The Value of Crowdsourced Earnings Forecasts

Russell Jame
University of Kentucky

Rick Johnston*
University of Alabama at Birmingham

Stanimir Markov
Southern Methodist University

Michael C. Wolfe
Virginia Tech

March 23, 2016

Abstract: Crowdsourcing — when a task normally performed by employees is outsourced to a large network of people via an open call — is making inroads into the investment research industry. We shed light on this new phenomenon by examining the value of crowdsourced earnings forecasts. Our sample includes 51,012 forecasts provided by Estimize, an open platform that solicits and reports forecasts from over 3,000 contributors. We find that Estimize forecasts are incrementally useful in forecasting earnings and measuring the market's expectations of earnings. Our results are stronger when the number of Estimize contributors is larger, consistent with the benefits of crowdsourcing increasing with the size of the crowd. Finally, Estimize consensus revisions generate significant two-day size-adjusted returns. The combined evidence suggests that crowdsourced forecasts are a useful supplementary source of information in capital markets.

Keywords: Analyst, Forecast, Earnings Response Coefficients, Crowdsourcing

JEL Classification: G28, G29, M41, M43

Acknowledgments: Accepted by Christian Leuz. Johnston thanks Leigh Drogen from Estimize for providing the Estimize data and answering questions about the business. This paper is a revision of an earlier working paper titled "Crowdsourcing Forecasts: Competition for Sell-Side Analysts?" by Johnston and Wolfe. Johnston conceived of and initiated this project while at Purdue University. Johnston thanks Constantin Cosereanu, V. Shah and B. Markovich from Bloomberg for answering related questions about Bloomberg. We also thank an anonymous referee, Shail Pandit, Phil Stocken, Stephen Taylor, Tony Kang and workshop participants at Cass Business School City University London, Hong Kong Poly University, London Business School Conference, Loyola Marymount University, McGill University, McMaster University, Rice University, Temple University Conference, UAB, University of Illinois at Chicago, University of Kansas, University of San Francisco, University of Technology Sydney Conference, Wake Forest University and York University for their helpful comments and suggestions. An Online Appendix to this paper can be downloaded at http://research.chicagobooth.edu/arc/journal-of-accounting-research/online-supplements.

*Corresponding author, rickj@uab.edu

Bolstered by the low cost of online publishing and the rising popularity of blogs, discussion forums and commenting, a growing number of niche web sites are creating opportunities for new forms of investment analysis to emerge — and for buy-side professionals, even those at rival firms, to collaborate and learn directly from one another. These social media web sites are supplementing, and in some cases supplanting, the traditional Wall Street information ecosystem that transmits sell-side investment research and stock calls to the buy side.

Costa (2010) Institutional Investor Magazine

1. Introduction

In the last two decades, technology has significantly lowered information and communication costs and bolstered the creation of new information sources (e.g., blogs, message boards, Facebook, and Twitter), thereby changing the process by which investors acquire information. According to a recent survey, nearly one in three individuals in the US relies on investment advice transmitted via social media outlets.[1] Recognizing the increased importance of this new source of information in the capital markets, the Securities and Exchange Commission (SEC) now allows firms to disclose news through social media.

Technology is also altering interactions between organizations and outsiders in other ways. Increasingly, businesses use technology to capture the collective intelligence of online participants. This blend of bottom-up, open, creative process to meet organizational goals is called crowdsourcing (Brabham, 2013). Various entities, such as Seeking Alpha and Estimize seek to supplement or disrupt sell-side research with crowdsourcing. Seeking Alpha crowdsources investment research and publishes it on its website. Estimize seeks to create an alternative to the sell-side earnings consensus by crowdsourcing forecasts from analysts, investors, corporate finance professionals, students and others. Prior to Estimize, whisper forecasts were an alternative source of earnings forecasts. Whisper forecasts emerged in the 1990s, as concerns with sell-side bias and strategic non-updating in the period prior to earnings

[1] http://www.experiencetheblog.com/2013/04/four-recent-studies-on-rapid-adoption.html

announcements increased. Subsequently websites dedicated to publishing whisper forecasts were established.

This paper offers a first look at the value of crowdsourced earnings forecasts from Estimize. These forecasts warrant research attention because they have unique attributes relative to other sources of alternative earnings information (e.g., Whisper sites and Seeking Alpha).[2] Specifically, a whisper site distributes a single forecast that aggregates information from various sources using a proprietary approach. Thus the role of crowdsourcing is both limited and unidentified, and prior evidence on the value of whisper forecasts may not extrapolate to the crowdsourced forecast setting.[3] Social media finance sites (e.g., Motley Fool, StockTwits, and Seeking Alpha) have crowdsourcing features, but offer unstructured data (i.e., commentaries), limiting their usefulness as a source of earnings information. Therefore, a crowdsourcing site able to attract and retain a large number of capable earnings forecasters may become integral to the sourcing and dissemination of earnings forecasts.

We assess the value of Estimize forecasts by investigating whether they are incrementally useful in forecasting earnings and measuring the market expectation of earnings, and whether they convey new information. Our analyses are guided by two non-mutually exclusive hypotheses. The first hypothesis is that crowdsourced forecasts are incrementally useful only because they are less biased and incorporate more public information. The second hypothesis implies a greater role for crowdsourced forecasts in capital markets: by capturing the collective wisdom of a large and diverse group of individuals, they impart new information to the markets.

[2] Section 2 offers a more in-depth comparison of Estimize to other sources of crowdsourced research as well as whisper forecasts.

[3] Prior evidence on whether whisper numbers convey information to the market is mixed. Analyzing a sample of 262 forecasts, Bagnoli et al. (1999) find affirmative evidence, but their findings haven't been replicated in more recent and larger samples (Bhattacharya et al., 2006; Brown and Fernando, 2011).

Our sample consists of 51,012 quarterly earnings forecasts for 1,874 firms submitted to Estimize by 3,255 individuals in 2012 and 2013. Firms covered by Estimize contributors are generally in the IBES universe but are larger, more growth oriented, and more heavily traded than the average IBES firm. Relative to IBES forecasts, individual Estimize forecasts tend to be less accurate at long horizons, but equally accurate at shorter horizons; they are less biased and bolder (further from the combined IBES and Estimize consensus). Approximately half of Estimize forecasts are issued in the two days prior to the earnings announcement date, while less than 2% of IBES forecasts are issued in the same period. The stark difference in forecast timing suggests a complementary relation between IBES analysts and Estimize contributors.

First, we explore whether Estimize forecasts are incrementally useful in predicting earnings by quantifying the accuracy benefits from combining Estimize forecasts with the IBES consensus or a statistical forecast based on firm characteristics (So, 2013). Using either benchmark, we find that incorporating Estimize forecasts yields significant improvements in accuracy over all forecast horizons during the quarter. To explore whether the incremental usefulness of Estimize forecasts is robust to controlling for differences in timing and bias, we estimate a regression of actual earnings per share (EPS) on contemporaneous Estimize and IBES consensus forecasts.[4] The coefficient on the Estimize consensus is significantly greater than zero, indicating that Estimize has incremental information. More importantly, this coefficient is increasing in the number of Estimize contributors suggesting that this incremental information increases with the size of the crowd.

Next, we assess whether Estimize forecasts add value as a measure of the market's earnings expectation based on a regression of three-day size-adjusted earnings announcement

[4] By focusing on the slope coefficient from a regression, we abstract from differences in usefulness that stem from differences in forecast bias.

returns on the IBES and Estimize consensus earnings surprise. We find that Estimize is incrementally useful in measuring the market's expectations, and the relative importance of Estimize as a measure of the market's expectations is increasing in the size of the contributor base. When the number of Estimize contributors is greater than five, the Estimize consensus fully subsumes the IBES consensus.

Finally, we estimate two-day size-adjusted returns following Estimize consensus forecast revisions to address the question of whether Estimize forecasts convey new information to the market. After filtering out revisions that occur around confounding news events, we document abnormal returns of 0.26% following large upward revisions (the top half of upward revisions) and -0.15% following large downward revisions. The difference of 0.41% is statistically significant, and it does not appear to reverse over the subsequent two weeks, suggesting that new information, rather than investor overreaction or price pressure, explains the return differential.

Our primary contribution is to introduce a new phenomenon, crowdsourced earnings forecasts, and explore its significance. Our findings that Estimize forecasts provide incremental information for forecasting earnings and for measuring the market expectation of earnings provide support for crowdsourced forecasts as a supplemental source of information. However, these results are partially attributable to compensating for sell-side forecast deficiencies. The incremental usefulness of Estimize forecasts in predicting earnings, after controlling for differences in forecast bias and horizon with IBES and the evidence of significant price reactions to Estimize revisions corroborate our second hypothesis that they convey new information. Finally, our evidence that the incremental usefulness of Estimize in forecasting earnings and proxying for the market expectation is increasing in the number of contributors illustrates that the value of crowdsourcing is a function of crowd size.

Our study also contributes to the literature that explores different approaches to forecasting earnings (Brown et al., 1987; Bradshaw et al., 2012; So, 2013). Specifically, it compares crowdsourced forecasts to sell-side and statistical forecasts. Crowdsourced forecasts are available for fewer stocks and generally at much shorter horizons than sell-side forecasts, but they are less biased, bolder, and incrementally useful in predicting future earnings. Statistical forecasts suffer from a significant timing disadvantage, but they are available for all stocks. They also have incremental predictive power relative to sell-side forecasts (So, 2013) but not relative to crowdsourced forecasts concentrated in the period before earnings are announced (this study). Finally, sell-side forecasts are available throughout the forecast period and incrementally useful in forecasting earnings at all horizons.

This paper fits in a broader literature that explores how technological and institutional changes influence the sourcing and dissemination of financial information in today's capital markets.[5] Surveying this literature, Miller and Skinner (2015) observe that social media provide firms with new ways to disseminate information but also reduce firms' ability to tightly manage their information environments, since external users have the ability to create and disseminate their own content (p. 13). Our paper provides evidence that technology has empowered external users to create and disseminate useful information.

2. Background and Hypotheses

2.1. Crowdsourcing

"Crowdsourcing" was first coined by Jeff Howe of *Wired Magazine* in 2006. It is the act of a company or institution taking a function once performed by employees and outsourcing it to

[5] E.g., Crawford, Gray, Johnson, and Price, 2014; Blankenspoor, Miller, and White, 2014; Giannini, Irvine, Shu, 2014; Jung, Naughton, Tahoun, and Wang, 2014; Lee, Hutton, and Shu, 2015.

an undefined, generally large network of people in the form of an open call.[6] The key ingredients of crowdsourcing are an organization that has a task it needs performed, a community that is willing to perform the task, an online environment that allows the work to take place and the community to interact with the organization, and mutual benefit for the organization and the community (Brabham, 2013).

A well-known example of successful crowdsourcing is Wikipedia: a web-based, encyclopedia project, initiated in 2001 by the Wikimedia foundation, where content is freely contributed and edited by a large number of volunteers rather than by a small number of professional editors and contributors. Wikipedia is among the top ten most visited web sites.[7] It not only covers more topics than *Encyclopedia Britannica*, it is also surprisingly accurate. According to a 2005 study by the scientific journal *Nature* comparing 42 science articles by Wikipedia and *Encyclopedia Britannica*, the average Wikipedia science article has about four inaccuracies while the average *Encyclopedia Britannica* article has about three (Giles, 2005).

2.2. Estimize

2.2.1. Institutional details

Estimize is a private company founded in 2011 by Leigh Drogen, a former quantitative hedge fund analyst, with the objective of crowdsourcing earnings and revenue forecasts and thus providing an alternative to sell-side forecasts. Estimize contributors include independent, buy-side, and sell-side analysts, as well as private investors and students. Contributors are asked but not required to provide a brief personal profile. Forecasts are available on the Estimize web site and Bloomberg; they are also sold as a data feed to institutional investors. The availability of

[6] Crowdfunding is a related concept in which firm financing is solicited from a large network of people via the internet.
[7] http://en.wikipedia.org/wiki/List_of_most_popular_websites

Estimize data on Bloomberg, the most widely used (by professionals) financial information system, is evidence of the market's interest in crowdsourced financial information. Bloomberg representatives reveal that Bloomberg makes Estimize data available without an upcharge, but that it does not monitor its use. Other social media data available on Bloomberg terminals include StockTwits and Twitter.

