Visualizing Income Statements

Jon A. Garfinkel[a], Paul Hribar[b], Lawrence Hsiao[c]

Abstract

A firm's comparative income statement is conceptually a grid with line items down the rows and successive quarters across the columns, a structure well suited to image-based deep learning. We render each firm's income statement as an image that preserves both the numerical values and the spatial structure of its major line items and train a convolutional neural network to extract features that predict post-filing stock returns. We find that firms with higher income statement buy probability (ISBP) earn higher post-filing returns out-of-sample. This remains robust after standard risk adjustments, is incremental to established return anomalies, and persists over time. We further show that ISBP is positively associated with future earnings growth, yet stock prices do not fully incorporate these earnings implications, consistent with investor underreaction and the resulting return predictability. Overall, our findings suggest that the visual representation of income statements contains economically meaningful information for predicting returns.

[a]Henry B. Tippie Research Professor of Finance, Tippie College of Business, University of Iowa, jon-garfinkel@uiowa.edu

[b]Henry B. Tippie Excellence Chair in Accounting and Professor, Tippie College of Business, University of Iowa, paul-hribar@uiowa.edu

[c]Assistant Professor of Finance, College of Management, National Taiwan University, lawrencehsiao@ntu.edu.tw

1 Introduction

A central goal of financial statement analysis is to use the information reported in financial statements to assess a firm's value and to forecast its future performance and stock returns. The fundamental analysis literature pursues this goal by combining numbers drawn from across the financial statements into composite signals of future profitability and value. For example, Ou and Penman (1989) distill dozens of financial statement items into a single summary measure that predicts future earnings and returns; Lev and Thiagarajan (1993) and Abarbanell and Bushee (1997, 1998) identify a set of fundamental signals on which a trading strategy earns abnormal returns; and Piotroski (2000) aggregates nine accounting indicators into a composite score that separates winners from losers among value firms (see also Holthausen and Larcker, 1992; Mohanram, 2005). A related literature instead extracts price-relevant signals from transformations of individual income statement items, such as standardized unexpected earnings (Ball and Brown, 1968), revenue surprises (Jegadeesh and Livnat, 2006), gross profitability (Novy-Marx, 2013), operating profitability (Fama and French, 2015), tax expense surprises (Thomas and Zhang, 2011), and earnings acceleration (He and Narayanamoorthy, 2020).

Recent advances in machine learning and artificial intelligence (AI) have allowed researchers to revisit the prediction of earnings and returns from financial statements. A growing body of work applies machine learning models to financial statement data rather than a few hand-selected ratios. Chen et al. (2022) use tree-based models on detailed financial-statement items to predict earnings changes and form profitable hedge portfolios; Binz et al. (2025) apply machine learning to profitability decomposition frameworks to sharpen out-of-sample forecasts of profitability and returns; and Kim et al. (2024) show that a large language model can forecast future earnings from standardized financial statements about as well as professional analysts. What these studies share is that the financial statement is supplied to the model as numerical data. We extend this literature by asking a different question: what can AI learn when the financial statement is presented not as a set of numbers but as a visual image of the statement itself?

Specifically, we investigate whether an AI model, trained from scratch on raw income statement items arranged in their natural two-dimensional, grid-like form, with rows representing line items and columns representing quarters, can learn features that predict stock returns following firms' 10-K/10-Q filings.[1] Several features of this approach may allow us to uncover information that traditional fundamental analysis discards. Rendering the income statement as an image preserves two types of adjacency. First, down each column, the vertical articulation within a quarter from sales revenue through cost of goods sold, gross profit, operating income, and income before extraordinary items retains the proportionality across related items. Second, across each row the trajectory of a single item over the most recent nine quarters captures its trend, acceleration, and volatility. A convolutional neural network (CNN) is designed to exploit exactly this local two-dimensional structure as its filters scan small blocks of the image and learn configural patterns that span several items and several quarters at once (for example, revenue accelerating while gross margin holds steady and operating costs stay flat). Such patterns are nonlinear and relational, describing how groups of items co-move and evolve together rather than the level of any single ratio, and are difficult to represent through a fixed combination of researcher-chosen inputs. To the extent that these configurations carry leading information about the persistence of a firm's operating performance, they may help forecast future earnings and the returns that follow.

Like other deep-learning models, however, the CNN does not deliver a transparent mapping from inputs to predictions, and the features it learns are not directly observable. Rather

[1]The AI model we employ is a convolutional neural network (CNN), which is particularly effective at extracting spatial patterns and features from images. Inspired by the human visual system, CNNs have achieved remarkable success in a wide range of applications, including object detection, video recognition, and medical image analysis. See, for example, LeCun et al. (1989).

than treat the model as a black box, we confront this limitation directly. Using SHAP values from the explainable AI literature (e.g. Lundberg and Lee, 2017) we reverse-engineer the model's predictions to quantify how much each quarterly income statement item contributes to its output, recovering an economically interpretable account of what the CNN finds predictive and confirming that it draws on sensible accounting information rather than spurious artifacts.

Our research proceeds as follows. First, for each firm filing 10-Ks/10-Qs between 1996Q1 and 2005Q4 (model training window), we construct its "income statement image" that captures both the magnitude and the spatial structure of income statement items over the most recent nine quarters. Each column contains nine quarterly income statement items, ordered from top to bottom as sales revenue, cost of goods sold, gross profit, operating expenses, operating income, interest income/expenses, pre-tax income, taxes and minority interests, and income before extraordinary items, all scaled by total assets in the preceding quarter.[2] Taken together, the income statement image not only encodes the value of each income statement item across the past nine quarters, but also preserves their spatial arrangement within each period.

Next, each income statement image is assigned with one of the three labels ("sell", "hold", or "buy") based on the relative performance of the firm's 63-day buy-and-hold abnormal returns after 10-K/10-Q filings among the cross-section of firms in the same quarter. We train a convolutional neural network (CNN) on these in-sample income statement images to learn features that best distinguish among the three assigned labels. Lastly, we construct out-of-sample income statement images beginning in 2006Q2 and apply the trained model to generate a firm-specific income statement buy probability (ISBP) for each filing, which can be thought of as the CNN-predicted likelihood for an image to be a "buy" when "sell" and "hold" options are also available.[3]

[2]Scaling by total lagged assets fosters comparability between firms of different sizes. In addition, quarter q appears in the leftmost column, followed by $q-1$, …, $q-7$, and $q-8$ in the rightmost column.
[3]We apply the trained CNN to income statement images from 2006Q2 onward to guard against look-ahead bias. See Section 3 for details.

If CNN is capable of extracting return-predictive information from income statement images, firms with higher income statement buy probabilities (ISBP) should, on average, earn higher post-filing abnormal returns. To test this prediction, in each quarter we sort firms filing 10-Ks and 10-Qs into decile portfolios based on ISBP, using cutoffs derived from the cross-sectional distribution of ISBP in the previous quarter. We find that a high-minus-low hedge portfolio sorted on ISBP earns positive and highly significant post-filing buy-and-hold abnormal returns in the out-of-sample period (2006Q3 to 2023Q2). Specifically, the average hedge portfolio returns 2.8% to 3.5% when we use a variety of buy-and-hold abnormal return measures.[4] Hence, AI's ability to detect return-predicting features from income statement images cannot be attributed to lack of risk controls.

Next, we compare the average hedge portfolio returns sorted on ISBP to those sorted on 16 well-known income-statement-based return predictors and firm characteristics, including the standardized unexpected earnings (Ball and Brown, 1968; Foster et al., 1984; Bernard and Thomas, 1989), revenue surprise (Jegadeesh and Livnat, 2006), tax expense surprise (Thomas and Zhang, 2011), earnings acceleration (He and Narayanamoorthy, 2020), trend in gross profitability (Akbas et al., 2017), firm size (Fama and French, 1992, 1993), book-to-market ratio (Fama and French, 1992, 1993), filing return, pre-filing return, earnings persistence (Francis et al., 2004), earnings volatility (Cao and Narayanamoorthy, 2012), gross profitability (Novy-Marx, 2013), operating profitability (Ball et al., 2016), operating accruals (Sloan, 1996; Hribar and Collins, 2002), total accruals (Richardson et al., 2005), and asset growth (Cooper et al., 2008). We find that the average hedge portfolio returns sorted on ISBP continue to stand out across all measures of buy-and-hold abnormal returns.

We proceed to examine whether the return-predicting power of ISBP resembles that of the

[4]We employ the market-adjusted buy-and-hold returns, the size-adjusted buy-and-hold returns, and the buy-and-hold returns adjusted by factor models including the Fama-French four- and six-factor models (Fama and French, 1993, 2015, 2018; Carhart, 1997), the q^5-model (Hou et al., 2015; Hou et al., 2021), and the risk-and-behavioral model (Daniel et al., 2020). See Table 2 and the Appendix for more details.

firm characteristics discussed above. In two-way sorting analyses, we find that the return-predicting ability of ISBP persists across almost all quintiles of the other firm characteristics, suggesting that ISBP captures information that is largely distinct from existing firm characteristics known to predict returns. In contrast, the return-predicting power of the considered firm characteristics largely disappears across ISBP quintiles. In addition, we estimate Fama and MacBeth (1973) regressions of the 63-day post-filing buy-and-hold abnormal returns on ISBP along with the 16 variables. We find that the average coefficients on ISBP remain positive and highly significant, suggesting that ISBP provides incremental predictability for post-filing returns beyond the existing return anomalies.

Next, we leverage recent advances in explainable AI to identify which quarterly income statement items are most influential in determining ISBP. Specifically, we estimate SHAP values (Lundberg and Lee, 2017), which measure the marginal contribution of each quarterly income statement item to the model's predictions. We find that the four most influential inputs are sales revenue in quarters q, $q-1$, and $q-4$, as well as cost of goods sold in quarter $q-2$. When aggregating importance across the most recent nine quarters, sales revenue emerges as the most important income statement item, followed by cost of goods sold and gross profit. We further document that these importance rankings are highly stable across firms of different sizes, book-to-market ratios, and product-market characteristics, including firm fluidity (Hoberg et al., 2014), product similarity (Hoberg and Phillips, 2016), vertical integration (Frésard et al., 2020), and firm scope (Hoberg and Phillips, 2025).

We find that ISBP positively predicts one-quarter-ahead earnings growth, suggesting that the visual features identified by the CNN contain information about firms' future operating performance. To investigate whether this information explains ISBP's return predictability, we employ the Mishkin test (Mishkin, 1983; Abel and Mishkin, 1983), which jointly examines the relation between ISBP and future earnings growth and the extent to which stock prices reflect these earnings implications. The results indicate that stock prices do not fully incorporate the favorable earnings implications of ISBP at the time of filing. Consequently, the market reaction to ISBP is weaker than warranted by its association with future earnings growth, generating predictable post-filing returns.[5]

We perform two additional analyses. First, we find that alternative machine learning models that take the same set of one-dimensional quarterly income statement items as inputs underperform our baseline CNN model in predicting post-filing returns, thereby highlighting the importance of the spatial layout of accounting items in the input representation. Second, we show that implementing a more conservative monthly-rebalanced long–short strategy based on our predictions yields an average monthly return of approximately 1.1%.

Our study makes several contributions. Its primary contribution is the idea that a firm's income statement can be represented as an image capturing a grid of accounting items arrayed across successive quarters, and demonstrating how a convolutional neural network can extract return-relevant information from this spatial structure. This idea connects to two established literatures. The first examines the price-relevant information in transformations of individual income statement items, such as sales revenue (Jegadeesh and Livnat, 2006), gross profits (Novy-Marx, 2013; Akbas et al., 2017), operating profits (Fama and French, 2015), tax expenses (Thomas and Zhang, 2011), and earnings (Ball and Brown, 1968; He and Narayanamoorthy, 2020). But rather than hand-engineering a signal from one item at a time, we let the model learn from a snapshot of the statement as a whole. The second literature applies machine learning to asset pricing, typically by supplying the model a curated set of firm characteristics or return predictors (Rapach et al., 2013; Kelly et al., 2019; Feng et al., 2020; Freyberger et al., 2020; Kozak et al., 2020; Gu et al., 2020, 2021; Leippold et al., 2022; Cao et al., 2024; Chen et al., 2024; Murray et al., 2024), and extends to mutual fund alphas (DeMiguel et al., 2023; Kaniel et al., 2023), hedge fund returns (Wu et al., 2021), bond risk premia (Bianchi

[5]Prior studies employing the Mishkin test framework include Sloan (1996), Dechow and Sloan (1997), Rangan and Sloan (1998), Collins and Hribar (2000), Narayanamoorthy (2006), Cao and Narayanamoorthy (2012), Chen and Shane (2014), Hui et al. (2016), Ma and Markov (2017), and He and Narayanamoorthy (2020).

et al., 2021; Kelly et al., 2023), and option returns (Büchner and Kelly, 2022; Bali et al., 2023). Rather than feeding the model predictors chosen in advance, we let it learn directly from raw accounting data. This design also mitigates concerns that investors may overestimate AI's ability to extract return-predictive information from financial statements (Li et al., 2025).

