal of Finance* 47 (5): 1811–1836.

Stickel, Scott E. 1995. The anatomy of the performance of buy and sell recommendations. *Financial Analysts Journal* 51 (5): 25–39.

Womack, Kent L. 1996. Do brokerage analysts' recommendations have investment value? *The Journal of Finance* 51 (1): 137–167.