Estimize takes steps to incentivize accuracy and ensure the integrity of its data. By asking contributors to provide a personal profile as well as tracking and reporting contributor accuracy, Estimize encourages accurate forecasting and also allows investors to form their own assessment of contributor accuracy. Further, all forecasts are limited to a certain range based on a proprietary algorithm. Estimates by new analysts are manually reviewed. Forecasts whose reliability is believed to be low are flagged and excluded from their reported consensus. Finally, to encourage participation and accurate forecasting, Estimize recognizes top contributors with prizes and features them in podcasts.

Motivations for contributing estimates to Estimize are numerous and varied. For instance, some portfolio managers and retail investors may contribute forecasts because they want to ensure that prices more quickly reflect their information – a practice known among practitioners as "talking your book" (Crawford et al., 2014); others because they want to manipulate prices.[8] Students and industry professionals may participate because they want to develop their

[8] Analyzing a sample of 142 stock market manipulation cases pursued by the SEC from January 1990 to November 2001, Aggarwal and Wu (2006) report that approximately 83% concern stocks traded in relatively inefficient markets (OTC Bulletin Board, Pink Sheets, regional exchanges, or unidentified markets) which Estimize contributors shy away from. Among these cases is the highly publicized case of 14-year-old Jonathan Lebed who successfully manipulated the price of 11 thinly-traded micro-cap stocks by posting messages on Yahoo Finance message boards (http://www.sec.gov/news/press/2000-135.txt).

forecasting skills. Finally, all individuals may derive utility from sharing information, competing against the experts, and potentially being recognized as accurate forecasters.[9]

Since crowdsourced research is a new phenomenon that has received limited attention in the academic literature, we next discuss similarities and differences between Estimize and select information sources with crowdsourcing features: whisper sites, Seeking Alpha, SumZero, StockTwits, and Motley Fool.

2.2.2. Comparison to Other Sources of Crowdsourced Research

Whisper sites share Estimize's general objective to create an alternative source of earnings estimates, but we view these sites as a predecessor rather than a variant of crowdsourcing. Specifically, while Estimize outsources the task of providing earnings forecasts to a community of contributors, whisper sites gather information by various means and then distill it into a whisper forecast (Brown and Fernando, 2011). Thus, generating an earnings forecast is performed by the whisper site, not the contributors. Further complicating any comparison is the fact that each site's process is unique and proprietary, thus opaque (Bhattacharya et al., 2006).[10]

The evidence on whether whisper forecasts convey new information to the market is limited and mixed. The only study that finds such evidence analyzes a hand-collected sample of 262 forecasts gathered from the World Wide Web, *The Wall Street Journal*, and financial

---

[9] Surveying the crowdsourcing literature, Estelles-Arolas and Gonzalez-Ladron-DeGuevara (2012) conclude that individuals contribute to "satisfy one or more of the individual needs mentioned in Maslow's pyramid: economic reward, social recognition, self-esteem, or to develop individual skills" (p. 7).

[10] In a December 6, 2011 blog post, Leigh Drogen identifies dissatisfaction with the whisper number's opaqueness as an impetus for founding Estimize. "No longer will the whisper number be a secret back stage Wall Street product, we're throwing it in the open where everyone can see it. We're going to provide transparency to the process, and measurement of those who contribute to that whisper number. We're going to connect the buy side with independent analysts, traders, and the social finance community in order to find out what the market truly expects these companies to report."

newswires over the period 1995-1997 (Bagnoli et al., 1999). The small, heterogeneous, and pre-Regulation FD sample raises questions about the generalizability and current relevance of the evidence. In fact, Rees and Adut (2005) find that whisper forecasts are generally more accurate than analysts' forecasts prior to Reg. FD but less accurate after Reg. FD. Similarly, Bhattacharya et al. (2006) analyze the post-Reg. FD period and find that whisper forecasts are not more informative than analysts' forecasts and do not contain any incrementally useful information.

While whisper sites use a different approach to offer a similar product to Estimize, Seeking Alpha uses a similar approach to offer a different product. Seeking Alpha provides an open platform for investment research (rather than earnings estimates) contributed by investors and industry experts. Efforts to promote valuable research include vetting the quality of research commentaries, paying contributors based on the number of page views their commentaries receive, and recognizing most-read contributors as "Opinion Leaders" on the site. Chen et al. (2014) find robust evidence that the tone of commentaries posted on Seeking Alpha predicts stock returns, consistent with crowdsourced research having investment value and Seeking Alpha being a distinct source of new information.

SumZero is similar to Seeking Alpha, but its distinguishing feature is that it aims to crowdsource buy-side research for the benefit of the buy-side. Contributors and users must verify buy-side employment, which makes SumZero considerably less open than Seeking Alpha or Estimize. Crawford et al. (2014) find that recommendations posted on SumZero have investment value, consistent with buy-siders having the capacity to produce new information and validating SumZero as a separate source of new information.

An increasingly popular information source is StockTwits, an open platform that allows individuals to post 140 character messages about stocks. StockTwits differs from the sites discussed above in that it crowdsources two distinct tasks: the task of searching and reporting for market-moving news (typically conducted by editors and reporters employed by financial newswires) and the task of providing research (typically conducted by Wall Street analysts). Early evidence shows that, on average, StockTwits' contributors have negative stock picking skill, suggesting that their messages reflect investor sentiment unrelated to firm fundamentals (Giannini et al., 2014).

Founded in 1993 at the dawn of the internet era as an investment newsletter, The Motley Fool has become a multimedia financial services company, offering investment advice and financial news and products, as well as a platform for subscribers to contribute their own stock picks. Avery et al. (2011) and Hirschey et al. (2000) find that Motley Fool's crowdsourced stock picks and the site's own stock picks, respectively, have investment value, but neither study explores whether these recommendations add value to an investor who is aware of sell-side research and the post-earnings announcement drift anomaly (Chen et al., 2014).[11]

In sum, technological change has spurred the development of new sources of investment research. As a source of earnings estimates, Estimize offers unique advantages. Compared to whisper sites, Estimize is more transparent and open, thus potentially reflecting a more diverse set of contributors. Users of social finance sites (e.g., Seeking Alpha) have access to stock opinions and commentaries from a diverse set of contributors, but these opinions and commentaries must be further processed to generate a quantitative earnings forecast. By examining the significance of crowdsourced earnings forecasts, our study contributes to the

[11] An earlier literature examines opinions posted on internet message boards and chatrooms and finds little or no evidence that these opinions are value-relevant (Wysocki, 1998; Tumarkin and Whitelaw, 2001; Antweiler and Frank, 2004; Das and Chen, 2007).

understanding of the process by which earnings forecasts are sourced and disseminated in capital markets.

2.3. Hypotheses

The demand for crowdsourced earnings forecasts is likely driven by (1) the known shortcomings of sell-side forecasts, such as bias, inefficiency, and tendency not to update immediately before earnings announcements,[12] (2) the apparent failure of the whisper sites to become a pervasive source for earnings forecasts,[13] and (3) the belief that the forecasts of a larger, more independent, and more diverse collection of people can bring new information to the market.[14]

Our empirical analyses of forecasts provided by Estimize, the first genuine supplier of crowdsourced forecasts, are guided by two broad hypotheses. The first hypothesis is that crowdsourced forecasts only compensate for sell-side forecasts' bias and reluctance to update in the period immediately prior to earnings announcements. Under this hypothesis, crowdsourced forecasts may provide incremental earnings information over and above the sell side simply by incorporating more public information and being less biased.

---

[12] See Sections 3.4 and 3.5 in Ramnath et al. (2008) for a survey of studies documenting analyst forecast inefficiency and bias, respectively. Bhattacharya et al. (2006, p. 16) identifies the sell-side's reluctance to update earnings forecasts as a contributing factor to the whisper forecast phenomenon. Bagnoli et al. (1999) document that sell-side (whisper) forecasts are relatively more frequent earlier (later) in the quarter. Berger et al. (2016) conclude that the relative absence of sell-side forecasts late in the quarter is explained by analysts strategically disseminating earnings information without adjusting the earnings forecast and frictions limiting the frequency of revisions to the current quarter forecasts.

[13] Bhattacharya et al. (2006) discuss why whisper forecasts are popular with individual investors but not with institutional investors and present results which "suggest that institutional investors do not pay much attention to whisper numbers." (p. 17)

[14] In an interview with Business Insider, Leigh Drogen, founder of Estimize says: "The other part of it is, and this may be even more important than the fact that we believe that for many stocks the Estimize community will be more accurate, but they'll be more representative of the market. That's the most important part, it's that the sell side is a very narrow set of people whose incentive structure is geared toward producing data in a very specific way. We believe if we open it up to all the different people out in the financial sphere including hedge fund analysts, independent analysts, regular traders, regular investors, people in corporate finance… Having all of those disparate groups contribute to one estimate will get a more representative view of what the market believes."

The second and more consequential hypothesis asserts that crowdsourced forecasts convey new information to the market. Our hypotheses are not mutually exclusive. Crowdsourced forecasts may correct sell-side deficiencies and increase the amount of information. One cannot presume that crowdsourced forecasts have information content for two reasons. First, there is mixed prior evidence on whether research with crowdsourcing features conveys new information. For instance, opinions posted on Seeking Alpha convey new information (Chen et al., 2014), but those posted on StockTwits do not (Giannini et al., 2014). Also, Bagnoli et al.'s (1999) results that whisper forecasts convey new information have not been replicated by later studies (Bhattacharya et al., 2006; Brown and Fernando, 2011). Second, our ability to draw inferences about crowdsourced forecasts on the basis of prior evidence is limited given the substantial differences between Estimize and the sources of crowdsourced research and whisper forecasts examined in prior work.

3. Data and Descriptive Statistics

3.1. Sample

We outline the sample selection in Table 1. The initial Estimize sample includes 51,012 non-GAAP earnings per share forecasts where both the estimate and the earnings announcement dates occur in the 2012 or 2013 calendar year. The sample includes 1,874 unique firms, 7,534 firm-quarters, and 3,255 Estimize contributors. We exclude forecasts issued more than 90 days prior to the earnings announcement — a rarity for Estimize — and forecasts issued after earnings are announced (likely data errors). We eliminate forecasts "flagged" by Estimize as less reliable (see Section 2.2.1.).[15] Finally, when a contributor made multiple forecasts on a single day, we replace those forecasts with the contributor's average for that day.[16] The final Estimize sample includes 45,569 forecasts for 1,870 firms contributed by 3,054 individuals.

An important objective of our study is to conduct a comparative analysis of crowdsourced forecasts, provided by Estimize, and sell-side forecasts, provided by IBES. We therefore create an Estimize–IBES matched sample by requiring that (1) an Estimize firm-quarter include at least one IBES earnings per share forecast and (2) Estimize and IBES report actual EPS that match to two decimal places. The second filter is needed to conduct proper accuracy comparison and imposed only when needed.[17] The final Estimize–IBES matched sample includes 2,835 contributors providing 37,031 forecasts for 1,601 firms.

3.2. Characteristics of Firms Covered by Estimize and IBES

Panel A of Table 2 contrasts the characteristics of firms covered by (1) both Estimize and IBES, (2) IBES only, and (3) Estimize only.[18] The number of firm-quarters in the three categories are 6,580, 18,041, and 750, respectively, revealing a considerable gap in breadth of coverage between Estimize and IBES. There is also a gap, although a smaller one, in depth of

---

[15] Data quality is a valid concern given that Estimize is an open platform that includes non-professionals. In the Internet Appendix, we repeat our main tests after 1) including flagged forecasts and 2) including flagged forecasts but excluding estimates more than three standard deviations away from the mean of all existing Estimize and IBES forecasts. Our results suggest that excluding Estimize-flagged observations or statistical outliers enhances the value of crowdsourced forecasts.

[16] An alternative approach would be to use the last forecast, in effect assuming the last forecast is a sufficient forecast for a contributor's information set. However, in many cases the time stamps for the two forecasts are identical. When the time stamps differ, using the last forecast yields similar results.