Second, our study contributes to a fast-growing literature that applies machine learning and artificial intelligence to financial statement analysis. Chen et al. (2022) use tree-based models on detailed financial-statement items to forecast earnings changes; Binz et al. (2025) estimate profitability-decomposition frameworks with machine learning to improve out-of-sample forecasts of profitability and returns; and Kim et al. (2024) show that a large language model can predict future earnings directly from standardized financial statements (see also Amel-Zadeh et al., 2020). Each of these studies supplies the financial statements to the model as numerical data or text, discarding its two-dimensional layout. We depart by preserving that layout. Representing the statement as an image allows the CNN to exploit spatial relationships among line items and across time. Our contribution, in this sense, is not merely that AI can predict returns from financial statements, but that the visual, grid-like structure of the statement is itself a source of incremental return-predictive information.

Third, our paper contributes to a burgeoning literature on visualized data. For example, researchers have examined how market participants respond to visuals in firms' Twitter earnings announcements (Nekrasov et al., 2022), visualized earnings surprises displayed on Robinhood (Moss, 2022), visual information in corporate executive presentations (Cao et al., 2022), and persuasion in start-up pitch videos (Hu and Ma, 2025). In addition, Gu et al. (2025) construct a sentiment measure derived from GIFs embedded in messages posted on Stocktwits. Unlike these studies, which focus on the information content of existing visual disclosures, we develop a standardized procedure for converting raw income statement data into visual representations. The fact that the return-predictive power of ISBP persists after controlling for conventional return anomalies further underscores the incremental informational value of visualization.

Finally, our study adds to a recent literature employing CNN models to uncover information from image-like inputs. For example, Obaid and Pukthuanthong (2022) construct a daily investor sentiment index by extracting information from photos embedded in Wall Street Journal articles, and Jiang et al. (2023) utilize charts depicting historical prices and trading volume to predict future stock returns. We differ from the CNN literature in two respects. First, our image inputs are not news photos, bar charts, or line graphs, but raw accounting items arranged to mirror their presentation in income statements. Second, we go beyond demonstrating the predictive performance of CNN models to investigate their decision-making process using recent advances in explainable AI, quantifying the importance of each input variable.

2 Data and Variables

We obtain data from Compustat, CRSP, and Edgar, and focus on U.S. common stocks traded on the three major exchanges (NYSE, AMEX, and NASDAQ). We describe the sample selection process in Table 1 and summarize the definitions of all the variables in the Appendix. First, we collect Compustat firm-quarters whose earnings announcement date (Compustat item RDQ) is between January 1996 and June 2023, and delete observations whose current quarter RDQ is before the quarter fiscal period end date.[6] Then, we remove observations with missing/non-positive book equity in the most recent nine quarters (quarters q - 8 to q) or missing/non-positive total assets in the most recent ten quarters (quarters q - 9 to q).

Next, we eliminate observations with missing sales revenue (SALES), cost of goods sold (COGS), gross profits (GP), operating expenses (OPEXP), operating income (OPINC), interest income/expenses (INT), pre-tax income (PTINC), taxes and minority interests (TAXMI), or income before extraordinary items (IBEI) in the most recent nine quarters. In addition, we

[6] Following Arif et al. (2019), our sample starts at 1996 because of limited filing data in Edgar before 1996.

require IBEI to equal to the Compustat item IBQ to avoid drawing inferences from observations that are potentially subject to data errors (Casey et al., 2016).

We further drop observations with missing filing data in Edgar or observations whose 10-K/10-Q filing dates in the current quarter precede the earnings announcement dates. We also remove late filers (i.e., 10-K filings greater than 90 days after fiscal year end or 10-Q filings greater than 45 days after fiscal quarter end). In addition, firms are required to have a CRSP daily price higher than one dollar at the current quarter's 10-K/10-Q filing date and at least 90 non-missing daily return observations in the [−150, −31] window relative to the current quarter's 10-K/10-Q filing date. Finally, we drop observations with missing firm size in the previous fiscal year end. After applying the above filters, we are left with 241,698 firm-quarter observations.

Next, we define the in-sample dataset and the out-of-sample dataset. The in-sample dataset consists of 90,544 firm-quarter observations between January 1996 to December 2005, and is used for CNN model training and validation. The corresponding out-of-sample dataset is used for predicting and testing the out-of-sample CNN model performance and thus serves as the dataset for our empirical analyses. It consists of 144,131 firm-quarter observations between July 2006 to June 2023 with non-missing firm characteristics.[7] These firm characteristics include standardized unexpected earnings (SUE), revenue surprise (RS), tax expense surprise (TES), earnings acceleration (EA), trend in gross profitability (TREND), filing return (RET[−1, +1]), pre-filing return (RET[−30, −2]), earnings persistence (PERSIST), earnings volatility (VOL), firm size (SIZE), book-to-market ratio (BM), gross profitability (GPA), operating profitability (OPA), operating accruals (OA), total accruals (TA), and asset growth (AG). These firm characteristics serve as benchmark and control variables throughout the paper. Detailed definitions of all variables used in this study are provided in the Appendix.

[7]The reason to set up a six-month lag between the end of the in-sample dataset and the out-of-sample dataset is to guard against look-ahead bias. We discuss this in more detail in Section 4.1.

To mitigate the impact of outliers, we follow prior research (e.g., Rangan and Sloan, 1998; Livnat and Mendenhall, 2006; Garfinkel and Sokobin, 2006; He and Narayanamoorthy, 2020) to transform all non-return variables into decile ranks (numbered 0 to 9, from low to high).[8] The cutoff points for these variables are based on the distribution of these variables in the previous quarter. Then, we convert all the decile ranks to scaled ranks by dividing by 9 and subtracting 0.5. The resulting scaled ranks vary from −0.5 to 0.5 with a mean of zero and a range of one. This variable transformation approach is to facilitate comparison of the economic magnitudes of firm characteristics. For example, the coefficient on a variable of interest (in scaled rank) in a return regression represents the return from a zero-investment strategy of going long on the highest variable decile and short on the lowest variable decile.

3 CNN Model Training

In this section, we describe the training procedures of our CNN model. Our goal is to have machines automatically extract the collective return-predicting information from firms' income statements in a nonlinear fashion. In particular, income statements are two-dimensional, with rows specifying income statement items and columns specifying time periods. As a result, both the values and the relative positioning of accounting items in an income statement matter, and the grid-like nature of income statements makes CNN particularly suitable for our goal.

Our use of CNN resembles its image classification task, where machines are trained to learn the relationship between images and their assigned labels. In our context, the input is an "income statement image" summarizing a firm's operating performance in the most recent nine quarters (quarters q - 8 to q), and the corresponding label is based on the firm's relative abnormal return performance following its most recent 10-K/10-Q filing dates.

[8]Variables that are not transformed into decile ranks are the six measure of the 63-day post-filing buy-and-hold abnormal return, including size-adjusted return (SAR), market-adjusted return (MAR), and four factor-adjusted returns (FF4, FF6, HMXZ5, and DHS3). See Section 4.1 for more details.

3.1 Plotting income statement images

To begin, we locate for each firm-quarter observation, its nine major income statement items. These include sales revenue (SALES), cost of goods sold (COGS), gross profits (GP), operating expenses (OPEXP), operating income (OPINC), interest income/expenses (INT), pre-tax income (PTINC), taxes and minority interests (TAXMI), and income before extraordinary items (IBEI). We do this for the most recent nine quarters, including the current one.[9]

Next, we divide the nine income statement items in a given quarter by the firm's total assets in the prior quarter to facilitate comparison of income statements across different firms in different quarters. After scaling, we record the values of each quarter's nine income statement items in the columns of the 9×9 pixels matrix, with the values of quarter q's (quarter q - 8's) accounting items appearing in the leftmost (rightmost) column.[10] Figure 1 displays an example of such an image. By construction, in each column (each quarter), the nine income statement items have the following relationships:

$$\mathrm{SALES} - \mathrm{COGS} = \mathrm{GP} \tag{1}$$

$$\mathrm{GP} - \mathrm{OPEXP} = \mathrm{OPINC} \tag{2}$$

$$\mathrm{OPINC} + \mathrm{INT} = \mathrm{PTINC} \tag{3}$$

$$\mathrm{PTINC} - \mathrm{TAXMI} = \mathrm{IBEI} \tag{4}$$

In particular, INT is positive (negative) when total interest income is greater (lower) than total interest expenses.

Unlike earnings, most income statement items are not available on earnings announcement dates, but eventually become available on firms' 10-K/10-Q filing dates. Since 10-K/10-Q filings dates are not easily accessible in early periods previous studies (e.g., Hou et

[9] This window spans two full annual cycles, allowing the CNN model to capture potential seasonality in income statement items. In addition, the construction of certain variables in our analysis (i.e., standardized unexpected earnings and trends in gross profitability) also requires two years of past data

[10] This format aims to follow those of firms' 10-K/10-Q income statements, where accounting items measured over the most recent year (quarter) are reported in the leftmost column.

al., 2020; Jensen et al., 2023) conservatively assume that accounting items are available six months (or four months) after fiscal period end dates to prevent hindsight bias.[11] However, Bowles et al. (2024) show that this convention underestimates return predictability of anomaly trading signals in more recent periods since it uses stale information.

3.2 Assigning labels to income statement images

Next, we assign each income statement image a label indicative of the firm's post-filing return performance. Specifically, we use the 63-day size-adjusted buy-and-hold return (SAR) to measure the abnormal return performance following a firm's 10-K/10-Q filing.[12] In particular, $\mathrm{SAR}_{i,q+1}$ is defined as the difference between the buy-and-hold return of firm i and that of its size-matched portfolio over the windows ([+2, +64]) in trading days relative to firm i's 10-K/10-Q filing date t in quarter q:

$$\mathrm{SAR}_{i,q+1} = \prod_{k=t+2}^{t+64} \left(1 + R_{i,k}\right) - \prod_{k=t+2}^{t+64} \left(1 + R_{p,k}\right) , \tag{5}$$

where R_i is the delisting-adjusted return of firm i and R_p is the return of firm i's size-matched portfolio.[13] We use the monthly NYSE size decile breakpoints at the end of June in year t to determine the size-matched portfolio for a firm whose 10-K/10-Q filing date is between July of year t and June of year $t+1$. The monthly size breakpoints and daily size portfolio returns are obtained from Kenneth French's website.

To mitigate the impact of bid-ask bounce and other market microstructure effects, we follow previous studies (Vega, 2006; Engelberg et al., 2012; Frank and Sanati, 2018) to compute SAR from day 2, and our results are robust to SAR defined using the trading window of [+1, +63]. Then, for each quarter in the in-sample period (1996Q1 to 2005Q4), we sort firms

[11] Before EDGAR became available online, obtaining firms' 10-K and 10-Q filings typically required visiting the SEC's Public Reference Room in Washington, D.C., and reviewing the filings in person.
[12] The 63-day holding window corresponds to the total number of trading days in three months.
[13] We replace missing delisting-adjusted returns with market returns, which is equivalent to reinvesting any remaining proceeds in the market portfolio until the end of the holding period.

filing 10-Ks/10-Qs into terciles based on their 63-day post filing SAR, with the bottom, mid, and top terciles labeled "sell", "hold", and "buy", respectively.[14] In this way, each income statement image is paired with a return label based on the relative performance of its SAR among the cross-section of firms in the same quarter.[15] Note that constructing in-sample income statement images (along with their return labels) requires return information from 1996Q1 to 2006Q1.[16]

3.3 Training the CNN

We feed in-sample income statement images, along with their return labels, into the CNN model for training and validation. The CNN uses convolutional filters to "scan" input images; these filters are small matrices (typically 3 × 3 pixels) applied across pixels to detect local features. We employ two convolutional blocks, with 64 and 128 filters of 3 × 3 pixels, respectively. During training, we follow the CNN literature and minimize the cross-entropy loss function. We randomly split the sample into a 70% training set and a 30% validation set, and apply regularization techniques similar to those in Gu et al. (2020) and Jiang et al. (2023) to mitigate overfitting.

For brevity, we report the detailed model architecture and training procedure in the Internet Appendix. Since CNN training is stochastic due to optimization algorithms and dropout, it can produce different outcomes even with the same architecture and data; accordingly, we independently train the same CNN ten times and store the resulting parameters for subsequent use.

3.4 Applying the Trained CNN for Prediction

After in-sample CNN training, we apply the estimated model parameters capturing the

---

[14]Since the number of training images for each label is about the same, we mitigate the class imbalance issue in CNN training that arises with a disproportionate ratio of labels.
[15]While we assign return labels based on SAR, our results are robust to using alternative definitions of abnormal returns such as market-adjusted or factor-adjusted returns in the label-assigning process.
[16]Income statement images (along with their return labels) constructed in a given quarter q requires return information in both quarters q and $q + 1$.

learned relationship between image features and outcome labels to out-of-sample observations. For each out-of-sample income statement image, the trained model generates the predicted probability that the image belongs to the "buy" class, which is the outcome of interest in our analysis. We define this probability as the income statement buy probability (ISBP).

Next, for firms filing 10-Ks or 10-Qs between 2006Q2 and 2023Q2, we construct their income statement images and apply the trained CNN to generate ISBP for each image. Because the CNN is estimated using information only up to 2006Q1, all out-of-sample predictions begin in 2006Q2, ensuring they are strictly post-training and free of look-ahead bias. To enhance robustness, we further average ISBP across the ten independently trained CNN models for each out-of-sample image.