[17] Since Estimize reports only historical (unadjusted for splits) data, we use historical IBES data throughout the study. Estimize obtains actuals from Briefing.com, whereas IBES evaluates company-reported actuals "to determine if any Extraordinary or Non-Extraordinary Items (charges or gains) have been recorded by the company during the period… If one or more items have been recorded during the period, actuals will be entered based upon the estimates majority basis at the time of reporting." (See Methodology for Estimates: A Guide to Understanding Thompson Reuters Methodologies, Terms and Policies for the First Call and I/B/E/S Estimates Databases (October 2009) available on www.wharton.upenn.edu/wrds/.) Because there is no generally accepted definition of operating earnings, IBES-reported actual EPS may differ from Estimize-reported actual EPS.

[18] The sample analyzed in Table 2 is larger than the Final IBES-Matched Sample because we drop the requirement that IBES and Estimize report identical non-GAAP EPS actuals.

coverage. Specifically, conditional on the two groups of forecasters covering the same firm, the average number of Estimize (IBES) forecasters in the same firm-quarter is 6.07 (10.45). The smaller number of Estimize contributors, relative to IBES analysts, likely reflects the fact that Estimize is still a relatively young venture. The small number of firm-quarters with Estimize-only coverage, 750, suggests that for all practical purposes, firms covered by Estimize contributors are a subset of the firms covered by IBES analysts. Additionally, we observe systematic and statistically significant differences in the characteristics of firms covered by both Estimize and IBES and those covered only by IBES. In particular, the former are larger, less volatile but more growth-oriented, and more liquid.

Panels B and C focus on firm-quarters with both Estimize and IBES coverage. In Panel B, we sort observations into quartiles based on depth of Estimize coverage (number of contributors in a firm-quarter). We document significant differences in depth of coverage across firms. For instance, only observations in the top quartile have coverage higher than the cross-sectional mean of 6.07; all observations in the bottom quartile have coverage of one. Further, we observe a strong, monotonic relation between Estimize coverage and IBES coverage, the latter ranging from 8.54 (bottom quartile) to 13.87 (top quartile), suggesting common factors drive Estimize and sell-side coverage decisions. A similar monotonic relation exists between depth of Estimize coverage and a firm's size, growth, and turnover, consistent with the notion that large, growth-oriented, and liquid firms attract more Estimize coverage. After sorting observations into quartiles based on depth of IBES coverage, we find that the same firm characteristics, plus low volatility, appear attractive to IBES analysts (Panel C).[19]

[19] In the Internet Appendix, we confirm that the univariate patterns documented in Panel C hold in a regression setting.

3.3. Comparison of Estimize and IBES Forecasts

Panels A and B of Table 3 examine Estimize contributor and IBES analyst activities during the quarter. The sample is the Estimize-IBES matched sample. Most Estimize contributors issue one forecast per quarter for each firm they cover. Estimize forecasts concentrate in the period immediately prior to earnings announcements, as evidenced by mean (median) forecast horizon of five days (two days). Finally, we observe that the mean (median) number of firms covered is 8.41 (1), suggesting that most Estimize contributors cover a single company.

IBES analysts are slightly more active. Specifically, the average IBES analyst issues 1.37 forecasts in a firm-quarter. IBES analysts issue their forecasts considerably earlier, as evidenced by mean (median) forecast age of 59 (65) days. The average (median) IBES analyst covers 3.92 (3) firms in the Estimize–IBES sample.

To further explore the difference in forecast horizon, Figure 1 plots the fraction of total Estimize and total IBES forecasts with horizon longer than or equal to *t*, where *t* ranges from 90 to zero. We find that 7% of the Estimize forecasts have horizons longer than 30 days, and 30% of Estimize forecasts have horizons longer than 5 days. In contrast, the corresponding figures for IBES are 70% and 95%. The stark difference in forecast horizons across the Estimize and IBES samples suggests that Estimize and IBES complement each other as sources of information in the short-term and long-term, respectively. In particular, IBES forecasts are more timely while Estimize forecasts are likely to reflect more recent information (Cooper, Day, and Lewis, 2001).[20]

[20] See Guttman (2010) and Shroff et al. (2014) for analyses of the trade-off between timeliness and accuracy.

Next, we compare individual Estimize and IBES forecasts in terms of accuracy, bias, and boldness. Our goal in this section is only to offer stylized facts about a new source (Estimize) of earnings forecasts, rather than to test formal hypotheses about differences in forecast quality between Estimize and IBES.

Following Clement (1999), we define forecast accuracy as the proportional mean absolute forecast error (PMAFE) measured as:

$$PMAFE_{i,j,t} = \left(AFE_{i,j,t} - \overline{AFE}_{j,t} \right) \Big/ \overline{AFE}_{j,t} , \tag{1}$$

where $AFE_{i,j,t}$ is the absolute forecast error for analyst *i*'s forecast of firm *j* for quarter *t* earnings, and $\overline{AFE}_{j,t}$ is the mean absolute forecast error for firm *j* in quarter *t*. Note that *PMAFE* is a measure of inaccuracy; therefore, large values indicate lower accuracy. Since *PMAFE* is a relative measure of accuracy, we only include firm-quarters with more than five unique (Estimize or IBES) forecasters (eliminating 646 Estimize forecasts and 453 firm-quarters). Given the significant difference in forecast horizon between Estimize and IBES, we partition observations into five groups based on forecast horizon. Further, we require that each group includes only firm-quarters with at least one Estimize and one IBES forecast. In the case of multiple Estimize (or IBES) forecasts, we compute an accuracy measure for each forecast and average individual accuracy measures to produce a single accuracy measure. In sum, for each firm-quarter in a given forecast horizon group, we calculate one Estimize accuracy measure and one IBES measure. Accuracy measures for forecasts in different horizon groups are standardized the same way, which makes it possible to document and interpret accuracy improvement over time.

Panel A of Table 4 reports average *PMAFE* for Estimize and IBES, their difference, and the corresponding t-statistic.[21] When the forecast horizon ranges from 90 to 30 days, the Estimize *PMAFE* is significantly larger than the IBES *PMAFE* (0.21 vs. 0.11), consistent with Estimize contributors being less accurate. At shorter horizons there is no significant difference in the accuracy of Estimize and IBES forecasts.

We measure forecast bias as:

$$BIAS_{i,j,t} = \frac{Forecast_{i,j,t} - Actual_{j,t}}{Price_{j,t-1}} * 100. \tag{2}$$

Panel B of Table 4 reports average forecast bias for Estimize and IBES, their difference, and the corresponding t-statistics. We find that both Estimize and IBES forecasts are relatively pessimistic (i.e., forecasts tend to be lower than actuals).[22] However, IBES forecasts exhibit greater pessimism, consistent with sell-side analysts' incentives to issue easy-to-beat forecasts (Richardson et al., 2004).[23]

Boldness, typically defined as the extent to which a forecast deviates (in absolute value) from the current consensus, is a key forecast attribute in theories of reputation and herding. Following Hong, Kubik, and Solomon (2000), we measure boldness as

$$Boldness_{i,j,t} = \left| Forecast_{i,j,t} - \overline{Forecast}_{j,t} \right| \Big/ \overline{Forecast}_{j,t}, \tag{3}$$

where $Forecast_{i,j,t}$ is analyst *i*'s forecast of firm *j* for quarter *t* earnings, and $\overline{Forecast}_{j,t}$ is the consensus forecast for firm *j* in quarter *t*, which we compute by averaging across all IBES and

[21] Throughout the paper, t-statistics are computed based on standard errors clustered by firm. Results are very similar if standard errors are double-clustered by both firm and quarter.
[22] Much of the analyst literature subtracts the forecast from the actual, resulting in positive pessimism measures.
[23] This finding appears at odds with prior work that finds that sell-side analysts are often optimistic, particularly at longer horizons (Richardson, Teoh, and Wysocki, 2004). Much of the difference stems from time-series variation in forecast bias. In particular, over the period 1984-2001 (the period studied in Richardson et al., 2004), we find that the average bias for forecasts of horizons of greater than 30 days is 0.24 (optimism), compared to -0.07 over the period 2002-2014 (pessimism). These results are provided in the Internet Appendix.

Estimize forecasts available at the time of the forecast. We drop the first forecast for each firm-quarter because we are not able to estimate a prior consensus. If an analyst has issued multiple forecasts in the same firm-quarter, we include her most recent forecast.

We find that Estimize forecasts are generally bolder than IBES forecasts (Panel C), consistent with the view that Estimize contributors have more diverse information sets and stronger forecasting incentives than the sell-side. While only descriptive, our findings that Estimize forecasts are reasonably accurate, less biased, and generally bolder than IBES forecasts provide preliminary evidence that Estimize forecasts could be a useful supplementary source of information.[24]

4. The Value of Estimize Forecasts

We investigate whether Estimize forecasts are useful in predicting earnings, measuring the market's expectation, and facilitating price discovery.

4.1. Predicting Earnings

We first examine whether a consensus forecast that combines Estimize and IBES forecasts is more accurate than an IBES-only consensus (Section 4.1.1). The IBES consensus is a natural benchmark as Estimize aims to provide "both a more accurate and more representative view of expectations compared to sell side only data sets which suffer from several severe biases."[25] Statistical forecasts have been found to be both superior (Bradshaw et al., 2012) and incrementally useful (So, 2013) to sell-side analysts in forecasting earnings at longer horizons, prompting us to also benchmark Estimize forecasts against two statistical forecasts: a de-biased

[24] In the Internet Appendix, we examine whether differences in accuracy, bias, and boldness between Estimize and IBES forecasts are related to firm characteristics (size, book-to-market, volatility, and turnover) and the number of IBES and Estimize contributors.
[25] https://www.estimize.com/about

IBES forecast and a statistical forecast computed from firm characteristics (So, 2013) (Section 4.1.2). Finally, we examine factors contributing to the incremental usefulness of Estimize forecasts (Sections 4.1.3 and 4.1.4).

4.1.1 Combining Estimize and IBES forecasts

We first test whether a consensus forecast that combines Estimize and IBES forecasts is more accurate than an IBES-only consensus. Consistent with prior literature, we construct an *Estimize*, *IBES*, and *Combined Consensus* forecast with a *t*-day horizon by averaging corresponding individual forecasts with horizons longer than or equal to *t* days. If a forecaster has issued multiple forecasts within the horizon, we include only the most recent one. We measure the accuracy of a consensus forecast (*PMAFE*) for firm *j* in quarter *q* as the difference between the consensus absolute error and the mean absolute forecast error (*MAFE*) across all forecasts for firm *j* in quarter *q*, scaled by the mean absolute forecast error (*MAFE*).

Table 5 presents the results for horizons that range from 60 to zero days.[26] We find that at the 60-day horizon, the *Estimize Consensus* is significantly less accurate than the *IBES Consensus* (*PMAFE* of 0.28 vs. -0.07), consistent with Panel A, Table 4's findings that individual Estimize forecasts are less accurate than individual IBES forecasts at longer horizons. However, accuracy is significantly improved by combining Estimize and IBES forecasts even at this horizon. Specifically, the difference between the *Combined Consensus* and the *IBES Consensus* is -0.03, and the *Combined Consensus* is more accurate than the *IBES Consensus* approximately 57% of the time.

As the forecast horizon decreases, the benefits from combining Estimize and IBES forecasts increase. For example, when the forecast horizon is 30 (1) days, the *Combined*

[26] We note that the corresponding increase in number of observations from 430 to 5,002 is due to the scarcity of long-term Estimize forecasts.

Consensus is more accurate than the *IBES Consensus* 60% (64%) of the time. The documented pattern is not surprising in view of our Figure 1 evidence that Estimize forecasts are infrequent at long horizons and common at short horizons. In untabulated analysis, we find that the average number of forecasts included in the *Estimize Consensus* increases from 1.83 when the horizon is 60 days to 5.86 when the horizon is one day. Our results are consistent with the accuracy of a consensus generally increasing with the number of forecasts.[27]

4.1.2. Combining Estimize and Statistical Forecasts

Given the well documented bias in sell-side forecasts, one way to improve upon them may be to simply remove the bias. We compute the de-biased IBES forecast ($IBES^D$) of analyst i for firm j in quarter t as:

$$IBES^{D}{}_{i,j,t} = \alpha_t + \beta_t * IBES_{i,j,t}\,, \tag{4}$$

where α_t and β_t are the estimated intercept and slope coefficient from a cross-sectional regression of actual quarterly earnings on IBES forecasted earnings across all four quarters in year *t-1*. The cross-sectional regression is estimated on a sample of firms with at least one Estimize forecast in quarter t. Each year the intercept is 0.02 and the slope coefficient is 1.02, meaning each IBES forecasts must be increased by adding a constant, 0.02, and scaled up by a factor of 0.02.

After de-biasing IBES forecasts, we repeat the analysis conducted in Table 5. The results, reported in Panel A of Table 6, show that the *Combined Consensus* continues to be significantly more accurate than the $IBES^D$ *Consensus*. For example, at the 30-day (1-day) horizon, the *Combined Consensus* is more accurate than the $IBES^D$ *Consensus* 56% (59%) of the time. These estimates are lower than the corresponding estimates of 60% (64%) reported in Table 5. The

[27] The timing advantage of Estimize forecasts likely plays a role as well, which we explore in Section 4.1.3.

accuracy benefits from combining the Estimize consensus and the de-biased IBES consensus are approximately 40% smaller than those from combining the Estimize consensus and the unadjusted IBES consensus. This result suggests Estimize forecasts' lower bias is an important but incomplete explanation for their incremental usefulness.

We next compute a characteristic forecast (*CF*) of earnings based on firm characteristics similar to So (2013).[28] We outline the approach and report descriptive statistics for *CF* in the Internet Appendix. As in Panel A, the accuracy of a forecast (*PMAFE*) is measured as the difference between the forecast absolute error and the mean absolute forecast error (*MAFE*) across all IBES and Estimize forecasts, scaled by the mean absolute forecast error (*MAFE*).[29] The *Combined Consensus* is computed as the equally weighted average of the *Estimize Consensus* and *CF*.

Panel B of Table 6 reports the results. We find that the *Estimize Consensus* is more accurate than the *CF* as well as the *Combined Consensus* at all horizons. In the Internet Appendix, we find that weighting schemes that weight the *CF* at 5% (for all horizons) and 10% (for 30 and 60 day horizons) deliver small improvements over the *Estimize Consensus*. We conclude that at shorter horizons, where Estimize forecasts are more prevalent and enjoy a greater timing advantage over the statistical forecast, the incremental usefulness of the *CF* is relatively small. Therefore, our remaining tests benchmark Estimize to IBES forecasts only.

[28] We attempt to minimize the timing advantage of Estimize by computing a statistical forecast that also exploits information in stock returns up to the day before the earnings are announced. We acknowledge that including stock returns to bring the statistical forecast up to date is an admittedly imperfect approach to address the disparity in information sets. We leave it to future research to develop superior techniques.

[29] The distribution of the *CF* error has fat tails. To reduce the influence of outliers, we trim the PMAFE of the *CF* at 10.

4.1.3. Determinants of the Incremental Usefulness of the Estimize Consensus

The results from the prior two sections suggest that the Estimize forecasts are incrementally useful in predicting earnings, and that this usefulness is only partially explained by a difference in bias between Estimize and IBES forecasts. In this section, we further explore the factors that influence the incremental usefulness of Estimize forecasts. We are particularly interested in the effect of the number of Estimize contributors (the benefits of crowdsourcing are likely increasing in the size of the crowd) and the low Estimize forecast age (recent forecasts are generally more accurate than older forecasts). By the same reasoning, many IBES analysts and low IBES forecast age are likely factors working against this outcome.

We model the likelihood that the PMAFE of the *Combined Consensus* is less than the PMAFE of the *IBES Consensus* as a function of *Log* (*Estimize Contributors), Log (IBES Contributors), Estimize Age* defined as the average age of Estimize forecasts, *IBES Age* defined similarly, and control variables: *Size, BM, Turn,* and *Vol,* defined in Table 2. We standardize all variables to have a mean of zero and a standard deviation of one.

Specifications 1 and 2 of Table 7 report the odds ratios from a logistic regression when the forecast horizon is one day and five days, respectively. In Specification 1, we find that a one-standard-deviation increase in *Log* (*Estimize Contributors)* increases the likelihood that the *Combined Consensus* is more accurate than the *IBES Consensus* by 13%. This is consistent with the value of crowdsourced forecasts increasing in the size of the crowd. We also find that a one-standard-deviation increase in *Estimize Age* reduces the same likelihood by roughly 9%. Specification 2 presents analogous results for a five-day horizon.[30] The results are generally

[30] We have explored horizons of longer than five days and generally find less significant results. As the horizon increases, we have less power because both the sample size and the variance of our main independent variables of interest shrink.

similar, although the coefficient on *Log* (*Estimize Contributors)* is reduced and no longer significant. There is some evidence that the value of Estimize is stronger for larger companies.

In Specifications 3 and 4, we report the slope coefficients from an OLS regression of the difference between the Estimize PMAFE and the IBES PMAFE. We now find stronger evidence that the relative value of Estimize is increasing in the number of Estimize contributors and declining in the number of IBES contributors. Specifically, a one-standard-deviation increase in *Log (Estimize Contributors)* results in a 14% reduction in relative PMAFE, while a one-standard-deviation increase in *Log (IBES Contributors)* results in a 12% increase in relative PMAFE. We continue to find that Estimize is relatively more accurate when it issues forecasts closer to the announcement date (i.e., as *Estimize Age* declines) and when IBES issues earlier forecasts.

4.1.4. Combining Concurrent Estimize and IBES Forecasts

The preceding results suggest that Estimize forecasts are incrementally useful in forecasting earnings because they are less pessimistic and they incorporate more public information by virtue of being less timely. In this section, we control for these differences in order to examine another possible explanation for the incremental usefulness of crowdsourced forecasts: they provide information useful in forecasting earnings that is incremental to the information provided in concurrent IBES forecasts, and in that sense "new" information.

We begin by constructing a sample of concurrent Estimize and IBES forecasts. There are 3,005 days when at least one Estimize and one IBES forecast were issued for the same firm-quarter. We compute an Estimize (or IBES) consensus by averaging across same-day Estimize (IBES) forecasts. The average (median) same-day Estimize consensus includes 2.8 (1) unique forecasts, and the corresponding values for the IBES consensus are 1.7 (1). The mean and median forecast age for the sample is 13.3 days and 4 days, respectively. The sample is skewed toward short-term forecasts because short-term IBES forecasts are more prevalent than long-term Estimize forecasts. Thus, our tests examine the incremental usefulness of Estimize forecasts late in the quarter when Estimize contributors are relatively more active.

We regress *Actual EPS* on the *Estimize Consensus*, the *IBES Consensus*, or the *Combined Consensus* and compare model fit. By including only same-day forecasts, we control for differences in forecast timing between the two groups of forecasters. By focusing on goodness of fit, a statistic which does not depend on the independent variable's mean value, we address the concern that Estimize forecasts improve upon the IBES consensus because they are less biased.[31] Thus, only the hypothesis that Estimize forecasts convey new information predicts that the *Combined Consensus* model will have higher r-squared than the *IBES Consensus* model.

Table 8 reports the results. A comparison of Specifications 2 and 3 shows that *Combined Consensus* explains *Actual EPS* better than *IBES Consensus* does (r-squared of 97.66% vs. r-squared of 97.24%). To assess the significance of this r-squared difference, we examine the fraction of Specification 3's residuals whose absolute value is smaller than that of Specification 2's residuals. We find that 54.11% of Specification 3's residuals have absolute values smaller than those of Specification 2, an amount significantly different from the null hypothesis value of 50% (t=2.83). Therefore, we conclude that even after controlling for differences in timing and bias, Estimize forecasts are incrementally useful in predicting actual EPS.

In Specification 4, we include both *Estimize Consensus* and *IBES Consensus*, in effect relaxing Specification 3's constraint that each is equally weighted in constructing a *Combined*

[31] A limitation of our approach is that it does not address the case of a time-varying IBES forecast bias. On the other hand, it is not obvious that users can easily adjust for a time-varying IBES forecast bias, which would create investor demand for an alternative source of information.

Consensus.[32] *Estimize Consensus* is weighted more than *IBES Consensus* (0.57 vs. 0.45), but the coefficients are not statistically different from each other. Both coefficients are statistically different from zero, suggesting that neither consensus subsumes the other in predicting future earnings.

In Specification 5, we explore whether the slope coefficients on *Estimize Consensus* and *IBES Consensus* (the optimal combination weights) are a function of the number of contributors in the consensus. We interact *Estimize* Consensus and *IBES Consensus* with *Log* (*Estimize Contributors) [EC]* and *Log (IBES Contributors) [IC]*, each standardized to have a mean of zero and standard deviation of one. We find that as the number of Estimize contributors increases, the weight placed on the *Estimize (IBES) Consensus* significantly increases (decreases), highlighting the importance of crowd size as a determinant of the benefits of crowdsourcing.

4.2 Market Earnings Expectation

A related but distinct question is whether Estimize forecasts help in measuring the market's expectations of earnings. A superior measure of the market expectation exhibits a stronger association with returns at the time the actual is announced: that is, a higher Earnings Response Coefficient (ERC) (Brown, Hagerman, Griffin, and Zmijewski, 1987).[33] Thus, we explore the role of the Estimize consensus in measuring the market's expectation by estimating the regression:

$$BHAR = \alpha + \beta ConsensusError + \varepsilon. \qquad (5)$$

[32] This approach dates back to a seminal study by Bates and Granger (1969). See section 8.5 in Elliott and Timmermann's (2008, JEL) survey of the literature on economic forecasting.

[33] There is a long tradition in accounting to infer differences in earnings quality based on differences in Earnings Response Coefficients (Dechow, Ge, and Schrand, 2010). Since the Earnings Response Coefficient is also a function of the error with which the market expectation is measured (Brown, Hagerman, Griffin, and Zmijewski, 1987), reducing this measurement error is critical to improving inferences about earnings quality on the basis of evidence about differences in Earnings Response Coefficients.

BHAR is the three-day buy-and-hold size-adjusted return around the earnings announcement date (day 0), defined as:

$$BHAR = \prod_{t=-1}^{1}(1+R_{j,t}) - \prod_{t=-1}^{1}(1+R_{j,t}^{Size}). \tag{6}$$

$R_{j,t}$ is the raw return on stock *j* on day *t*, and $R_{j,t}^{Size}$ is the equally-weighted return on day *t* of a benchmark portfolio that consists of all other stocks in the same NYSE size decile. *Consensus Error* is the difference between actual earnings and the consensus forecast computed on day *t-2*.

We estimate five specifications of Equation 5, reported in Table 9. In Specifications 1-3, the independent variable is *Estimize Consensus Error*, *IBES Consensus Error*, and *Combined Consensus Error*, respectively. All three consensus forecast errors are winsorized at the 1st and 99th percentile and scaled to have a standard deviation of one. The corresponding ERCs are 2.14, 2.04, and 2.16, not statistically different from one another. When we include *Estimize Consensus Error* and *IBES Consensus Error* (Specification 4), we find that both measures are related to earnings announcement returns. The point estimate is slightly larger for *Estimize Consensus Error* (1.39 vs. 0.98), but the coefficients are not significantly different from each other. These results suggest that the Estimize and the IBES consensus forecasts are similarly accurate market expectation proxies, and that neither proxy subsumes the other.

Finally, Specification 5 augments Specification 4 by interacting *Estimize Consensus Error* and *IBES Consensus Error* with *Log (Estimize Contributors) [EC]* and *Log (IBES Contributors) [IC]*. We find that the market reaction to the *Estimize (IBES) Consensus Error* is increasing (decreasing) in the number of Estimize contributors, suggesting the Estimize consensus is better aligned with the market expectation when the Estimize contributor base is larger.

To get a better sense of the economic significance of this effect, we estimate and plot (see Figure 2) the slope coefficients from Specification 4 when the number of Estimize contributors is Low (less than three), Medium (three to five), and High (greater than five). As the number of Estimize contributors varies from Low to High, we document a strong systematic variation in the slope coefficients on the *Estimize Consensus Error* (1.10, 1.38, and 3.16) and the *IBES Consensus Error* (1.26, 0.72, and -0.82). Thus, when the number of Estimize contributors is greater than five, the Estimize consensus fully subsumes the IBES consensus as a proxy for the market expectation.