4 CNN Performance

4.1 Univariate Portfolio Analysis

In each quarter starting from 2006Q3, we sort firms filing 10-Ks/10-Qs into decile portfolios based on their income statement buy probability (ISBP), with breakpoints determined by the cross-sectional distribution of ISBP in the previous quarter. This approach (Foster et al., 1984; Bernard and Thomas, 1989) mitigates hindsight bias by ensuring that portfolio formation relies only on information available at the time the strategy is implemented. Accordingly, the out-of-sample period begins in 2006Q3, six months after the end of the in-sample period (2005Q4). Then, we compute the average 63-day size-adjusted buy-and-hold returns (SAR) for each ISBP decile. If the CNN model is capable of detecting useful return-predicting information from income statement images, firms with higher ISBP should significantly outperform firms with ISBP in the 63-day SAR after the 10-K/10-Q filing dates.

Table 2 presents the results. We find that firms in the highest (lowest) ISBP decile have an average income statement probability of 39.9% (28.2%). In addition, firms in the highest decile on average outperform firms in the lowest decile by 3.5% (t-statistic = 8.046) in the post-filing 63-day SAR, which corresponds to an annualized return of 14%. In addition, when plotting the time-series of the return differential in the out-of-sample period in Figure 2, we find that the return differential is positive in 60 out of 68 quarters. This result suggests that the return-predicting power of ISBP is highly stable over time.

The significant return differential between firms in the highest and lowest deciles of income statement buy probability may exist due to lack of proper risk controls. In order to address this concern, we employ alternative measures of abnormal returns, including market-adjusted buy-and-hold returns (MAR) as well as factor-adjusted buy-and-hold returns. We compute the former by replacing $R_{p,k}$ in equation (5) with CRSP value-weighted daily return, and compute the latter by replacing $R_{p,k}$ in equation (5) with daily return $\widehat{R_{F,k}}$ predicted by factor models. To compute $\widehat{R_{F,k}}$, we first estimate individual stock factor loadings by regressing returns on the factors in a 120-day rolling window from $t-150$ to $t-31$ for each stock:

$$r_{i,t} = \alpha_i + \beta_i' F_t + \epsilon_{i,t}, \tag{6}$$

where $r_{i,t}$ is the excess return on stock i and F_t is a vector of factors. The predicted return $\widehat{R_{F,k}}$ is then computed as $\widehat{\beta'}_\iota F_k$.[17] We consider the factors in the Fama-French four- and six-factor models (Fama and French, 1993, 2015, 2018; Carhart, 1997;), the q[5]-model (Hou et al., 2015; Hou et al., 2021), and the risk-and-behavioral model (Daniel et al., 2020). The 63-day factor-adjusted buy-and-hold returns following the 10-K/10-Q filings of these models are denoted FF4, FF6, HMXZ5, and DHS3, respectively.[18]

---

[17]See, for example, Savor (2012) and Kapadia and Zekhnini (2019).

[18]Fama and French (2015) extends the Fama-French three-factor model (Fama and French, 1993) to control for operating profitability (RMW) and investment (CMA). After the inclusion of a momentum factor (Carhart, 1997), we have Fama-French four-factor and six-factor models (Fama and French, 2018). Hou et al. (2015) propose the q-model to control for market, size (ME), investment (IVA), and profitability (return on equity, ROE), and Hou et al. (2021) further includes an expected growth factor (EG) into the q[5]-model. Daniel et al. (2020) propose a 3-factor risk-and-behavioral model that accounts for market, long-term financing (FIN), and short-term earnings surprise (PEAD). Fama-French factors are obtained from Kenneth French's website, q[5]-model factors are obtained from Lu Zhang's website, and DHS3 factors are obtained from Lin Sun's website.

In the remaining columns of Table 2, we report the average 63-day MAR and factor-adjusted buy-and-hold returns (FF4, FF6, HMXZ5, and DHS3) for each income statement buy probability decile, respectively. The return differential between the highest and lowest income statement buy probability deciles range from 2.8% to 3.4%, with *t*-statistics all statistically significant at the 1% level. Overall, Table 2 shows that the significantly positive relation between income statement buy probability and the post-filing buy-and hold abnormal returns is robust to various risk adjustments.[19]

We proceed to examine whether the return-predicting power of income statement buy probability is comparable to those of the existing return anomalies known to predict returns. In particular, we consider a total of 16 firm characteristics, including standardized unexpected earnings (SUE, Ball and Brown, 1968; Foster et al., 1984; Bernard and Thomas, 1989), revenue surprise (RS, Jegadeesh and Livnat, 2006), tax expense surprise (TES, Thomas and Zhang, 2011), earnings acceleration (EA, He and Narayanamoorthy, 2020), trend in gross profitability (TREND, Akbas et al., 2017), filing return (RET[-1, +1]), pre-filing return (RET[-30, -2]), earnings persistence (PERSIST, Francis et al., 2004), earnings volatility (VOL, Cao and Narayanamoorthy, 2012), firm size (SIZE, Fama and French, 1992, 1993), book-to-market ratio (BM, Fama and French, 1992, 1993), gross profitability (GPA, Novy-Marx, 2013), operating profitability (OPA, Ball et al., 2016), total accruals (TA, Richardson et al., 2005), operating accruals (OA, Sloan, 1996; Hribar and Collins, 2002), and asset growth (AG, Cooper et al., 2008).

We follow the approach in the previous section to assign firms filing 10-Ks/10-Qs into decile portfolios based on one of the above 16 firm characteristics, where the cutoffs are based on the distribution of the previous quarter's firm characteristic. Then, we examine the average

[19] The results are qualitatively the same if we assign income statement images with labels based on firms' 21-day or 42-day post-filing size-adjusted buy-and-hold returns, train the CNN model, and examine the 21-day or 42-day post-filing buy-and-hold abnormal returns for ISBP deciles. In addition, the return differential between the highest and lowest ISBP deciles continues to be significant when we use larger convolutional filters (5×5).

hedge portfolio returns, which are defined as the difference in 63-day buy-and-hold abnormal returns following 10-K/10-Q filings between firms in the highest and lowest deciles of each variable over the out-of-sample period.

Table 3 reports the results. The first row reports the average hedge portfolio returns sorted on ISBP (ranging from 2.8% to 3.5%), which is the same as the last row in Table 2. Turning to the remaining rows, we find that the average hedge portfolio returns range from 1.4% to 1.8% for standardized unexpected earnings (SUE), from 1.4% to 1.7% for revenue surprise (RS), from 0.9% to 1.2% for tax expense surprise (TES), from 1.0% to 1.5% for earnings acceleration (EA), from 0.7% to 1.4% for filing return (RET[-1, +1]), from 1.6% to 2.6% for gross profitability (GPA), and from 0.7% to 2.0% for operating profitability (OPA). In particular, we find that average hedge portfolio returns sorted on ISBP are always the largest across all measures of buy-and-hold abnormal returns. For the other eight firm characteristics (TREND, RET [-30, -2], PERSIST, VOL, SIZE, BM, OA, and AG), the average hedge portfolio returns fail to remain statistically significant at the 5% level across at least two of the six abnormal return measure. Overall, Table 3 suggests that the income statement buy probability is more successful in predicting post-filing abnormal return performance compared to known return anomalies.

4.2 Bivariate Portfolio Analysis

In this section, we examine whether the return-predicting power of income statement buy probability (ISBP) overlaps with that of the 16 return predictors. In particular, we construct 5 × 5 portfolios sorted independently (Liu et al., 2018; He and Narayanamoorthy, 2020) on ISBP and one of the 16 predictors. Again, to alleviate hindsight bias, we use the distribution of each firm characteristic in the previous quarter to form the quintile cutoffs. Then, we examine the average 63-day SAR of the 25 portfolios, and report the double sorts results in Table IA1 in the

Internet Appendix.[20] We find that the average hedge portfolio returns sorted on ISBP persist to be positively significant across almost all quintiles of 16 predictors. In contrast, the average hedge portfolio returns based on the 16 predictors mostly appear insignificant across different quintiles of ISBP.[21] The results suggest that ISBP's return-predictive power is orthogonal to that of the 16 variables.

4.3 Cross-sectional regression

Next, we perform a cross-sectional regression analysis to simultaneously control for the firm characteristics that may affect the positive relation between the income statement buy probability (ISBP) and the 63-day buy-and-hold abnormal returns following 10-K/10-Q filings. We estimate Fama and MacBeth (1973) regressions in which the dependent variable is the firm's 63-day SAR. Each quarter, we run the following cross-sectional regression:

$$\text{SAR}_{i,q+1} = \alpha_q + \beta_q \text{Income statement buy probability}_{i,q} + \sum \beta_{c,q} \text{Controls}_{i,q} + \varepsilon_{i,q+1}, \quad (7)$$

where i refers to the stock, q refers to the quarter, and the income statement buy probability and control variables are converted into scaled ranks ranging from −0.5 to 0.5 with a mean of zero. Then, we average the cross-sectional coefficients across all quarters, where the weights correspond to the number of observations in each quarterly cross-sectional regression. In addition to using SAR as the dependent variable, we also employ MAR and four factor-adjusted buy-and-hold returns (FF4, FF6, HMXZ5, and DHS3).

Table 4 presents the regression results. The coefficient on ISBP in Column 1 is 0.013 (*t*-statistic = 3.176), suggesting that a long-short strategy of going long on the highest income statement buy probability decile and short on the lowest decile generates a 63-day SAR of around 1.3%, controlling for other firm characteristics simultaneously. In columns 2 to 5 where

[20]The results are robust to alternative return measures (MAR, FF4, FF6, HMXZ5, and DHS3).
[21]For example, the average hedge portfolio returns sorted on ISBP are significantly positive (ranging from 1.4% to 3.1%) across all five SUE quintiles. In contrast, the hedge portfolio returns sorted on SUE are statistically indistinguishable from zero in four out of five income statement buy probability quintiles.

we replace SAR with MAR and factor-adjusted buy-and-hold returns, the coefficients on ISBP range from 0.011 to 0.016 and are all statistically significant at the 1% level, suggesting that our results are not driven by omission of risk factors. For the other 16 firm characteristics, we find that standardized unexpected earnings (SUE) and revenue surprise (RS) positively predict post-filing returns, and higher asset growth (AG) is associated with lower post-filing returns. In addition, while gross profitability (GPA) exhibits significant return predictability in univariate portfolio analysis (the average hedge portfolio returns based on GPA are the largest among the 16 firm characteristics, only next to those based on ISBP), its return-predicting power is largely subsumed when other firm characteristics are included.

Overall, the results in Table 4 provide strong support for the out-of-sample return-predicting power of the ISBP, which cannot be accounted for by several known return anomalies or lack of risk controls.[22]

5 The Decision-Making Process of CNN Predictions

In this section, we examine the importance of quarterly income statement items in shaping ISBP via SHAP values (Lundberg and Lee, 2017).[23] Based on cooperative game theory, SHapley Additive exPlanations (SHAP) values approximate changes in the model predictions had one excluded a certain characteristic in its estimation process. In our context, each quarterly income statement item is a characteristic contributing to the final predicted ISBP of an income statement image. The SHAP value of a quarterly income statement item with respect to ISBP quantifies such contribution, with the 81 SHAP values summing to the difference between the image's ISBP and the average ISBP of randomly drawn income statement images.

To quantify the importance of a quarterly income statement item, we compute the average

[22]In untabaluted tests, we estimate cross-sectional regressions of ISBP on the 16 predictors. The resulting adjusted R-squared is 38.6%, indicating that more than 60% of the variation in ISBP remains unexplained after jointly controlling for all 16 predictors. This result is consistent with the evidence in Table 4 and suggests that ISBP captures a substantial amount of information that is distinct from these established return predictors.

[23]Prior studies employing SHAP values include Bali et al. (2023), Demiguel et al. (2023), and Geertsema and Lu (2023).

absolute SHAP value of the quarterly income statement item across all income statement images in the out-of-sample period.[24] For better interpretation, the 81 importance values are scaled such that they sum to one. Figure 3 displays the results in a bar chart (where the quarterly income statement items are presented by order of their importance) as well as in a heatmap (where darker colors represent higher importance). In particular, sales revenue in quarters q - 1, q - 4, and q as well as cost of goods sold in quarter q - 2 are the four most important characteristics, accounting for 3.03%, 2.73%, 2.67%, and 2.58% in shaping the income statement buy probability, respectively. As a comparison, if each quarterly income statement is equally important, the contribution should be around 1.23% (= 100%/81).

Next, we explore how different values of a quarterly income statement item affect ISBP in a nonlinear fashion. In particular, Figure 4 displays SHAP plots for the above four most important characteristics (i.e., quarterly income statement items) in shaping ISBP predictions: sales revenue in quarter $q - 1$ (top left panel), sales revenue in quarter $q - 4$ (top right panel), sales revenue in quarter q (bottom left panel), and cost of goods sold in quarter $q - 2$ (bottom right panel). Each characteristic is winsorized at the 1st and 95th percentiles.[25] The observations in the out-of-sample period are then sorted into 20 groups based on the value of the characteristic. For each group, the conditional mean SHAP value (i.e., the average SHAP value across all observations within the group) is plotted on the y-axis against the group's mean characteristic value on the x-axis. We find that there exists a substantial degree of nonlinearity in the relation between the four quarterly income statement items and their conditional mean SHAP values. Panel A shows a negative relation between $SALE_{q-1}$ and its conditional mean SHAP value, while this relation becomes insignificant for higher values of $SALE_{q-1}$. In Panel B, the conditional mean SHAP value decreases approximately linearly when $SALE_{q-4}$ is below

[24]While SHAP values can be computed for either in-sample or out-of-sample observations, we calculate them on the testing sample so that the resulting feature attributions reflect the model's out-of-sample predictive behavior rather than its in-sample fit.