4.3. Facilitating Price Discovery

In this section, we examine the market reaction to Estimize consensus revisions. If Estimize forecasts contain information that is not already incorporated into prices, then upward (downward) revisions should be associated with positive (negative) abnormal returns.[34]

We compute the Estimize consensus revision for firm *j* on day *t* as the Estimize consensus for firm *j* on day *t* less the Estimize consensus for firm *j* on day *t-1*, scaled by the share price at the end of the prior quarter (*Rev/Price*). We winsorize *Rev/Price* at the 1st and 99th percentile, and we scale *Rev/Price* to have a standard deviation of one. Our measure of abnormal return is the size-adjusted buy-and-hold return over a two-day event window [0, 1], where day zero is the day of the Estimize consensus revision.

Our sample contains 13,798 consensus forecast revisions.[35] To better identify the effect of Estimize consensus revisions on prices, we follow Loh and Stulz (2011) and exclude revisions

[34] Our Table 8 findings only speak to the question of whether Estimize forecasts incorporate information that contemporaneous IBES forecasts fail to incorporate.

[35] Three factors explain the difference in observations between the final Estimize sample, 45,569 observations, and the sample analyzed here, 13,798. The Final Estimize Sample includes individual forecasts, many of which occur on the same day, whereas the sample analyzed here includes forecast revisions at the consensus level. We drop the first

that fall in the two-day window (-1, 0) around earnings announcements (5,860 observations), earnings guidance (72 observations), IBES recommendation changes (954 observations), and IBES forecast revisions (2,424 observations).

Specification 1 of Table 10 reports the results of the regression of abnormal returns (*BHAR*) on *Rev/Price*. We find that a one-standard-deviation increase in *Rev/Price* is associated with a 0.15% increase in two-day abnormal returns. The point estimate of 0.15% is statistically and economically significant. As a comparison, using the same approach, we find that a one-standard-deviation increase in the IBES consensus revision is associated with a 0.23% increase in abnormal returns (untabulated).

Next, we examine the price impact of upward and downward revisions to the Estimize consensus. In Specification 2, we regress *BHAR* on *Upward*, a dummy variable equal to one for upward consensus revisions. We find that upward revisions are associated with a statistically significant 0.19% *BHAR*, while downward revisions, as captured in the intercept, are associated with a statistically insignificant abnormal return of -0.07%. Since many consensus revisions are small, we explore variables indicating whether the absolute magnitude of the revision is in the top half of all upward revisions, *Large Upward*, or in the bottom half of all downward revisions, *Large Downward*. We document that extreme upward revisions are associated with a 0.26% *BHAR* while extreme downward revisions are associated with a -0.15% *BHAR*. The difference of 0.41% is statistically significant (t=3.42).

In Specification 4, we explore whether Estimize revisions are more informative when sell-side analyst coverage is low (*Low Coverage)* and when they are issued at short horizons (*Short Horizon*) where Estimize contributors are relatively more active and accurate. We also

forecast in each firm-quarter because we cannot estimate a prior consensus. We drop observations where the new forecast confirms the prior consensus forecast (i.e., consensus revision is zero).

examine whether Estimize revisions have relatively more impact when the Estimize- and IBES-reported actuals differ (*Differing Actuals*). If Estimize and IBES analysts are forecasting different measures of earnings (i.e., they differ on exclusions from GAAP earnings), then the Estimize revision may capture value relevant information excluded from the revisions of sell-side analysts (Gu and Chen, 2004). We estimate the following regression:

$$\begin{aligned} BHAR = \alpha + \beta_1 Rev/Prc + \beta_2 LowCoverage + \beta_3 Rev/Prc * LowCoverage \\ + \beta_4 ShortHorizon + \beta_5 Rev/Prc * ShortHorizon \\ + \beta_6 DifferingActuals + \beta_7 Rev/Prc * DifferingActuals + \varepsilon . \end{aligned} \tag{7}$$

Low Coverage is a dummy variable equal to one if the firm is covered by fewer than 10 IBES analysts (the sample median). *Short Horizon* is a dummy variable equal to one if the forecast horizon is less than 8 days (the sample median). *Differing Actuals* is a dummy variable equal to one if the IBES-provided actual earnings differ from Estimize-provided actual earnings.

We find that a one-standard-deviation increase in *Rev/Prc* is associated with an incremental, statistically significant 0.27% increase in *BHAR* for firms with low IBES coverage, suggesting Estimize forecasts are particularly informative for stocks with low sell-side analyst coverage. Short horizon and differing reported actuals seem to have no incremental effect, as neither of these estimates are significantly different from zero.

Finally, we plot the cumulative size-adjusted returns for *Large Upward* and *Large Downward* revisions in the 20 trading days surrounding the revision (-10, 10) in Figure 3. We observe that *Large Upward* (*Large Downward*) revisions are preceded by positive (negative) abnormal returns, consistent with Estimize contributors revising their forecasts to incorporate the arrival of public information. As documented in Table 10, we find a large return differential of 0.41% on days 0 and 1 between *Large Upward* and *Large Downward* revisions. We find no evidence that this return differential reverses over the subsequent 10 trading days.[36] The lack of reversal helps alleviate a concern that the significant two-day BHAR is driven by market participants overreacting to Estimize consensus revisions. [37]

5. Conclusions

Crowdsourcing is taking root in the investment research industry. We contribute to the understanding of this phenomenon by examining the value of crowdsourced earnings forecasts, specifically forecasts available on Estimize, an open financial estimates platform. Our sample includes forecasts submitted to Estimize by analysts, portfolio managers, and independent investors, as well as corporate finance professionals and students.

We find substantial accuracy benefits from combining IBES and Estimize forecasts at all horizons; these benefits are smaller but still significant when we restrict the sample to contemporaneous forecasts and control for differences in forecast bias between IBES and Estimize. Also, we find that the Estimize consensus is incrementally useful as a measure of the market's earnings expectation. The usefulness of the Estimize consensus in forecasting earnings and proxying for the market's earnings expectations is increasing in the number of Estimize contributors. Finally, Estimize consensus revisions appear to induce a statistically and economically significant market reaction, especially for stocks with below-median IBES coverage. We conclude that crowdsourced forecasts are incrementally useful in predicting earnings and measuring the market's expectation of earnings, and also improve price discovery.

[36] The average daily abnormal return for large downgrades over the (2, 10) period is 0.03% (t=0.75). The average daily abnormal return for large upgrades over the same period is 0.02% (t=0.91).

[37] The evidence that Estimize forecasts contain information not fully reflected in contemporaneous IBES forecasts (Table 8) or market prices (Table 10) raises the possibility that Estimize forecasts incorporate information earlier than some IBES forecasts. Consistent with this view, in the Internet Appendix, we show that Estimize revisions predict the sign of subsequent IBES revisions.

Our results are subject to several caveats. First, sell-side earnings estimates are informative, widely disseminated, and considerably timelier (released earlier) than Estimize estimates, which suggests that Estimize contributors likely learn from the sell-side, and that without the sell-side, Estimize's ability to provide new information may be compromised.[38]

Second, we acknowledge that the superiority of Estimize forecasts over statistical forecasts generated using So's (2013) regression approach (which we augment to include stock returns as an earnings predictor) may be attenuated by the use of more sophisticated statistical approaches which we leave for future research to explore.

Finally, the long-term success of the crowdsourcing model is still an open question. Information goods are notoriously difficult to price and sell, and only time will tell whether Estimize can recover the costs of operating an open financial estimates platform. Further, existing competitors may change their behavior to erode the value of Estimize. For instance, sell-side analysts may reduce bias and increase information production in the period prior to earnings announcements, and whisper sites may increase transparency and use more sophisticated methods to mine the ever-growing world of social media. To successfully address these competitive threats, Estimize may have to further grow its contributor base—a key value driver for Estimize—or successfully diversify into areas where competition is nonexistent or weak: the sourcing of private company estimates, introduced in 2013, and macroeconomic forecasts and merger predictions, introduced in 2014.

[38] In the Internet Appendix, we examine this possibility by studying Estimize forecasts when IBES coverage is absent. The results are inconsistent with Estimize conveying less new information to the market when IBES is not present; however, our analysis is based on a very small sample.

References:

AGGARWAL, R., and G. WU. 'Stock Market Manipulations.' *Journal of Business* 79 (2006): 1915-1953.

ANTWEILER, W., and M. Z. FRANK. 'Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards.' *The Journal of Finance* 59 (2004): 1259-1294.

AVERY, C., J. A. CHEVALIER, and R. J. ZECKHAUSER. 'The 'CAPS' Prediction System and Stock Market Returns.' NBER Working Paper No. 17298, National Bureau of Economic Research, 2011.

BAGNOLI, M., M. D. BENEISH, and S. G. WATTS. 'Whisper Forecasts of Quarterly Earnings per Share.' *Journal of Accounting and Economics* 28 (1999): 27-50.

BATES, J. M., and C. W. J. GRANGER. 'The Combination of Forecasts.' *Operational Research Society* 20 (1969): 451-468.

BERGER, P.G., C. HAM, and Z. KAPLAN. 'Do Analysts Say Anything About Earnings Without Revising Their Earnings Forecasts?' Working Paper, University of Chicago, 2016.

BHATTACHARYA, N., A. SHEIKH, and S. R. THIAGARAJAN. 'Does the Market Listen to Whispers?' *Journal of Investing* 15 (2006): 16-24.

BLANKENSPOOR, E., G. S. MILLER, and H. WHITE. 'The Role of Dissemination in Market Liquidity: Evidence from Firms' Use of Twitter.' *The Accounting Review* 89 (2014): 79-112.

BRABHAM, D. C. *Crowdsourcing*. Cambridge, MA: MIT Press, 2013.

BRADSHAW, M. T., M. DRAKE, J. MYERS, and L. MYERS. 'A Re-Examination of Analysts' Superiority over Time-Series Forecasts of Annual Earnings.' *Review of Accounting Studies* 17 (2012): 944–968.

BROWN, L., P. GRIFFIN, R. HAGERMAN, and M. ZMIJEWSKI. 'Security Analyst Superiority Relative to Univariate Time-Series Models in Forecasting Quarterly Earnings.' *Journal of Accounting and Economics* 9 (1987): 61-87.

BROWN, W. D., and G. D. FERNANDO. 'Whisper Forecasts of Earnings per Share: Is Anyone Still Listening?' *Journal of Business Research* 64 (2011): 476-482.

CHEN, H., P. DE, Y. J. HU, and B. H. HWANG. 'Wisdom of Crowds: The Value of Stock Opinions Transmitted through Social Media.' *Review of Financial Studies* 27 (2014): 1367-1403.

CLEMENT, M. B. 'Analyst Forecast Accuracy: Do Ability, Resources and Portfolio Complexity Matter?' *Journal of Accounting and Economics* 27 (1999): 285-303.

COOPER, R. A., T. E. DAY, and C. M. LEWIS. 'Following the Leader: A Study of Individual Analysts' Earnings Forecasts.' *Journal of Financial Economics* 61 (2001): 383–416.

COSTA, L. ‘Facebook for Finance.’ *Institutional Investor* 44 (October, 2010): 54-93.

CRAWFORD, S., W. GRAY, B. JOHNSON, and R. A. PRICE. ‘What Motivates Buy-side Analysts to Share Recommendations Online?’ Working paper, University of Houston, 2014.

DAS, S. R., and M. Y. CHEN. ‘Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web.’ *Management Science* 53 (2007): 1375-1388.

DECHOW, P., W. GE, and C. SCHRAND. ‘Understanding Earnings Quality: A Review of the Proxies, Their Determinants and Their Consequences.’ *Journal of Accounting and Economics* 50 (2010): 344-401.

ELLIOTT, G., and A. TIMMERMANN. ‘Economic Forecasting.’ *Journal of Economic Literature* 46 (2008): 3-56.

ESTELLES-AROLAS, E., and F. GONZALEZ-LADRON-DEGUEVARA. ‘Towards an Integrated Crowdsourcing Definition.’ *Journal of Information Science* 38 (2012): 189–200.

GIANNINI, R. C., P. J. IRVINE, and T. SHU. ‘The Convergence of Investors’ Opinions around Earnings News: Evidence from a Social Network.’ Working paper, Texas Christian University, 2014.