[25]Since the four characteristics are highly positively skewed, we winsorize their values at the bottom 1% and top 5% using all observations in the testing sample prior to sorting the observations into 20 groups.

0.5, but drops sharply once $SALE_{q-4}$ exceeds 0.5. In Panel C, we observe a positive albeit nonlinear relationship between $SALE_q$ and ISBP predictions, with a big V-shaped relation when $SALE_q$ is between 0.4 and 0.5. Finally, Panel D indicates a somewhat negative relationship between $COGS_{q-2}$ and ISBP predictions, although the relation is highly nonlinear.

We proceed to examine the importance of an income statement item by adding up the scaled importance values of the quarterly income statement items in the most recent nine quarters (quarters q - 8 to q), and display the results in Figure 5. We find that the most important income statement item is sales revenue (SALES, 22.33%), followed by cost of goods sold (COGS, 19.14%) and gross profits (GP, 13.68%). On the other hand, income before extraordinary items (IBEI, 6.97%), interest income/expenses (INT, 3.19%), and taxes and minority interests (TAXMI, 2.39%) are the least influential income statement items. The fact that earnings (income before extraordinary items) do not play an important role in ISBP predictions is somewhat expected, as the majority of the price reactions to earnings should be concentrated in days immediately following earnings announcements.

Next, we examine whether the relative importance of income statement times varies across different firm size and book-to-market ratio. In addition, since sales revenue, cost of goods sold, and gross profits are the top three important income statement items, a natural follow-up question is whether the importance of these income statement items in shaping ISBP predictions vary with different product market attributes. For example, is sales revenue considered more important when the firm is in a more competitive industry? Or, does cost of goods sold contribute more to the income statement buy probability when the firm is more vertically integrated? To answer these questions, each quarter we classify firms in the out-of-sample period into quintiles based on firm size (SIZE), book-to-market ratio (BM), and four product market attributes, including firm fluidity (Hoberg et al., 2014), product similarity (Hoberg and Phillips, 2016), level of vertical integration (Frésard et al., 2020), and firm scope

(Hoberg and Phillips, 2025), respectively.[26] Next, we compute the importance of the nine income statement items for the five quintiles of each variable of interest

Figure IA2 in the Internet Appendix report the results. We find that the importance of sales revenue and cost of goods sold in shaping ISBP predictions is mostly increasing in firm size and book-to-market ratio. In addition, the importance of cost of goods sold increases when the firm is more vertically integrated, but decreases when the product market around a firm is changing more intensively (high firm fluidity) or when the firm has lower market power (high product similarity). Despite of these observations, we find that the importance rankings of the nine income statement items remain the same across different firm size, book-to-market ratio, or product market attribute quintiles.

6 The Source of ISBP's Return Predictability

In this section, we investigate the source of ISBP's return predictability. We conjecture that this predictability arises because investors do not fully incorporate the implications of ISBP for future earnings growth. To examine this mechanism, we implement the Mishkin test (Mishkin, 1983; Abel and Mishkin, 1983), which jointly evaluates the relation between ISBP and future earnings growth and the extent to which stock prices reflect this earnings implication. Under this hypothesis, ISBP should positively predict subsequent earnings growth. However, if investors do not fully incorporate these earnings implications into prices, the price response will be weaker than implied by the earnings relation, generating predictable post-filing returns.

To begin with, we run Fama and MacBeth (1973) regressions of one-quarter-ahead earnings growth on ISBP in the out-of-sample period. We use the previously defined SUE as the earnings growth measure (Bernard and Thomas, 1989; Ball and Bartov, 1996), and run the following regression:

[26] Firm fluidity measures how intensively the product market around a firm is changing in a year, and firm scope refers to the number of distinct product markets a firm operates in a year. The data is obtained from the Hoberg-Phillips Data Library at https://hobergphillips.tuck.dartmouth.edu/.

$$\mathrm{SUE}_{q+1} = \alpha + \gamma_1 \mathrm{ISBP}_q + \sum \gamma_c \mathrm{Controls}_q + \delta_{q+1}, \tag{8}$$

where SUE_{q+1} represents one-quarter-ahead earnings growth. Control variables are the 16 return predictors considered before.[27] In column 1 of Table 5, the average coefficient on ISBP are positive and statistically significant at the 1% level, suggesting that the income statement buy probability is a significant predictor of earnings growth in the subsequent quarter. In other words, then investors do not appear to fully incorporate the implications of income statement buy features for future earnings announcement. Economically, moving from the lowest decile to the highest decile of ISBP in the current quarter leads to a 7.3% incremental increase in the decile of one-quarter-ahead earnings growth, controlling for past earnings growth and other firm characteristics. On the other hand, the coefficients on SUE are significantly positive, consistent with the previous findings in the literature.

Next, we examine whether ISBP is positively related to the three-day abnormal return around the one-quarter-ahead earnings announcement ($\mathrm{RET}[-1, +1]_{q+1}$). One advantage of using shorter-window returns is that they are less susceptible to risk considerations (Bernard and Thomas, 1990; Sloan, 1996; Narayanamoorthy, 2006; Cao and Narayanamoorthy, 2012). In column 2 of Table 5, the average coefficient on ISBP is positive and statistically significant at the 1% level, indicating a positive relation between ISBP and the three-day abnormal return around the next earnings announcement date. In terms of economic magnitude, moving from the lowest decile to the highest decile of ISBP in the current quarter leads to a 0.5% incremental increase in $\mathrm{RET}[-1, +1]_{q+1}$.

Since ISBP has positively predicts both one-quarter-ahead earnings growth and for the three-day abnormal return surrounding the next earnings announcement date, we structure an econometrics framework, i.e., the Mishkin test (Mishkin, 1983; Abel and Mishkin, 1983), to

[27]In particular, SUE_q serves as the control for the well-documented earnings autocorrelation pattern in prior studies.

test whether the market fully understands the implications of the income statement buy features for SUE_{q+1}.[28] In particular, we simultaneously estimate the following two equations via iterative-weighted non-linear least squares (Mishkin, 1983), and the coefficients with * represent the coefficients inferred from market investors' expectation of SUE_{q+1}:

$$\text{Forecasting equation: } \text{SUE}_{q+1} = \alpha + \gamma_1 \text{ISBP}_q + \sum \gamma_c \text{Controls}_q + \delta_{q+1} \quad (9)$$

$$\text{Rational pricing equation: } \text{AR}_{q+1} = \beta(\text{SUE}_{q+1} - \alpha^* - \gamma_1^* \text{ISBP}_q - \sum \gamma_c^* \text{Controls}_q) + \varepsilon_{q+1}. \quad (10)$$

Equation (9) is the same as equation (8) and characterizes the evolution of earnings growth, and equation (10) is based on a linear abnormal return (AR) model (e.g., Sloan, 1996).[29] In other words, if the market correctly interprets the implications of the income statement buy features for future earnings growth as depicted in equation (9), AR_{q+1} should only be related to the earnings growth surprise $(\text{SUE}_{q+1} - E_q[\text{SUE}_{q+1}] = \delta_{q+1})$, but not related to the income statement buy probability in quarter q.

We test whether $\gamma_1 = \gamma_1^*$ holds or not, i.e., whether the observed relation between SUE_{q+1} and income statement buy features is the same as the relation between SUE_{q+1} and income statement buy features implicit in AR_{q+1}. In other words, $\gamma_1 = \gamma_1^*$ indicates that investors are fully aware of the implications of income statement buy features for SUE_{q+1}, and this restriction yields a likelihood ratio test statistic that has a chi-square distribution with one degree of freedom.[30]

We report the estimated coefficients, *t*-statistics based on firm and quarter double-

[28]Studies that adopt the Mishkin test include or example, Sloan (1996), Dechow and Sloan (1997), Rangan and Sloan (1998), Collins and Hribar (2000), Narayanamoorthy (2006), Cao and Narayanamoorthy (2012), Chen and Shane (2014), Hui et al. (2016), Ma and Markov (2017), and He and Narayanamoorthy (2020).

[29]Kraft et al. (2007) shows that the exclusion of control variables from the forecasting and rational pricing equations leads to an omitted variables problem. That is, if the variables omitted are not rationally priced and are also correlated with the variable of interest in the forecasting equation, then the source of market inefficiency cannot be correctly identified. Hence, we include various control variables that may be related to ISBP.

[30]The test statistic of the Mishkin test is $2 \times n \times \ln(SSR^c/SSR^u)$ distributed asymptotically $\mathcal{X}^2(q)$, where q is the number of constraints imposed by market efficiency, n is the number of observations in each equation, SSR^c is the sum of squared residuals from the constrained weighted system, and SSR^u is the sum of squared residuals from the unconstrained weighted system.

clustered standard errors, and likelihood ratio test statistics of the Mishkin test in Table 6. In particular, AR_{q+1} is measured by either the buy-and-hold market-adjusted returns from a three-day window around quarter $q+1$'s earnings or the quarter-long window starting two days after the 10-K/10-Q filings in quarter q and ending on one day after the next earnings announcement date. All variables except for the abnormal return AR are converted into scaled ranks ranging from −0.5 to 0.5 with a mean of 0.

We find that $\gamma_1 = \gamma_1^*$ is rejected at the 1% level for the three-day-window (likelihood ratio statistic = 6.853) and rejected at the 5% level for the quarter-long window (likelihood ratio statistic = 4.369). Hence, investors do not fully understand the implications of ISBP for future earnings growth. In particular, both the three-day and quarter-long window γ_1^* are statistically indistinguishable from zero, implying that investors are completely ignoring the positive implications of ISBP for future earnings growth. Overall, the results suggest that the positive relation between ISBP and post-filing buy-and-hold abnormal returns is consistent with investors missing the implications of ISBP for one-quarter-ahead earnings growth.

7 Additional Analyses

We conduct two additional analyses in this section. First, a natural follow-up question is whether treating the 81 quarterly income statement items as one-dimensional data and applying alternative machine learning models would achieve return-predictive performance comparable to our baseline CNN model, which uses income statement images as inputs. We consider three alternative models: a logit model with elastic net regularization (EN), a random forest (RF) model, and a Long-Short Term Memory (LSTM) model. To foster comparison, we maintain the same in-sample and out-of-sample periods as well as the 70/30 training to validation ratio.[31]

Table IA2 in the Internet Appendix reports the average hedge portfolio returns during the out-of-sample period based on each machine learning model's buy probability. Compared with

[31]We describe the hyperparameter choices of these models in the Internet Appendix.

the baseline CNN model reported in the last row of Table 2, the alternative models produce 63-day post-filing hedge returns that are uniformly smaller in magnitude, suggesting that they are less effective at identifying return-predictive information. These results suggest that preserving the spatial structure of income statement items through image-based representations provides price-relevant information that is not fully captured when the same accounting data are treated as one-dimensional inputs.[32]

In addition, we examine whether sorting stocks based on the income statement probability to form a more conservative month-based rebalancing trading strategy (Hou et al., 2020; Jensen et al., 2023) can still generate profits. In particular, at the end of each month t in the out-of-sample period, we sort firms into deciles based on the income statement buy probability computed using the most recent income statement image. For a firm to enter the portfolio formation at the end of month t, we require that the most recent 10-K/10-Q filing date to be within three months prior to portfolio formation to exclude stale information. Then, we examine the average returns in the subsequent month $t + 1$ for each income statement buy probability decile. Table IA3 presents the average portfolio returns for each income statement buy probability decile. Panel A shows that a hedge portfolio going long in the top income statement buy probability decile and short in the bottom decile yields an average equal-weighted monthly return of 1.1%, with a t-statistic of 6.208. The factor-adjusted hedge returns range from 0.9% to 1.1% and are all statistically significant at the 1% level. Hence, the return-predicting power of the income statement buy features persists under a more conservative month-based rebalancing strategy.[33]

[32]While our results show that the CNN model outperforms several machine learning models in predicting post-filing returns, this finding should be interpreted with caution. There is unlikely to be a universally superior model for return prediction, and a carefully tuned alternative model may outperform CNNs in specific settings. Rather than attempting to identify the best return forecasting model—a goal that is arguably unattainable—our objective is to demonstrate that CNNs can effectively analyze financial disclosures with a grid-like structure and extract economically meaningful information from them.

[33]However, we find that the hedge portfolio return is insignificant if we form value-weighted portfolios. This is because CNN model is treating each input image equally during the training phase. Hence, CNN predictions are unsurprisingly more accurate when we employ CNN stored parameters to form out-of-sample portfolios with equal-weights.

8 Conclusion

In this study, we examine whether the visual representation of income statements contains price-relevant information that can be extracted by AI and used to predict future stock returns. To address this question, we construct "income statement images" that closely resemble the income statements reported in 10-K and 10-Q filings. Specifically, the images preserve both the values and spatial arrangement of major income statement items over the most recent nine quarters. We then apply CNNs to these images to automatically extract features that are predictive of abnormal stock returns following corporate filings.

In out-of-sample tests, we find that firms in the highest income statement buy probability (ISBP) decile significantly outperform firms in the lowest ISBP decile in the post-filing period across different measures of buy-and-hold abnormal returns. In addition, the average hedge portfolio returns are larger in magnitude than those generated by several well-known return anomalies, and the return-predicting power of ISBP remains robust after controlling for these anomalies. These findings suggest that the information captured by ISBP is not merely a repackaging of previously documented predictors but instead reflects a distinct source of return-relevant information embedded in income statements.

By estimating SHAP values, we quantify the contribution of each income statement item in shaping ISBP and find that the relative importance rankings of these items do not vary to a large extent across firm size, book-to-market ratio, or product market attributes. In particular, sales revenue, costs of goods sold, and gross profits are the three most important income statement items. In addition, we find that ISBP is positively associated with both one-quarter-ahead earnings growth and the three-day abnormal return surrounding the subsequent earnings announcement, suggesting that ISBP captures information about future firm performance. The Mishkin (1983) test indicates that investors do not fully incorporate these earnings implications into prices, resulting in a price response that is weaker than implied by the earnings relation and giving rise to predictable post-filing returns.

Overall, this paper demonstrates the potential of AI to extract economically meaningful information from visualized income statement data. More generally, financial statements naturally possess grid-like structures and spatial hierarchies that make them well suited to image-based analysis. Extending our approach to balance sheets, cash flow statements, and other corporate disclosures may yield additional insights into the information content of financial reporting and represents a promising direction for future research.

Declaration of generative AI and AI-assisted technologies: During the preparation of this work, the authors used a generative AI large language model (Anthropic's Claude) to assist with proofreading, to flag typographical and grammatical errors, and to suggest improvements to wording and style. All such suggestions were reviewed and approved by the authors, who retain full editorial control and responsibility for the final text.

References

Abarbanell, J.S., Bushee, B.J., 1997. Fundamental analysis, future earnings, and stock prices. J. Account. Res. 35 (1), 1–24.

Abarbanell, J.S., Bushee, B.J., 1998. Abnormal returns to a fundamental analysis strategy. Account. Rev. 73 (1), 19–45.

Abel, A.B., Mishkin, F.S., 1983. On the econometric testing of rationality–market efficiency. Rev. Econ. Stat. 65, 318–323.

Akbas, F., Jiang, C., Koch, P.D., 2017. The trend in firm profitability and the cross-section of stock returns. Account. Rev. 92 (5), 1–32.

Amel-Zadeh, A., Calliess, J.P., Kaiser, D., Roberts, S., 2020. Machine learning-based financial statement analysis. Working Paper.

Arif, S., Marshall, N.T., Schroeder, J.H., Yohn, T.L., 2019. A growing disparity in earnings disclosure mechanisms: The rise of concurrently released earnings announcements and 10-Ks. J. Account. Econ. 68 (1), 101221.

Bali, T.G., Beckmeyer, H., Moerke, M., Weigert, F., 2023. Option return predictability with machine learning and big data. Rev. Financ. Stud. 36 (9), 3548–3602.

Ball, R., Bartov, E., 1996. How naive is the stock market's use of earnings information? J. Account. Econ. 21 (3), 319–337.

Ball, R., Brown, P., 1968. An empirical evaluation of accounting income numbers. J. Account. Res. 6 (2), 159–178.

Ball, R., Gerakos, J., Linnainmaa, J.T., Nikolaev, V., 2016. Accruals, cash flows, and operating profitability in the cross section of stock returns. J. Financ. Econ. 121 (1), 28–45.

Bernard, V.L., Thomas, J.K., 1989. Post-earnings-announcement drift: delayed price response or risk premium? J. Account. Res. 27, 1–36.

Bernard, V.L., Thomas, J.K., 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. J. Account. Econ. 13 (4), 305–340.

Bianchi, D., Büchner, M., Tamoni, A., 2021. Bond risk premiums with machine learning. Rev. Financ. Stud. 34 (2), 1046–1089.

Binz, O., Schipper, K., Standridge, K.R., 2025. Estimating profitability decomposition frameworks via machine learning: Implications for earnings forecasting and financial statement analysis. J. Account. Econ. 80 (2–3), 1–20.

Bowles, B., Reed, A.V., Ringgenberg, M.C., Thornock, J.R., 2024. Anomaly time. J. Finance 79 (5), 3543–3579.

Büchner, M., Kelly, B., 2022. A factor model for option returns. J. Financ. Econ. 143 (3), 1140–1161.

Cao, S., Cheng, Y., Wang, M., Xia, Y., Yang, B., 2022. Visual information and AI divide: Evidence from corporate executive presentations. Available at SSRN 4490834.

Cao, S., Jiang, W., Wang, J., Yang, B., 2024. From man vs. machine to man + machine: The art and AI of stock analyses. J. Financ. Econ. 160, 103910.

Cao, S.S., Narayanamoorthy, G.S., 2012. Earnings volatility, post-earnings announcement drift, and trading frictions. J. Account. Res. 50 (1), 41–74.

Carhart, M.M., 1997. On persistence in mutual fund performance. J. Finance 52 (1), 57–82.

Casey, R.J., Gao, F., Kirschenheiter, M.T., Li, S., Pandit, S., 2016. Do Compustat financial statement data articulate? J. Financ. Report. 1 (1), 37–59.

Chen, J.Z., Shane, P.B., 2014. Changes in cash: persistence and pricing implications. J. Account. Res. 52 (3), 599–634.

Chen, L., Pelger, M., Zhu, J., 2024. Deep learning in asset pricing. Manag. Sci. 70 (2), 714–750.

Chen, X., Cho, Y.H., Dou, Y., Lev, B., 2022. Predicting future earnings changes using machine learning and detailed financial data. J. Account. Res. 60 (2), 467–515.

Collins, D.W., Hribar, P., 2000. Earnings-based and accrual-based market anomalies: one effect or two? J. Account. Econ. 29 (1), 101–123.

Cooper, M.J., Gulen, H., Schill, M.J., 2008. Asset growth and the cross-section of stock returns. J. Finance 63 (4), 1609–1651.

Daniel, K., Hirshleifer, D., Sun, L., 2020. Short- and long-horizon behavioral factors. Rev. Financ. Stud. 33 (4), 1673–1736.

Dechow, P.M., Sloan, R.G., 1997. Returns to contrarian investment strategies: Tests of naive expectations hypotheses. J. Financ. Econ. 43 (1), 3–27.

DeMiguel, V., Gil-Bazo, J., Nogales, F.J., Santos, A.A., 2023. Machine learning and fund characteristics help to select mutual funds with positive alpha. J. Financ. Econ. 150 (3), 103737.

Engelberg, J., Sasseville, C., Williams, J., 2012. Market madness? The case of Mad money. Manag. Sci. 58 (2), 351–364.

Fama, E.F., French, K.R., 1992. The cross-section of expected stock returns. J. Finance 47 (2), 427–465.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. J. Financ. Econ. 33 (1), 3–56.

Fama, E.F., French, K.R., 2015. A five-factor asset pricing model. J. Financ. Econ. 116 (1), 1–22.

Fama, E.F., French, K.R., 2018. Choosing factors. J. Financ. Econ. 128 (2), 234–252.

Fama, E.F., MacBeth, J.D., 1973. Risk, return, and equilibrium: Empirical tests. J. Polit. Econ. 81 (3), 607–636.

Feng, G., Giglio, S., Xiu, D., 2020. Taming the factor zoo. J. Finance 75 (3), 1327–1370.

Foster, G., Olsen, C., Shevlin, T., 1984. Earnings releases, anomalies, and the behavior of security returns. Account. Rev. 59, 574–603.

Francis, J., LaFond, R., Olsson, P.M., Schipper, K., 2004. Costs of equity and earnings attributes. Account. Rev. 79 (4), 967–1010.

Frank, M.Z., Sanati, A., 2018. How does the stock market absorb shocks? J. Financ. Econ. 129 (1), 136–153.

Freyberger, J., Neuhierl, A., Weber, M., 2020. Dissecting characteristics nonparametrically. Rev. Financ. Stud. 33 (5), 2326–2377.

Frésard, L., Hoberg, G., Phillips, G.M., 2020. Innovation activities and integration through vertical acquisitions. Rev. Financ. Stud. 33 (7), 2937–2976.

Garfinkel, J.A., Sokobin, J., 2006. Volume, opinion divergence, and returns: A study of post-earnings announcement drift. J. Account. Res. 44 (1), 85–112.

Geertsema, P., Lu, H., 2023. Relative valuation with machine learning. J. Account. Res. 61 (1), 329–376.

Gu, M., Hirshleifer, D., Teoh, S.H., Wu, S., 2025. GIFfluence: A visual approach to investor sentiment and the stock market. arXiv preprint arXiv:2512.20027.

Gu, S., Kelly, B., Xiu, D., 2020. Empirical asset pricing via machine learning. Rev. Financ. Stud. 33 (5), 2223–2273.

Gu, S., Kelly, B., Xiu, D., 2021. Autoencoder asset pricing models. J. Econometrics 222 (1), 429–450.

He, S., Narayanamoorthy, G.G., 2020. Earnings acceleration and stock returns. J. Account. Econ. 69 (1), 101238.

Hoberg, G., Phillips, G., 2016. Text-based network industries and endogenous product differentiation. J. Polit. Econ. 124 (5), 1423–1465.

Hoberg, G., Phillips, G., Prabhala, N., 2014. Product market threats, payouts, and financial flexibility. J. Finance 69 (1), 293–324.

Hoberg, G., Phillips, G.M., 2025. Scope, scale and competition: The 21st-century firm. J.

Finance 80, 415–466.

Holthausen, R.W., Larcker, D.F., 1992. The prediction of stock returns using financial statement information. J. Account. Econ. 15 (2–3), 373–411.

Hou, K., Mo, H., Xue, C., Zhang, L., 2021. An augmented q-factor model with expected growth. Rev. Finance 25 (1), 1–41.

Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: An investment approach. Rev. Financ. Stud. 28 (3), 650–705.

Hou, K., Xue, C., Zhang, L., 2020. Replicating anomalies. Rev. Financ. Stud. 33 (5), 2019–2133.

Hribar, P., Collins, D.W., 2002. Errors in estimating accruals: Implications for empirical research. J. Account. Res. 40 (1), 105–134.

Hu, A., Ma, S., 2025. Persuading investors: A video-based study. J. Finance 80 (5), 2639–2688.

Hui, K.W., Nelson, K.K., Yeung, P.E., 2016. On the persistence and pricing of industry-wide and firm-specific earnings, cash flows, and accruals. J. Account. Econ. 61 (1), 185–202.

Jegadeesh, N., Livnat, J., 2006a. Post-earnings-announcement drift: The role of revenue surprises. Financ. Anal. J. 62 (2), 22–34.

Jegadeesh, N., Livnat, J., 2006b. Revenue surprises and stock returns. J. Account. Econ. 41 (1–2), 147–171.

Jensen, T.I., Kelly, B., Pedersen, L.H., 2023. Is there a replication crisis in finance? J. Finance 78 (5), 2465–2518.

Jiang, J., Kelly, B., Xiu, D., 2023. (Re-)Imag(in)ing price trends. J. Finance 78 (6), 3193–3249.

Kaniel, R., Lin, Z., Pelger, M., Van Nieuwerburgh, S., 2023. Machine-learning the skill of mutual fund managers. J. Financ. Econ. 150 (1), 94–138.

Kapadia, N., Zekhnini, M., 2019. Do idiosyncratic jumps matter? J. Financ. Econ. 131 (3), 666–692.

Kelly, B., Palhares, D., Pruitt, S., 2023. Modeling corporate bond returns. J. Finance 78 (4), 1967–2008.

Kelly, B.T., Pruitt, S., Su, Y., 2019. Characteristics are covariances: A unified model of risk and return. J. Financ. Econ. 134 (3), 501–524.

Kim, A.G., Muhn, M., Nikolaev, V.V., 2024. Financial statement analysis with large language models. Working Paper.

Kozak, S., Nagel, S., Santosh, S., 2020. Shrinking the cross-section. J. Financ. Econ. 135 (2), 271–292.

Kraft, A., Leone, A.J., Wasley, C.E., 2007. Regression-based tests of the market pricing of accounting numbers: The Mishkin test and ordinary least squares. J. Account. Res. 45 (5), 1081–1114.

LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D., 1989. Backpropagation applied to handwritten ZIP code recognition. Neural Comput. 1 (4), 541–551.

Leippold, M., Wang, Q., Zhou, W., 2022. Machine learning in the Chinese stock market. J. Financ. Econ. 145 (2), 64–82.

Lev, B., Thiagarajan, S.R., 1993. Fundamental information analysis. J. Account. Res. 31 (2), 190–215.

Li, B., Rossi, A.G., Yan, X.S., Zheng, L., 2025. Machine learning from a "Universe" of signals: The role of feature engineering. J. Financ. Econ. 172, 104138.

Liu, J., Stambaugh, R.F., Yuan, Y., 2018. Absolving beta of volatility's effects. J. Financ. Econ. 128 (1), 1–15.

Livnat, J., Mendenhall, R.R., 2006. Comparing the post-earnings announcement drift for surprises calculated from analyst and time series forecasts. J. Account. Res. 44 (1), 177–205.

Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst. 30, 4765–4774.

Ma, G., Markov, S., 2017. The market's assessment of the probability of meeting or beating the consensus. Contemp. Account. Res. 34 (1), 314–342.

Mishkin, F., 1983. A rational expectations approach to macroeconomics: Testing policy effectiveness and efficient markets models. University of Chicago Press.

Mohanram, P.S., 2005. Separating winners from losers among low book-to-market stocks using financial statement analysis. Rev. Account. Stud. 10 (2–3), 133–170.