Giles, J**. ‘**Internet encyclopaedias go head to head.’ Nature 438 (2005): 900-901.

Gu, Z and T. Chen. ‘Analysts’ treatment of non recurring items in street earnings.’ Journal of Accounting and Economics 28 (2004): 129-170.

GUTTMAN, I. ‘The Timing of Analysts' Earnings Forecasts.’ *The Accounting Review* 85 (2010): 513-545.

HIRSCHEY, M., V. J. RICHARDSON, and S. SCHOLZ. ‘Stock-Price Effects of Internet Buy-Sell Recommendations: The Motley Fool Case.’ *Financial Review* 35 (2000): 147-174.

HONG, H., J. D. KUBIK, and A. SOLOMON. ‘Security Analysts’ Career Concerns and Herding of Earnings Forecasts.’ *The Rand Journal of Economics* 31 (2000): 121-144.

HOWE, J. ‘The Rise of Crowdsourcing.’ Wired (2006).

JUNG, M. J., J. P. NAUGHTON, A. TAHOUN, and C. WANG. ‘Corporate Use of Social Media.’ Working paper, Northwestern University, 2014.

LEE, L. F., A. HUTTON, and S. SHU. ‘The Role of Social Media in the Capital Market: Evidence from Consumer Product Recalls.’ *Journal of Accounting Research* 53 (2015): 367-404.

LOH, R. K., and R. M. STULZ. ‘When Are Analyst Recommendation Changes Influential?’ *Review of Financial Studies* 24 (2011): 593-627.

MILLER, G. S., and D. J. SKINNER. 'The Evolving Disclosure Landscape: How Changes in Technology, the Media, and Capital Markets Are Affecting Disclosure.' *Journal of Accounting Research* 53 (2015): 221-239.

RAMNATH, S., S. ROCK, and P. SHANE. 'The Financial Analyst Forecasting Literature: A Taxonomy with Suggestions for Future Research.' *International Journal of Forecasting* 24 (2008): 34-75.

REES, L., and D. ADUT. 'The Effect of Regulation Fair Disclosure on the Relative Accuracy of Analysts' Published Forecasts versus Whisper Forecasts.' *Advances in Accounting* 21 (2005): 173-197.

RICHARDSON, S., S. H. TEOH, and P. D. WYSOCKI. 'The Walk-Down to Beatable Analyst Forecasts: The Role of Equity Issuance and Insider Trading Incentives.' *Contemporary Accounting Research* 21 (2004): 885-924.

SO, E. C. 'A New Approach to Predicting Analyst Forecast Errors: Do Investors Overweight Analyst Forecasts?' *Journal of Financial Economics* 108 (2013): 615-640.

SHROFF, P., R. VENKATARAMAN, and B. XIN. 'Timeliness of Analysts' Forecasts: The Information Content of Delayed Forecasts.' *Contemporary Accounting Research* 31 (2014): 202-229.

TUMARKIN, R., and R. F. WHITELAW. 'News or Noise? Internet Postings and Stock Prices.' *Financial Analysts Journal* 57 (2001): 45-51.

WYSOCKI, P. D. 'Private Information, Earnings Announcements and Trading Volume or Stock Chat on the Internet. A public Debate about Private Information.' Working paper, University of Michigan,1998.

Table 1: Sample Selection

This table describes the sample selection process. The initial sample includes forecasts issued by Estimize contributors where both the forecast and the earnings announcement dates occur in the 2012 or 2013 calendar year. We eliminate forecasts issued 90 days or more before earnings are announced or after earnings are announced. We also eliminate forecasts "flagged" as unreliable based on quantitative filters developed by Estimize. Finally, forecasts issued by a contributor for a given firm on the same day are replaced with their average. The IBES-matched sample is obtained from the final Estimize sample after eliminating firm-quarters where: 1) there is no IBES coverage and 2) Estimize-reported actual EPS differ from IBES-reported actual EPS.

	Forecasts	**Firms**	**Firm-Quarters**	**Contributors**
Initial Sample	51,012	1,874	7,534	3,255
Less:				
Forecasts Issued Outside of [0, 90]	(1,512)	(4)	(53)	(67)
"Flagged" Observations	(1,090)	0	(50)	(134)
Duplicate Firm-Contributor-Day Observations	(2,841)	0	0	0
Final Estimize Sample	**45,569**	**1,870**	**7,431**	**3,054**
Less:				
Observations with no IBES coverage	(2,975)	(110)	(817)	(94)
Observations where Actual EPS reported differently in IBES and Estimize	(5,563)	(159)	(1,159)	(125)
Final IBES-Matched Sample	**37,031**	**1,601**	**5,455**	**2,835**

Table 2: Characteristics of Stocks Covered by Estimize and IBES

This table reports summary statistics for stocks covered by Estimize and IBES. Panel A reports stock characteristics for firm-quarters where 1) both Estimize and IBES provide at least one forecast, 2) only IBES issues a forecast, and 3) only Estimize issues a forecast. Panel B (C) sorts firm-quarters into quartiles based on the number of unique Estimize (IBES) contributors issuing an earnings forecast for the firm-quarter. *Size*: price times shares outstanding computed on the last day of the prior year. *BM*: book value of equity divided by size, computed on the last day of the prior year. *VOL*: the standard deviation of daily stock returns over the prior year. *Turnover*: the daily average of share volume divided by shares outstanding during the prior year.

Panel A: Sorts by Existence of IBES and Estimize Coverage

	Firm-Quarters	*Estimize Coverage*	*IBES Coverage*	*Size*	*BM*	*VOL*	*Turnover*
Estimize and IBES	6,580	6.07	10.45	13.48	0.49	2.26	12.62
IBES Only	18,041	0	5.11	2.77	0.77	2.80	8.79
Estimize Only	750	2.99	0	4.11	0.63	2.50	10.15

Panel B: Sorts by Magnitude of Estimize Coverage (Estimize and IBES Sample)

Quartile Rank	Firm-Quarters	*Estimize Coverage*	*IBES Coverage*	*Size*	*BM*	*VOL*	*Turnover*
4 (Coverage: >=7)	1,746	16.00	13.87	26.90	0.37	2.36	16.76
3 (Coverage: 4-6)	1,187	4.85	10.37	12.23	0.44	2.16	12.18
2 (Coverage: 2-3)	1,829	2.44	9.14	8.46	0.54	2.20	11.12
1 (Coverage: 1)	1,818	1.00	8.54	6.46	0.58	2.29	10.43

Panel C: Sorts by Magnitude of IBES Coverage (Estimize and IBES Sample)

Quartile Rank	Firm-Quarters	*Estimize Coverage*	*IBES Coverage*	*Size*	*BM*	*VOL*	*Turnover*
4 (Coverage: >=15)	1,729	10.13	21.31	29.83	0.47	2.18	15.50
3 (Coverage: 9-14)	1,572	5.26	11.28	12.23	0.47	2.18	12.76
2 (Coverage: 5-8)	1,527	4.56	6.37	6.68	0.50	2.33	11.76
1 (Coverage: <=4)	1,752	4.09	2.55	4.38	0.51	2.35	10.40

Table 3: Characteristics of Estimize and IBES Forecasts

This table reports summary statistics for the Final IBES-Matched Sample (See Table 1). Panel A reports summary statistics for forecasts issued by Estimize contributors. The first and second row report the distribution for the total number of forecasts and the total number of unique contributors for a firm-quarter. The third row reports the number of forecasts made by a contributor for a firm-quarter; the fourth row presents the unique number of firms covered by an Estimize contributor. The bottom row describes the distribution of forecast horizon across firm-quarters. We first compute the average forecast age across all forecasts issued for the same firm-quarter, and then describe the distribution across all firm-quarters. Panel B reports analogous statistics for forecasts issued by IBES analysts.

Panel A: Estimize Forecast Characteristics

	Obs.	Mean	Std. Dev	Q1	Median	Q3
Forecasts per Firm-Quarter	5,455	6.79	13.84	1.00	3.00	7.00
Contributors per Firm-Quarter	5,455	6.44	12.30	1.00	3.00	7.00
Forecasts per Firm-Quarter per Contributor	35,121	1.05	0.27	1.00	1.00	1.00
Estimize Firms covered by Contributor (per quarter)	4,168	8.41	36.73	1.00	1.00	3.00
Forecast Horizon	5,455	5.03	9.20	1.00	2.00	5.67

Panel B: IBES Forecast Characteristics

	Obs.	Mean	Std. Dev	Q1	Median	Q3
Forecasts per Firm-Quarter	5,455	14.62	13.62	5.00	11.00	20.00
Analysts per Firm-Quarter	5,455	10.70	7.94	4.00	9.00	15.00
Forecasts per Firm-Quarter per Analyst	58,357	1.37	0.69	1.00	1.00	2.00
Estimize Firms covered by Analyst (per quarter)	14,834	3.92	3.11	1.00	3.00	6.00
Forecast Horizon	5,455	59.30	27.01	35.50	65.00	85.00

Table 4: Comparison of Estimize and IBES Individual Forecasts – Accuracy, Bias, and Boldness

This table compares Estimize and IBES forecasts with similar horizons on three dimensions: Accuracy (Panel A), Bias (Panel B), and Boldness (Panel C). The table reports the results for five horizons ranging from 90 to 30 days prior to the earnings announcement (30, 90) to the earnings announcement day (0). Accuracy is defined as the proportional mean absolute forecast error (*PMAFE*): the forecast's absolute error less the mean absolute forecast error across all forecasts for the same firm-quarter, scaled by the mean absolute forecast across all forecasts for the same firm-quarter. *Bias* is the difference between forecasted earnings and actual earnings scaled by the stock price at the end of the previous quarter and multiplied by 100. *Boldness* is the absolute deviation of the forecast from the current consensus, scaled by the current consensus (Percent Absolute Deviation from the Consensus). The current consensus is defined as the average of individual Estimize and IBES forecasts. Each panel reports firm-quarter observations, the attribute's average value in the Estimize and IBES samples, the difference between the two samples, and the t-stats of the difference. Standard errors are clustered by firm. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels (two-tailed), respectively.

Panel A: Accuracy (Average PMAFE)

Horizon	Firm-Quarters	Estimize	IBES	Estimize-IBES	t(Estimize-IBES)
[30, 90]	959	0.21	0.11	0.10***	(2.74)
[10, 29]	1,006	0.00	-0.01	0.01	(0.29)
[5,9]	808	-0.02	-0.07	0.05	(1.59)
[1,4]	1,675	-0.09	-0.05	-0.04*	(-1.68)
[0]	159	-0.07	-0.15	0.08	(1.12)

Panel B: Bias (Forecast Error Scaled by Price)

Horizon	Firm-Quarters	Estimize	IBES	Estimize-IBES	t(Estimize-IBES)
[30, 90]	959	0.00	-0.08	0.08***	(9.12)
[10, 29]	1,006	-0.02	-0.08	0.06***	(6.33)
[5,9]	808	-0.03	-0.09	0.07***	(9.11)
[1,4]	1,675	-0.03	-0.08	0.06***	(9.00)
[0]	159	-0.03	-0.09	0.05***	(3.09)

Panel C: Boldness (Percent Absolute Deviation from Consensus)

Horizon	Firm-Quarters	Estimize	IBES	Estimize-IBES	t(Estimize-IBES)
[30, 90]	788	1.40	1.04	0.36***	(6.69)
[10, 29]	988	1.19	1.01	0.17***	(4.32)
[5,9]	801	1.10	0.94	0.17***	(4.18)
[1,4]	1,668	0.96	0.94	0.02	(0.78)
[0]	159	0.85	0.96	-0.11	(-1.32)

Table 5: Consensus Forecast Accuracy Across Different Horizons

This table examines the accuracy of the Estimize consensus, the IBES consensus, and a Combined consensus (an average across Estimize and IBES forecasts) for horizons ranging from 60 days prior to the earnings announcement (-60) to the day of the earnings announcement (0). For example, when horizon is -60 days, the Estimize consensus is the average across all Estimize forecasts issued at least 60 days before the earnings announcement. Estimize PMAFE is the absolute forecast error for the Estimize consensus of firm j for quarter t, less the mean absolute forecast error across all IBES analysts and Estimize contributors for firm j in quarter t (MAFE), scaled by the mean absolute forecast error across all analysts for firm j in quarter t (MAFE). IBES PMAFE and Combined PMAFE are calculated analogously. Combined – IBES reports the difference in accuracy between the Combined consensus and the IBES consensus, and % (Combined < IBES) is a dummy variable equal to 100% if the Combined Consensus is more accurate than the IBES consensus, and 0% otherwise. T-statistics, based on standard errors clustered by firm, are reported in parentheses. The null hypothesis is 0 or 50% (only in the last column). The sample is the Final IBES-Matched Sample (See Table 1). *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels (two-tailed), respectively.