Moss, A., 2022. How do brokerages' digital engagement practices affect retail investor information processing and trading? Ph.D. thesis, University of Iowa.

Murray, S., Xia, Y., Xiao, H., 2024. Charting by machines. FwJ. Financ. Econ. 153, 103791.

Narayanamoorthy, G., 2006. Conservatism and cross-sectional variation in the post-earnings announcement drift. J. Account. Res. 44 (4), 763–789.

Nekrasov, A., Teoh, S.H., Wu, S., 2022. Visuals and attention to earnings news on Twitter. Rev. Account. Stud. 27 (4), 1233–1275.

Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55 (3), 703–708.

Novy-Marx, R., 2013. The other side of value: The gross profitability premium. J. Financ. Econ. 108 (1), 1–28.

Obaid, K., Pukthuanthong, K., 2022. A picture is worth a thousand words: Measuring investor sentiment by combining machine learning and photos from news. J. Financ. Econ. 144 (1), 273–297.

Ou, J.A., Penman, S.H., 1989. Financial statement analysis and the prediction of stock returns. J. Account. Econ. 11 (4), 295–329.

Piotroski, J.D., 2000. Value investing: The use of historical financial statement information to separate winners from losers. J. Account. Res. 38, 1–41.

Rangan, S., Sloan, R.G., 1998. Implications of the integral approach to quarterly reporting for the post-earnings-announcement drift. Account. Rev. 73, 353–371.

Rapach, D.E., Strauss, J.K., Zhou, G., 2013. International stock return predictability: What is the role of the United States? J. Finance 68 (4), 1633–1662.

Richardson, S.A., Sloan, R.G., Soliman, M.T., Tuna, I., 2005. Accrual reliability, earnings persistence and stock prices. J. Account. Econ. 39 (3), 437–485.

Savor, P.G., 2012. Stock returns after major price shocks: The impact of information. J. Financ. Econ. 106 (3), 635–659.

Sloan, R.G., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? Account. Rev. 71 (3), 289–315.

Thomas, J., Zhang, F.X., 2011. Tax expense momentum. J. Account. Res. 49 (3), 791–821.

Vega, C., 2006. Stock price reaction to public and private information. J. Financ. Econ. 82 (1), 103–133.

Wu, W., Chen, J., Yang, Z., Tindall, M.L., 2021. A cross-sectional machine learning approach for hedge fund return prediction and selection. Manag. Sci. 67 (7), 4577–4601.

Appendix. Variable Definitions. This table summarizes variable definitions. Compustat quarterly items are colored in blue.

Variables	Descriptions
SALES	Quarterly sales revenue. We use quarterly net sales (SALEQ). If SALEQ is unavailable, we use quarterly revenue (REVTQ).
COGS	Quarterly cost of goods sold. We use quarterly costs of good sold (COGSQ). If COGSQ is unavailable, we use quarterly total operating expenses (XOPRQ) minus quarterly selling, general and administrative expenses (XSGAQ, zero if missing).
GP	Quarterly gross profits. We use SALES minus COGS.
OPEXP	Quarterly operating expenses. We use quarterly selling, general and administrative expenses (XSGAQ, zero if missing) plus quarterly depreciation and amortization expenses (DPQ, zero if missing).
OPINC	Quarterly operating income (after depreciation). We use GP minus OPEXP. The comparable item in Compustat is OIADPQ.
INT	Quarterly interest income/expenses. We use quarterly non-operating income/expenses (NOPIQ, zero if missing), minus quarterly interest and related expenses (XINTQ, zero if missing), plus quarterly special items (SPIQ, zero if missing).
PTINC	Quarterly pretax income. We use OPINC plus INT. The comparable item in Compustat is PIQ.
TAXMI	Quarterly income taxes and minority interests. We use quarterly total income taxes (TXTQ, zero if missing), plus quarterly minority interests (MIIQ, zero if missing).
IBEI	Quarterly income before extraordinary items. We use PTINC minus TAXMI. The comparable item in Compustat is IBQ.
SUE	Standardized unexpected earnings. We use the change the change in split-adjusted quarterly earnings per share ($\frac{\text{EPSPXQ}}{\text{AJEXQ}}$) from its value four quarters ago divided by the standard deviation of this change over the prior eight quarters (six quarters minimum). SUE also serves as the earnings growth proxy.
RS	Revenue surprise. We use the change in split-adjusted quarterly revenue per share ($\frac{\text{SALEQ}}{\text{CSHPRQ}\times\text{AJEXQ}}$) from its value four quarters ago divided by the standard deviation of this change over the prior eight quarters (six quarters minimum).

Variables	Descriptions
TES	Tax expense surprise. We use the change in split-adjusted tax expenses per share divided by total assets per share four quarters prior. The change in split-adjusted tax expense per share is computed as ($\frac{\text{TXTQ}}{\text{CSHPRQ}\times\text{AJEXQ}}$) in quarter q minus its value four quarters prior. Total assets per share is computed as ($\frac{\text{ATQ}}{\text{CSHPRQ}\times\text{AJEXQ}}$).
EA	Earnings acceleration. For firm i in quarter q, we use $$\frac{\text{EPS}_{i,q}-\text{EPS}_{i,q-4}}{\text{Stock Price}_{i,q-1}}-\frac{\text{EPS}_{i,q-1}-\text{EPS}_{i,q-5}}{\text{Stock Price}_{i,q-2}}.$$ $\text{EPS}_{i,q}$ is earnings per share for firm i in quarter q. Shares are adjusted for stock splits.
TREND	Trend in gross profitability. For firm i in quarter q, we use $\beta_{i,q}$ estimated from the following time-series regression: $$GPQ_{i,q}=\alpha_{i,q}+\beta_{i,q}t+\lambda_{1,i,q}D1+\lambda_{2,i,q}D2+\lambda_{3,i,q}D3+\epsilon_{i,q}\ ,$$ where $t=1,2,\ldots,8$ and represents a deterministic time trend covering quarter $q-7$ through q, and D1 to D3 represent quarterly dummy variables. GPQ is computed as GP divided by total assets (ATQ).
RET[-1, +1]	Filing return. We use the buy-and-hold market-adjusted return of a firm during the [-1, +1] window around its 10-K/10-Q filing date.
RET[-30,-2]	Pre-filing return. We use the buy-and-hold market-adjusted return of a firm during the [-30, -2] window around its 10-K/10-Q filing date.
PERSIST	Earnings persistence. For firm i in quarter q, we use $\beta_{i,q}$ estimated from the following time-series regression: $$\text{EARNINGS}_{i,q}=\alpha_{i,q}+\beta_{i,q}\text{EARNINGS}_{i,q-1}+\epsilon_{i,q}\ ,$$ with the most recent eight quarters (quarter $q-7$ to q) of earnings (IBQ).
VOL	Earnings volatility. We use the standard deviation of earnings (IBQ) in the most recent eight quarters (quarter q-7 to q).
SIZE	Firm size. We use market capitalization (From CRSP) of a firm on its 10-K/10-Q filing date.
BM	Book-to-market ratio. Book equity is shareholders' equity, plus balance sheet deferred taxes and investment tax credit (TXDITCQ, zero if missing), minus the book value of preferred stock (PSTKQ, zero if missing). Depending on availability, we use stockholders' equity (SEQQ), or common equity (CEQQ) plus the book value of preferred stock (PSTKQ, zero if missing), or total assets (ATQ) minus total liabilities (LTQ) in that order as shareholders' equity. BM is defined as the book equity divided by the firm's market capitalization (from CRSP) on its 10-K/10-Q filing date.

Variables	Descriptions
GPA	Gross profitability. We use the sum of GP in the most recent four quarters (quarters q-3 to q), divided by total assets (ATQ).
OPA	Operating profitability. We use the sum of operating profits in the most recent four quarters (quarters q-3 to q), divided by total assets (ATQ) four quarters prior. We compute operating profits as SALES minus COGS, minus selling, general and administrative expenses (XSGAQ, zero if missing), plus research and development expenses (XRDQ, zero if missing).
TA	Total accruals. We use the cash flow approach to measure TA as the sum of net income (NIQ) in the most recent four quarters (quarter q-3 to q) minus total operating, investing, and financing cash flows (OANCFY, IVNCFY, and FINCFY) plus sales of stocks (SSTKY, zero if missing) minus stock repurchases and dividends (PRSTKCY and DVY, zero if missing), divided by total assets (ATQ) four quarters prior.
OA	Operating accruals. We use the cash flow approach to measure OA as the sum of net income (NIQ) in the most recent four quarters (quarter q-3 to q) minus total operating cash flow (OANCFY), divided by total assets (ATQ) four quarters prior.
AG	Asset growth. We use the change in total assets (ATQ) from its value four quarters prior, divided by total assets four quarters prior.
SAR	Size-adjusted buy-and-hold return. We use the difference between the buy-and-hold return of a firm and that of its size-matched portfolio over the 63-day window ([+2, +64]) following its 10-K/10-Q filing date. We use the monthly NYSE size decile breakpoints at the end of June in year t to determine the size-matched portfolio for a firm whose 10-K/10-Q filing date is between July of year t and June of year t+1.
MAR	Market-adjusted buy-and-hold-return. We use the difference between the buy-and-hold return of affirm and that of the CRSP value-weighted market portfolio over the 63-day window [+2, +64] following its 10-K/10-Q filing date.
FF4	Fama-French three-factor and momentum-adjusted buy-and-hold return during the 63-day window ([+2, +64]) following earnings announcement date, with factor loadings estimated using the 120-day window ([-150, -31], 90 days minimum) prior to the earnings announcement date. The factors are market, size, value, and momentum.
FF6	Fama-French five-factor and momentum-adjusted buy-and-hold return during the 63-day window ([+2, +64]) following 10-K/10-Q filing date, with factor loadings estimated using the 120-day window ([-150, -31], 90 days minimum) prior to the 10-K/10-Q filing date. The factors are market, size, value, operating profitability, investment, and momentum.
HMXZ5	q^5-factor-adjusted buy-and-hold return during the 63-day window ([+2, +64]) following 10-K/10-Q filing date, with factor loadings estimated using the 120-day window ([-150, -31], 90 days minimum) prior to the 10-K/10-Q filing date. The factors are market, size, investment, return on equity, and expected growth.
DHS3	Behavioral-factor-adjusted buy-and-hold return during the 63-day window ([+2, +64]) following 10-K/10-Q filing date, with factor loadings estimated using the 120-day window ([-150, -31], 90 days minimum) prior to the 10-K/10-Q filing date. The factors are market, financing, and post earnings announcement drift.

	q	q-1	q-2	q-3	q-4	q-5	q-6	q-7	q-8
SALES	0.425	0.490	0.448	0.806	0.553	0.493	0.444	0.425	0.536
COGS	0.219	0.248	0.190	0.441	0.281	0.241	0.384	0.208	0.272
GP	0.206	0.242	0.258	0.365	0.271	0.251	0.061	0.218	0.263
OPEXP	0.176	0.183	0.197	0.291	0.202	0.189	0.036	0.186	0.231
OPINC	0.030	0.059	0.061	0.073	0.070	0.063	0.024	0.032	0.033
INT	0.042	0.002	0.001	-0.002	0.004	-0.003	-0.002	-0.003	-0.004
PTINC	0.072	0.062	0.063	0.072	0.073	0.060	0.023	0.029	0.029
TAXMI	0.029	0.024	0.006	0.027	0.027	0.022	0.008	0.012	0.012
IBEI	0.044	0.037	0.057	0.045	0.047	0.038	0.014	0.016	0.017

Figure 1. An Example Income Statement Image. This figure presents an example income statement image, with rows specifying nine income statement items and columns specifying the most recent nine quarters (quarters q-8 to q). The nine income statement items include sales revenue (SALES), cost of goods sold (COGS), gross profits (GP), operating expenses (OPEXP), operating income (OPINC), interest income/expenses (INT), pretax income (PTINC), taxes and minority interests (TAXMI), and income before extraordinary items (IBEI).

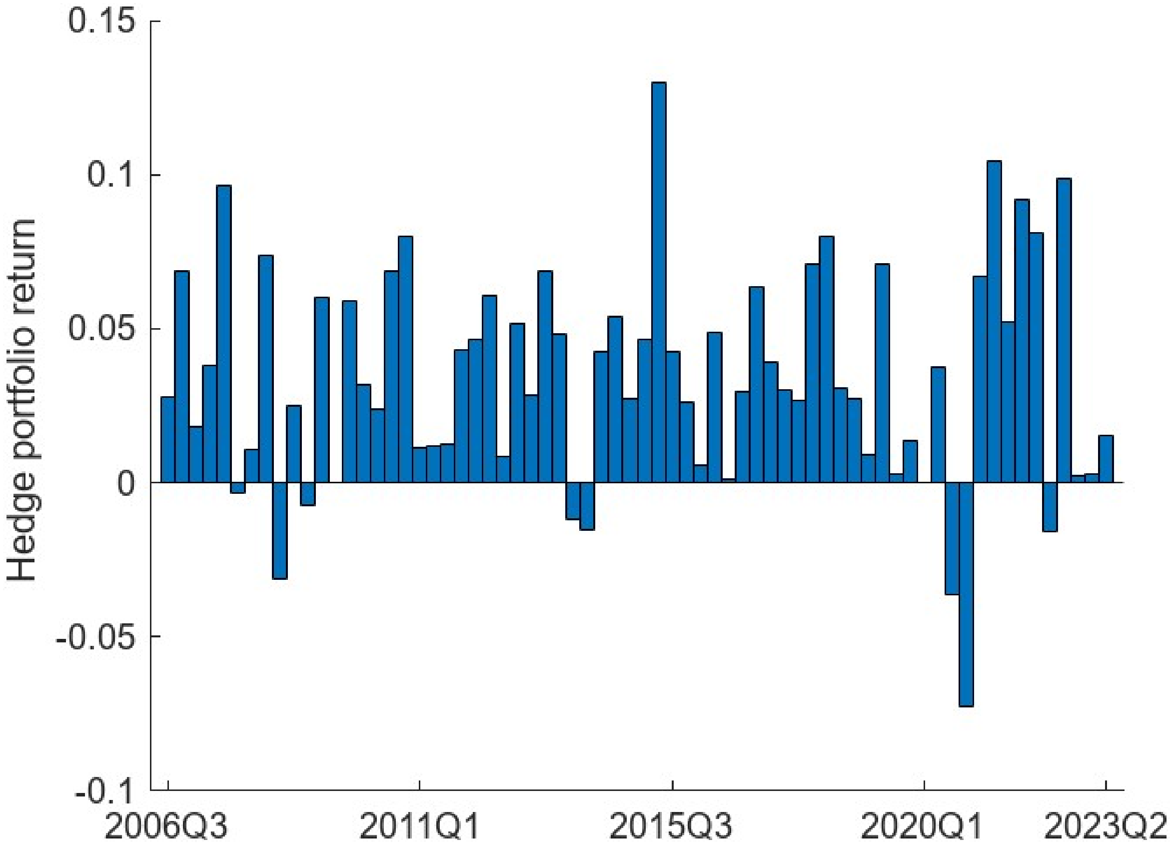

Figure 2. The Return Predictability of ISBP Over Time. This figure depicts the difference in the average 63-day size-adjusted buy-and-hold returns (SAR) between the highest decile and the lowest decile of income statement buy probability (ISBP) in each quarter during the out-of-sample period (2006Q3 to 2023Q2). The decile cutoffs are based on the distribution of the previous quarter's ISBP. See the Appendix for variable definitions.

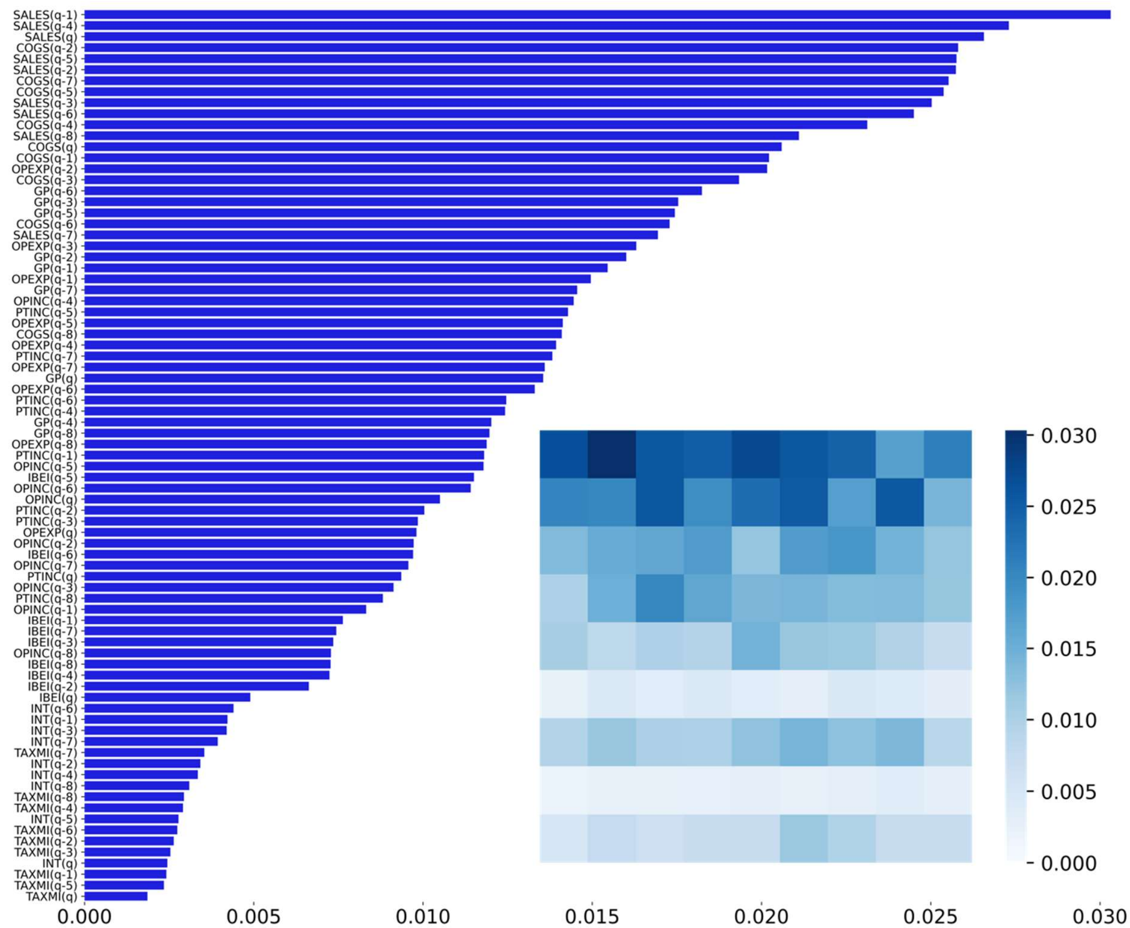

Figure 3. Importance of Quarterly income Statement Items. The bar chart and heatmap show the importance of 81 quarterly income statement items. The importance of a quarterly income statement item is measured by the mean absolute SHAP value of the item across all income statement images in the out-of-sample period (2006Q3-2023Q2). Importance values are normalized so that the total importance across all items sums to one. The 81 quarterly income statement items are sales revenue (SALES), cost of goods sold (COGS), gross profit (GP), operating expenses (OPEXP), operating income (OPINC), interest income and expenses (INT), pretax income (PTINC), taxes and minority interests (TAXMI), and income before extraordinary items (IBEI) in the most recent nine quarters.

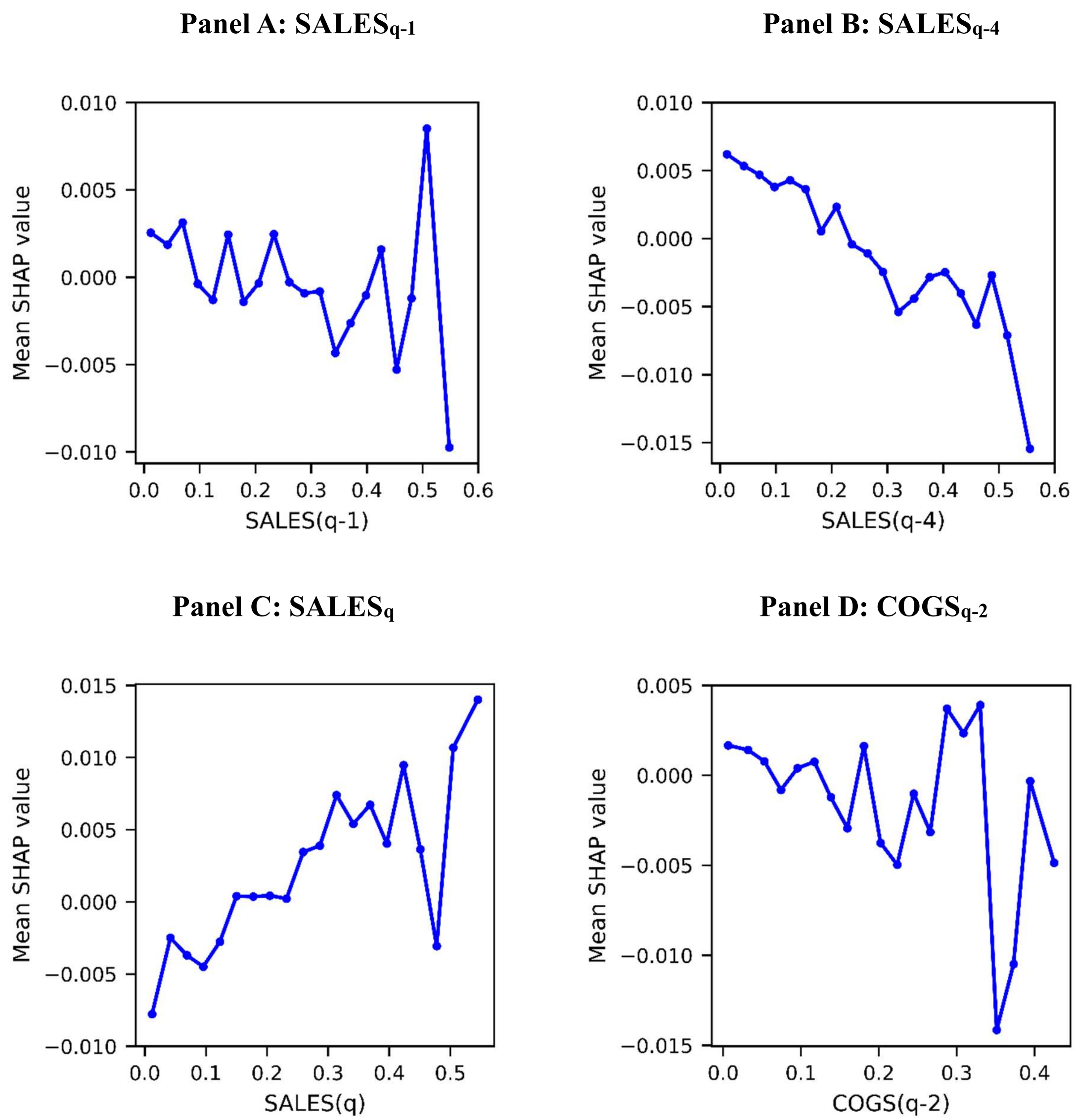

Figure 4. Nonlinearity in the Relation Between Quarterly Income Statement Items and ISBP. This figure presents SHAP plots for four characteristics (i.e., quarterly income statement items) used in CNN's income statement buy probability (ISBP) predictions: sales revenue in quarter $q-1$ (panel A), sales revenue in quarter $q-4$ (panel B), sales revenue in quarter q (panel C), and cost of goods sold in quarter $q-2$ (panel D). Each characteristic is winsorized at the 1st and 95th percentiles. The observations in the out-of-sample (2006Q3-2023Q2) are then sorted into 20 groups based on the value of the characteristic. For each group, the conditional mean SHAP value (i.e., the average SHAP value across all observations within the group) is plotted on the y-axis against the group's mean characteristic value on the x-axis.

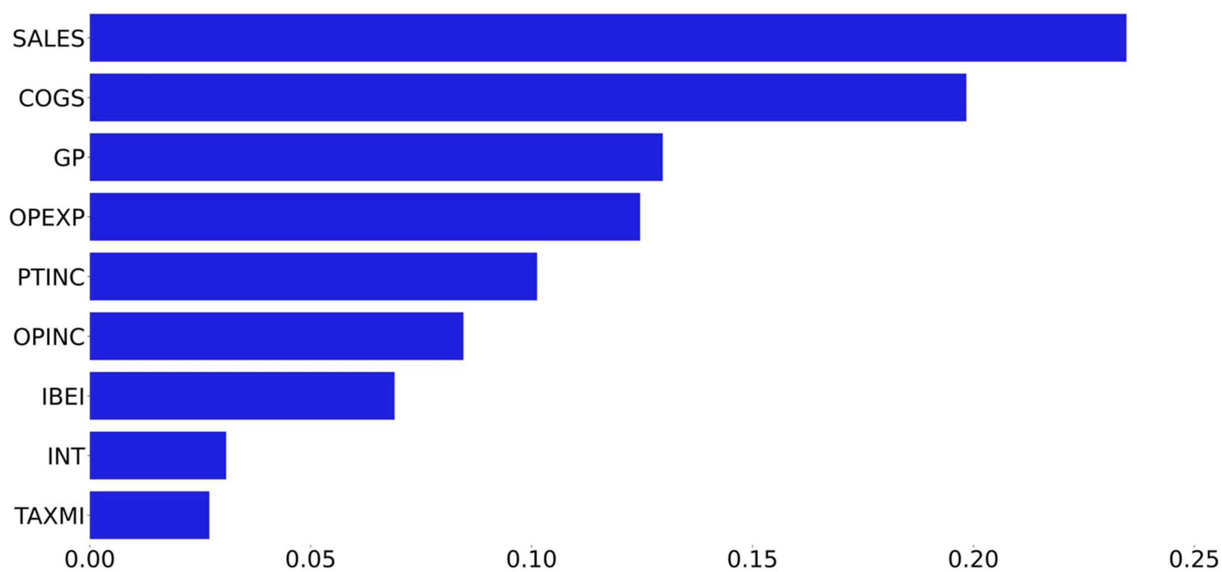

Figure 5. Importance of Income Statement Items. This figure presents the importance of nine characteristic groups in predicting the income statement buy probability (ISBP). The nine characteristic groups correspond to the nine income statement items: sales revenue (SALES), cost of goods sold (COGS), gross profit (GP), operating expenses (OPEXP), operating income (OPINC), interest income and expenses (INT), pretax income (PTINC), taxes and minority interests (TAXMI), and income before extraordinary items (IBEI). The importance of a characteristic group is computed as the sum of the importances of its constituent characteristics (i.e., quarterly income statement items in the most recent nine quarters). The importance of an individual characteristic is measured by its mean absolute SHAP value across all income statement images in the out-of-sample period (2006Q3-2023Q2). Group importance values are then normalized so that the total importance across all characteristic groups sums to one.