Horizon	Obs.	Estimize PMAFE	IBES PMAFE	Combined PMAFE	Combined - IBES	% (Combined < IBES)
-60	430	0.28***	-0.07***	-0.10***	-0.03**	57.44***
		(5.11)	(-2.91)	(-4.48)	(-2.22)	(2.82)
-30	941	0.16***	-0.07***	-0.11***	-0.04***	59.72***
		(5.07)	(-3.94)	(-7.55)	(-4.06)	(5.81)
-10	1,856	0.02	-0.13***	-0.18***	-0.05***	60.83***
		(0.79)	(-9.51)	(-18.51)	(-6.98)	(8.46)
-5	2,493	-0.03**	-0.13***	-0.20***	-0.06***	61.85***
		(-2.02)	(-11.89)	(-25.29)	(-8.62)	(10.57)
-1	4,568	-0.15***	-0.15***	-0.24***	-0.08***	63.86***
		(-15.02)	(-19.79)	(-45.57)	(-13.65)	(17.44)
0	5,002	-0.17***	-0.16***	-0.25***	-0.09***	64.05***
		(-18.37)	(-27.51)	(-56.85)	(-18.26)	(20.70)

Table 6: Consensus Forecast Accuracy Across Different Horizons - Alternative Benchmarks

This table examines the accuracy of the Estimize consensus, a benchmark consensus, and a Combined consensus across different horizons. In Panel A, the benchmark consensus is a de-biased IBES consensus ($IBES^D$) and the Combined consensus is an average across all individual Estimize and $IBES^D$ forecasts. Section 4.1.2 describes the construction of the de-biased IBES forecast. Estimize PMAFE is the absolute forecast error for the Estimize consensus of firm j for quarter t, less the mean absolute forecast error across all $IBES^D$ analysts and Estimize contributors for firm j in quarter t (MAFE), scaled by the mean absolute forecast error across all analysts for firm j in quarter t (MAFE). $IBES^D$ PMAFE and Combined PMAFE are calculated analogously. Combined – $IBES^D$ reports the difference in accuracy between the Combined consensus and the $IBES^D$ consensus, and % (Combined < $IBES^D$) is a dummy variable equal to 100% if the Combined Consensus is more accurate than the $IBES^D$ consensus, and 0% otherwise. In Panel B, the benchmark consensus is a statistical forecast that incorporates information in firm characteristics and the Combined consensus is an average of the Estimize consensus and the statistical forecast. The Internet Appendix describes how the statistical forecast is obtained. T-statistics, based on standard errors clustered by firm, are reported in parentheses. The null hypothesis is 0 or 50% (only in the last column). The sample is the Final IBES-Matched Sample (See Table 1). *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels (two-tailed), respectively.

Panel A: Comparing Estimize and De-biased IBES ($IBES^D$)

Horizon	Obs.	Estimize PMAFE	$IBES^D$ PMAFE	Combined PMAFE	COMBINED -$IBES^D$	% (Combined < $IBES^D$)
-60	430	0.23***	-0.10***	-0.12***	-0.01	54.65*
		(4.70)	(-4.70)	(-5.73)	(-1.04)	(1.94)
-30	941	0.14***	-0.11***	-0.15***	-0.04***	55.79***
		(4.77)	(-7.06)	(-11.72)	(-4.23)	(3.58)
-10	1,856	0.02	-0.16***	-0.20***	-0.05***	56.25***
		(0.82)	(-14.83)	(-24.61)	(-7.29)	(5.43)
-5	2,493	-0.03**	-0.16***	-0.21***	-0.05***	56.68***
		(-2.35)	(-18.31)	(-31.17)	(-8.73)	(6.73)
-1	4,568	-0.14***	-0.18***	-0.25***	-0.07***	59.02***
		(-13.99)	(-29.88)	(-55.02)	(-14.47)	(12.40)
0	5,002	-0.15***	-0.18***	-0.26***	-0.08***	58.74***
		(-16.59)	(-32.34)	(-60.46)	(-15.36)	(12.55)

Panel B: Comparing Estimize and Characteristic Forecast (CF)

Horizon	Obs.	Estimize PMAFE	CF PMAFE	Combined PMAFE	Combined - Estimize	% (Combined < Estimize)
-60	382	0.33***	2.38***	1.02***	0.69***	42.41%***
		(5.79)	(14.32)	(11.57)	(7.48)	(-3.00)
-30	840	0.19***	2.51***	0.99***	0.80***	41.43%***
		(5.60)	(21.53)	(16.47)	(12.82)	(-5.04)
-10	1,701	0.02	2.42***	0.91***	0.88***	36.74%***
		(1.12)	(30.04)	(22.15)	(20.39)	(-11.34)
-5	2,297	-0.02	2.38***	0.87***	0.89***	37.01%***
		(-1.45)	(34.53)	(25.04)	(24.31)	(-12.90)
-1	4,255	-0.15***	2.25***	0.78***	0.92***	34.78%***
		(-14.26)	(45.61)	(31.46)	(35.67)	(-20.84)
0	4,668	-0.16***	2.19***	0.75***	0.91***	35.07%***
		(-16.89)	(47.11)	(31.14)	(37.19)	(-21.38)

Table 7: Determinants of the Incremental Usefulness of the Estimize Consensus

This table explores the determinants of the relative forecast accuracy of the Estimize consensus. In Specifications 1 and 2, the dependent variable is a dummy variable equal to one if the combined consensus (an average across all individual Estimize and IBES forecasts) is more accurate than the IBES consensus. In Specifications 3 and 4, the dependent variable is the accuracy of the Estimize consensus less the accuracy of the IBES consensus. The consensus is computed either one day prior to the earnings announcement date (Specifications 1 and 3) or five days prior to the earnings announcement date (Specifications 2 and 4). Accuracy is measured as the proportional mean absolute forecast error (PMAFE) as defined in Table 5. *Estimize Age* is the average age of all forecasts in the Estimize consensus. *Estimize Contributors* is the number of unique individuals contributing to the Estimize consensus. *IBES Age and IBES Contributors* are defined analogously. *Size*, *Book-to-Market* (*BM*), *Turnover* (*Turn*) and Volatility (*Vol*) are defined as in Table 2. All variables are standardized to have a mean of zero and standard deviation of one. Specifications 1 and 2 are estimated using a logistic regression and the coefficients represent odds ratios. Specifications 3 and 4 are estimated using OLS. Standard errors are clustered by firm, and z-scores (in Specifications 1 and 2) and t-statistics (in Specifications 3 and 4) are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels (two-tailed), respectively.

	Logistic Regression		OLS	
	Combined < IBES PMAFE		Estimize – IBES PMAFE	
	[1]	[2]	[3]	[4]
Intercept			0.01	0.10***
			(0.35)	(4.36)
Estimize Age	0.91***	0.92**	0.07***	0.11***
	(-2.89)	(-2.11)	(3.46)	(4.03)
IBES Age	1.26***	1.24***	-0.08***	-0.10***
	(6.44)	(4.53)	(-5.43)	(-3.71)
Log (Estimize Contributors)	1.13***	1.07	-0.14***	-0.11***
	(2.98)	(1.25)	(-7.47)	(-4.84)
Log (IBES Contributors)	0.96	0.94	0.12***	0.17***
	(-0.90)	(-1.07)	(6.31)	(5.39)
Log (Size)	1.09	1.19**	-0.04*	-0.09**
	(1.64)	(2.55)	(-1.78)	(-2.52)
Log (BM)	1.03	1.09*	-0.01	-0.02
	(0.76)	(1.72)	(-0.71)	(-0.82)
Log (Turn)	1.00	1.02	-0.04**	-0.07**
	(-0.04)	(0.26)	(-2.14)	(-2.35)
Log (Vol)	1.01	1.04	0.02	0.01
	(0.17)	(0.44)	(0.78)	(0.18)
Horizon	1	5	1	5
Observations	4,264	2,312	4,264	2,312
Pseudo R^2 (R^2)	2.21%	2.13%	4.54%	5.34%

Table 8: Consensus Forecast Accuracy – Horizon Matched Sample

This table examines the accuracy of the Estimize consensus, the IBES Consensus, and the Combined Consensus (an average across all individual Estimize and IBES forecasts), holding forecast horizon constant. The number of firm-day observations where there is at least one Estimize and one IBES forecast is 3,005. The number of individual Estimize (IBES) forecasts is 8,321 (5,143). The table reports parameter estimates from panel regressions of actual EPS on *Estimize Consensus*, *IBES Consensus*, and *Combined Consensus*. Each consensus variable is constructed by averaging appropriate individual forecasts. Specification 5 interacts *Estimize Consensus* and *IBES Consensus* with the natural log of the number of Estimize contributors *(EC)* and the natural log of the number of IBES contributors (*IC*). T-statistics, based on standard errors clustered by firm, are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels (two-tailed), respectively.

	[1]	[2]	[3]	[4]	[5]
Intercept	-0.01	0.01	0.00	0.00	-0.01
	(-0.48)	(0.77)	(-0.05)	(-0.16)	(-0.71)
Estimize Consensus	1.01***			0.57***	0.40**
	(48.36)			(4.22)	(2.57)
IBES Consensus		1.03***		0.45***	0.65***
		(48.47)		(3.37)	(4.22)
Combined Consensus			1.02***		
			(47.10)		
*Estimize Consensus * EC*					0.26**
					(2.03)
*IBES Consensus * EC*					-0.29**
					(-2.30)
*Estimize Consensus * IC*					-0.02
					(-0.10)
*IBES Consensus * IC*					0.03
					(0.15)
Log (Estimize Contributors) [EC]					0.01*
					(1.95)
Log (IBES Contributors) [IC]					-0.01
					(-0.64)
Observations	3,005	3,005	3,005	3,005	3,005
R-squared	97.41%	97.24%	97.66%	97.65%	97.79%

Table 9: Market Reaction to Unexpected Earnings Proxy Variables

This table examines the market reaction to proxies for unexpected earnings. Market reaction is defined as the cumulative size-adjusted return for the three days surrounding the earnings announcement date (-1, 1). Unexpected earnings proxies include the *Estimize Consensus Error*, the *IBES Consensus Error*, and the *Combined Consensus Error*. The Estimize consensus includes all forecasts made by Estimize contributors on day t-2 or earlier. If a contributor issued multiple forecasts, we include only the most recent forecast. The IBES consensus is defined analogously. The Combined Consensus is the average of individual Estimize and IBES forecasts. For each consensus measure (Estimize, IBES, and Combined), we compute the forecast error as the actual earnings less the consensus forecast, scaled by the price at the end of the previous quarter. Consensus forecast errors are winsorized at the 1st and 99th percentile. Specification 5 interacts *Estimize Consensus Error* and *IBES Consensus Error* with the natural log of the number of Estimize contributors *(EC)* and the natural log of the number of IBES contributors (*IC*). All variables are standardized to have a mean of zero and standard deviation of one. T-statistics, based on standard errors clustered by firm, are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels (two-tailed), respectively.

	[1]	[2]	[3]	[4]	[5]
Intercept	0.25**	-0.26**	-0.16	0.00	0.55
	(2.05)	(-2.01)	(-1.22)	(0.02)	(1.30)
Estimize Consensus Error	2.14***			1.39***	1.07
	(11.53)			(5.35)	(1.53)
IBES Consensus Error		2.04***		0.98***	2.05***
		(11.44)		(4.06)	(3.25)
Combined Consensus Error			2.16***		
			(11.44)		
*Estimize Consensus Error * EC*					0.68**
					(2.25)
*IBES Consensus Error * EC*					-0.44*
					(-1.74)
*Estimize Consensus Error * IC*					-0.05
					(-0.18)
*IBES Consensus Error * IC*					-0.36
					(-1.46)
Log (Estimize Contributors) [EC]					-0.10
					(-0.70)
Log (IBES Contributors) [IC]					-0.20
					(-1.08)
Observations	3,429	3,429	3,429	3,429	3,429
R-squared	7.40%	6.74%	7.51%	8.05%	8.62%

Table 10: Market Reaction to Estimize Consensus Revisions

This table examines the market reaction to Estimize consensus revisions. The dependent variable is the cumulative size-adjusted return for the two days surrounding the change in the consensus (0, 1). *Rev/Price* is computed as the Estimize consensus on day *t* less the consensus on day *t-1*, scaled by the stock price as of the prior quarter. The day *t* consensus is the average across all forecasts issued on day *t* or earlier. If a contributor has issued multiple forecasts that meet this criteria, we select the most recent forecast. *Rev/Price* is winsorized at the 1st and 99th percentile, and scaled to have a standard deviation of one. *Upward* is a dummy variable equal to one if the change in the consensus is positive. *Large Upward* is a dummy variable equal to one if the change in the consensus is greater than the median breakpoint across all upward revisions. *Large Downward* is a dummy equal to one if the change in the consensus is less than the median breakpoint across all downward revisions. *Low Coverage* is a dummy equal to one if the firm is covered by fewer than 10 IBES analysts (the median breakpoint for analyst coverage). *Short Horizon* is a dummy equal to one if the forecast is made within 8 days of the earnings announcement (the median forecast age for this sample). *Differing Actuals* is a dummy equal to one if the IBES-provided actual earnings differ from Estimize-provided actual earnings. The sample includes 4,448 Estimize consensus revisions. The sample excludes Estimize consensus revisions that occur on the day of, or a day after, major events such as earnings announcements, earnings guidance, and published IBES research (i.e., forecast revisions or recommendation changes). T-statistics, based on standard errors clustered by firm, are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels (two-tailed), respectively.

	[1]	[2]	[3]	[4]
Intercept	0.04	-0.07	0.00	-0.04
	(0.72)	(-1.03)	(0.07)	(-0.59)
Estimize (Rev/Price)	0.15**			-0.03
	(2.31)			(-0.28)
Estimize Upward		0.19**		
		(2.32)		
Estimize Large Upward			0.26**	
			(2.30)	
Estimize Large Downward			-0.15	
			(-1.40)	
Low Coverage				-0.01
				(-0.06)
*Estimize * Low Coverage*				0.27**
				(2.40)
Short Horizon				0.13
				(1.43)
*Estimize * Short Horizon*				0.17
				(0.90)
Differing Actuals				0.03
				(0.21)
*Estimize * Differing Actuals*				0.09
				(0.48)
Observations	4,488	4,488	4,488	4,488
R-squared	0.30%	0.12%	0.28%	0.63%

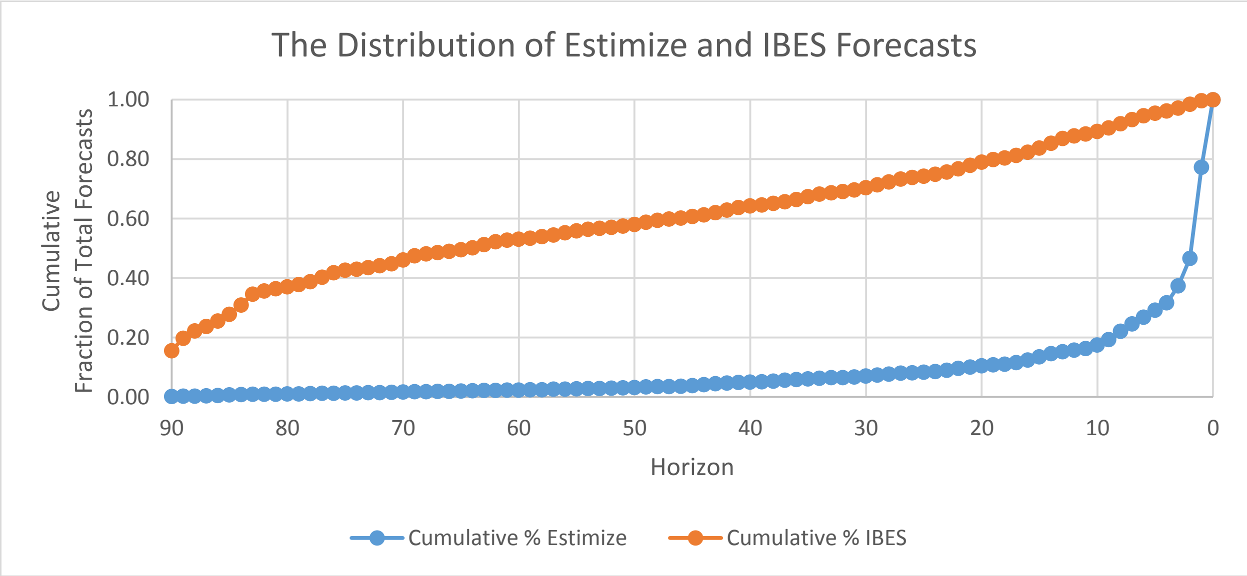

Figure 1: The Distribution of Individual Estimize and IBES Forecasts over a 90-Day Forecast Period
This figure plots the fraction of the total Estimize and IBES forecasts in the final Estimize-IBES matched sample with a horizon longer than or equal to *t*, where *t* ranges from day 90 to day 0 (earnings announcement day).

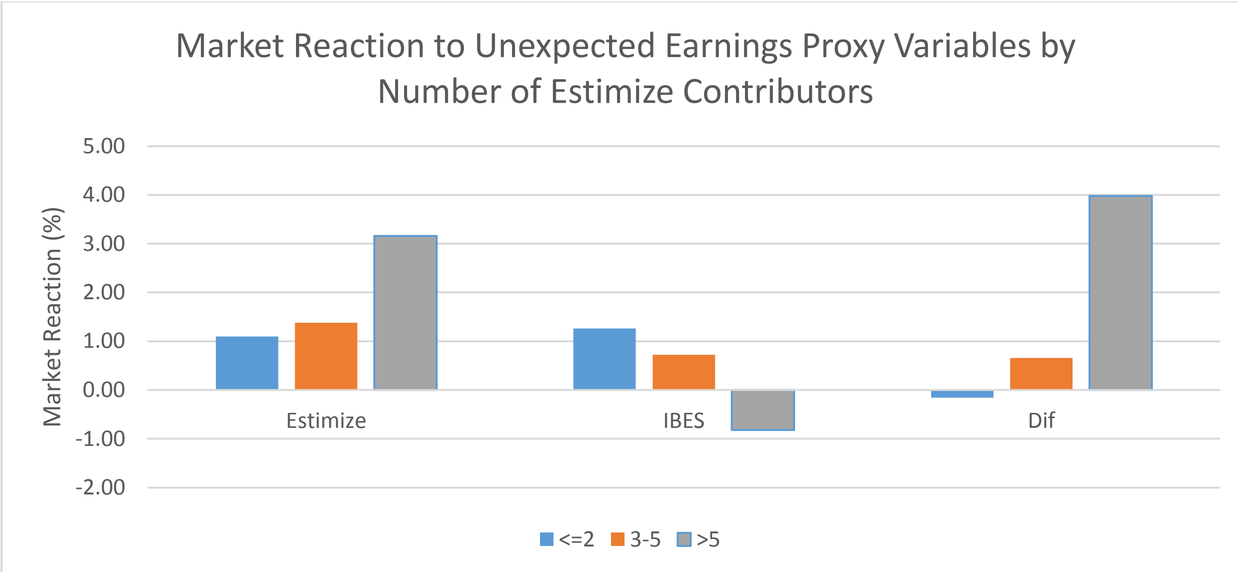

Figure 2: Market Reaction to Unexpected Earnings Proxy Variables Conditional on Number of Estimize Contributors

This table plots slope coefficients on the *Estimize Consensus Error* and the *IBES Consensus Error* from Specification 4 of Table 9 when the number of Estimize contributors in a firm-quarter is 2 or fewer (blue bar), 3-5 (orange bar), and more than 5 (gray bar). *Dif* measures the difference between the coefficients on *Estimize* and *IBES.*

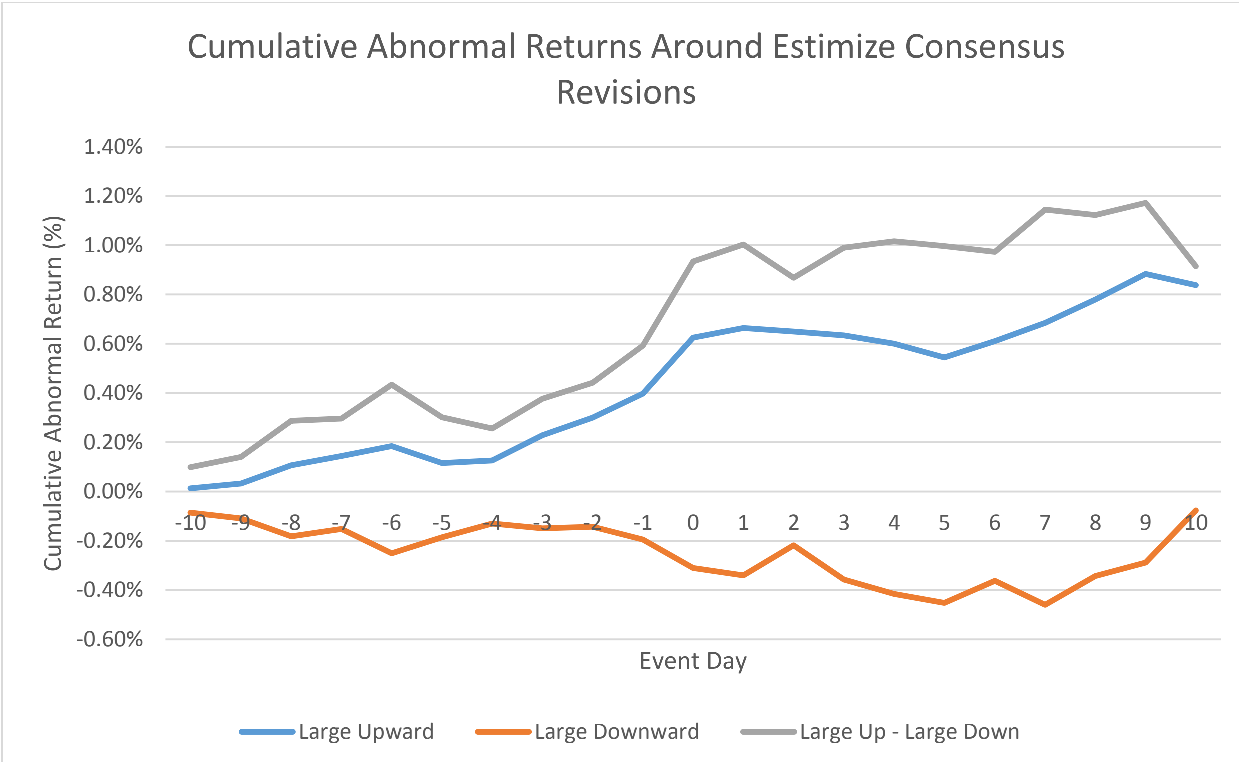

Figure 3: Cumulative Abnormal Returns around Estimize Consensus Revisions

This figure plots cumulative size-adjusted returns around large Estimize consensus revisions. We compute the day *t* consensus as the average of all Estimize forecasts issued on day *t* or earlier. If a contributor has issued multiple forecasts, we include only the most recent forecast. Estimize consensus revision is the change in the Estimize consensus from day *t-1* to day *t*, scaled by the stock price at the end of the previous quarter. Large Upward revisions include the top half of the upward revisions. Large Downward revisions are defined analogously. Day 0 is the day of the Estimize revision. The figure plots cumulative abnormal returns starting ten days prior to the revision (day -10) and ending 10 days after the revision (day 10). The sample includes 1,053 large upward revisions and 1,191 large downward revisions. Excluded are Estimize revisions that occur on the day of, or the day after, major events such as earnings announcements, earnings guidance, and published research (i.e., forecast revisions or recommendation changes) by IBES analysts.