Table 1.

Sample Selection

All Compustat firm-quarters with matched CRSP Permno (SHRCD = 10 or 11; EXCHCD = 1, 2 or 3) whose earnings announcement date (Compustat item RDQ) is between 1996/01/01 and 2023/06/30	556,937
Drop observations whose current quarter RDQ is before the quarter fiscal period end date	(117)
Drop observations with missing or non-positive book equity in the most recent nine quarters (quarters q-8 to q)	(116,549)
Drop observations with missing or non-positive total assets in the most recent ten quarters (quarters q-9 to q)	(8,871)
Drop observations with missing SALES, COGS, GP, OPEXP, OPINC, INT, PTINC, TAXMI, or IBEI in the most recent nine quarters	(15,030)
Drop observations with missing IBQ or IBEI not equal to IBQ in the most recent nine quarters	(13,689)
Drop observations with missing historical CIK codes or no corresponding 10-K/10-Q filing dates in EDGAR	(102,979)
Drop observations whose current quarter 10-K/10-Q filing date precedes the earnings announcement date	(3,012)
Remove late filers in the current quarter (i.e., *Filelag* > 90 days for 10-Ks or *Filelag* > 45 days for 10-Qs)	(15,777)
Drop observations whose CRSP daily price at the current quarter 10-K/10-Q filing date is missing or ≤$1	(38,516)
Drop observations with fewer than 30 trading days in the 120-day window ([-150, -31]) prior to the current quarter 10-K/10-Q filing date	(406)
Drop observations with missing firm size at the previous fiscal year end	(293)
Total observations	241,698
In-sample dataset: observations whose current quarter 10-K/10-Q filing date is between 1996Q1 and 2005Q4	90,544
Out-of-sample dataset: observations whose current quarter 10-K/10-Q filing date is between 2006Q3 and 2023Q2 and has non-missing SUE, RS, TES, EA, TREND, RET[-1, +1], RET[-30, -2], PERSIST, VOL, SIZE, BM, GPA, OPA, OA, TA, or AG	144,131

This table reports the sample selection procedures. Filelag is the number of days between the fiscal year (quarter) end and the 10-K (10-Q) filing date according to Edgar. See Appendix for variable definitions.

Table 2.
Income Statement Buy Probability and Post-Filing Return: Univariate Portfolio Analysis.

ISBP deciles	SAR	MAR	FF4	FF6	HMXZ5	DHS3
Low [28.2%]	−0.015***	−0.007	−0.008*	−0.007	−0.002	−0.005
	(−4.560)	(−0.765)	(−1.827)	(−1.660)	(−0.377)	(−0.712)
2 [30.4%]	−0.007	0.000	0.004	0.005	−0.002	−0.002
	(−1.303)	(0.014)	(0.966)	(1.048)	(−0.457)	(−0.302)
3 [30.8%]	−0.011	−0.005	0.004	0.006	−0.001	−0.001
	(−1.638)	(−0.714)	(1.049)	(1.404)	(−0.212)	(−0.213)
4 [31.3%]	−0.015**	−0.006	−0.004	−0.002	−0.007	−0.006
	(−2.383)	(−1.316)	(−0.770)	(−0.521)	(−1.266)	(−1.082)
5 [32.0%]	−0.010***	−0.002	−0.002	−0.002	−0.002	−0.003
	(−3.916)	(−0.540)	(−0.910)	(−0.776)	(−0.844)	(−0.905)
6 [32.9%]	0.001	0.007	0.004	0.004	0.006*	0.008
	(0.245)	(1.155)	(1.233)	(1.202)	(1.775)	(1.482)
7 [33.9%]	0.003	0.010	0.007**	0.007**	0.009***	0.010*
	(1.063)	(1.522)	(2.060)	(2.123)	(3.731)	(1.938)
8 [35.0%]	−0.001	0.006	0.004	0.006*	0.008***	0.006
	(−0.327)	(0.912)	(1.441)	(1.949)	(2.860)	(1.103)
9 [36.5%]	0.005	0.012*	0.008*	0.010**	0.012***	0.014**
	(1.474)	(1.729)	(1.798)	(2.519)	(2.820)	(2.372)
High [39.9%]	0.020***	0.027***	0.021***	0.024***	0.026***	0.029***
	(7.077)	(3.990)	(4.619)	(5.901)	(6.986)	(4.796)
High-Low	0.035***	0.034***	0.029***	0.030***	0.028***	0.034***
[11.6%]	(8.046)	(6.946)	(5.287)	(6.341)	(5.224)	(5.984)

This table reports the average 63-day buy-and-hold abnormal returns (BHARs) following 10-K/10-Q filings for portfolios sorted into deciles based on income statement buy probability (ISBP) during the out-of-sample period (2006Q3-2023Q2). BHARs are measured using size-adjusted returns (SAR), market-adjusted returns (MAR), factor-adjusted returns (FF4, FF6, HMXZ5, and DHS3). Decile cutoffs are determined using the previous quarter's distribution of ISBP. The average ISBP within each decile is reported in brackets. The final row reports the average hedge portfolio returns, calculated as the difference in BHARs between the highest and lowest ISBP deciles. See Appendix for variable definitions. Newey and West (1987) *t*-statistics with three lags are reported in parentheses, and ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Table 3.

Comparative Evaluation of Post-Filing Return Predictability

Variable	SAR	MAR	FF4	FF6	HMXZ5	DHS3
ISBP	0.035***	0.034***	0.029***	0.030***	0.028***	0.034***
	(8.046)	(6.946)	(5.287)	(6.341)	(5.224)	(5.984)
SUE	0.017***	0.017***	0.014***	0.015***	0.014***	0.018***
	(2.838)	(2.699)	(3.545)	(3.382)	(2.993)	(3.255)
RS	0.014**	0.016**	0.016***	0.017***	0.015***	0.017***
	(2.502)	(2.496)	(5.017)	(4.834)	(4.041)	(3.453)
TES	0.011*	0.011	0.009*	0.010**	0.011**	0.012**
	(1.697)	(1.628)	(1.938)	(2.242)	(2.191)	(2.052)
EA	0.013**	0.012**	0.012***	0.014***	0.010	0.015***
	(2.248)	(2.052)	(2.691)	(3.581)	(1.547)	(2.752)
TREND	−0.002	−0.001	−0.001	−0.001	0.000	0.000
	(−0.343)	(−0.157)	(−0.195)	(−0.278)	(−0.019)	(−0.114)
RET[-1, +1]	0.009	0.007	0.014***	0.014***	0.014**	0.011
	(1.276)	(0.963)	(2.749)	(2.860)	(2.528)	(1.538)
RET[-30, -2]	−0.002	−0.006	0.001	0.004	0.002	−0.003
	(−0.154)	(−0.475)	(0.105)	(0.553)	(0.190)	(−0.324)
PERSIST	0.003	0.002	−0.001	0.000	−0.001	0.000
	(0.813)	(0.685)	(−0.288)	(0.096)	(−0.327)	(0.041)
VOL	0.009	0.010	−0.009	−0.010	−0.009	−0.008
	(0.755)	(1.174)	(−1.018)	(−1.244)	(−1.008)	(−0.827)
ME	−0.007	−0.008	−0.022	−0.023	−0.027*	−0.026
	(−0.971)	(−0.505)	(−1.284)	(−1.468)	(−1.751)	(−1.368)
BM	0.003	0.005	0.013	0.009	0.007	0.008
	(0.256)	(0.341)	(1.349)	(0.864)	(0.695)	(0.593)
GPA	0.025***	0.025***	0.026***	0.021***	0.016**	0.023***
	(3.370)	(3.053)	(4.140)	(3.024)	(2.044)	(2.897)
OPA	0.020**	0.020	0.019**	0.013	0.007	0.015
	(2.047)	(1.504)	(2.138)	(1.666)	(0.702)	(1.164)
OA	−0.012	−0.013	−0.005	−0.008	−0.018***	−0.013
	(−1.319)	(−1.143)	(−0.680)	(−1.422)	(−3.062)	(−1.505)
TA	−0.007	−0.008	−0.002	−0.008*	−0.012***	−0.010
	(−0.989)	(−0.819)	(−0.362)	(−1.874)	(−2.681)	(−1.305)
AG	−0.012	−0.013	−0.010	−0.010	−0.010	−0.012
	(−1.253)	(−1.107)	(−1.027)	(−1.216)	(−1.357)	(−1.174)

This table reports the average hedge portfolio returns, defined as the difference in 63-day buy-and-hold abnormal returns (BHARs) following 10-K/10-Q filings between firms in the highest and lowest deciles of each variable, over the out-of-sample period (2006Q3-2023Q2). BHARs are measured using size-adjusted returns (SAR), market-adjusted returns (MAR), and factor-adjusted returns (FF4, FF6, HMXZ5, and DHS3). The variables include the income statement buy probability (ISBP), standardized unexpected earnings (SUE), revenue surprise (RS), tax expense surprise (TES), earnings acceleration (EA), trend in gross profitability (TREND), firm size (SIZE), book-to-market ratio (BM), filing return (RET[−1, +1]), pre-filing return (RET[−30, −2]), earnings persistence (PERSIST), earnings volatility (VOL), gross profitability (GPA), operating profitability (OPA), total accruals (TA), operating accruals (OA), and asset growth (AG). For each variable, decile cutoffs are determined using its distribution in the previous quarter. Newey and West (1987) *t*-statistics with three lags are reported in parentheses, and ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Table 4.

Income Statement Buy Probability and Post-Filing Return: Regression Analysis

	(1) SAR	(2) MAR	(3) FF4	(4) FF6	(5) HMXZ5	(6) DHS3
Intercept	0.004	0.010	0.008	0.009	0.018*	0.017
	(0.592)	(0.605)	(0.558)	(0.678)	(1.693)	(1.053)
ISBP	0.013***	0.014***	0.011**	0.013***	0.014***	0.016***
	(3.176)	(3.386)	(2.355)	(2.812)	(2.849)	(3.729)
SUE	0.013***	0.013***	0.006*	0.007*	0.010**	0.011**
	(3.483)	(3.342)	(1.688)	(1.928)	(2.194)	(2.564)
RS	0.011***	0.012***	0.012***	0.011***	0.010***	0.010***
	(3.863)	(4.247)	(4.745)	(3.846)	(3.783)	(4.073)
TES	0.002	0.002	0.002	0.002	0.002	0.002
	(0.838)	(0.786)	(0.955)	(0.988)	(0.566)	(0.597)
EA	0.004*	0.004*	0.004*	0.005**	0.003	0.004*
	(1.745)	(1.962)	(1.713)	(2.104)	(1.035)	(1.831)
TREND	-0.007***	-0.007***	-0.006**	-0.006**	-0.006**	-0.007***
	(-2.985)	(-3.025)	(-2.184)	(-2.092)	(-2.134)	(-2.730)
RET[-1, +1]	0.000	-0.002	0.005	0.006	0.005	0.002
	(0.072)	(-0.450)	(1.573)	(1.581)	(1.456)	(0.475)
RET[-30, -2]	-0.005	-0.007	-0.000	0.002	-0.002	-0.004
	(-0.676)	(-1.001)	(-0.078)	(0.432)	(-0.325)	(-0.647)
PERSIST	0.001	0.002	-0.001	-0.000	-0.000	0.000
	(0.603)	(0.611)	(-0.261)	(-0.132)	(-0.043)	(0.166)
VOL	0.007	0.006	-0.009	-0.011	0.002	0.001
	(0.392)	(0.392)	(-0.787)	(-1.042)	(0.183)	(0.060)
SIZE	-0.001	-0.001	-0.001	-0.001	-0.002	-0.002
	(-1.058)	(-0.562)	(-0.392)	(-0.389)	(-1.432)	(-1.035)
BM	0.013	0.014	0.023***	0.018***	0.011	0.013*
	(1.204)	(1.306)	(5.196)	(3.647)	(1.331)	(1.764)
GPA	0.014	0.013	0.013	0.010	0.007	0.012
	(1.380)	(1.338)	(1.039)	(0.862)	(0.660)	(1.074)
OPA	0.018	0.019	0.015	0.011	0.016	0.016
	(1.508)	(1.539)	(1.425)	(1.143)	(1.405)	(1.477)
OA	-0.011*	-0.011*	-0.003	-0.003	-0.013***	-0.010**
	(-1.962)	(-1.887)	(-0.943)	(-0.862)	(-2.862)	(-2.321)
TA	-0.006	-0.007	-0.007	-0.009**	-0.005	-0.007*
	(-1.430)	(-1.507)	(-1.549)	(-2.014)	(-1.260)	(-1.731)
AG	-0.011**	-0.011**	-0.009***	-0.008***	-0.007**	-0.008*
	(-2.018)	(-2.002)	(-3.086)	(-2.691)	(-2.086)	(-1.965)
Adj. R^2	0.046	0.051	0.032	0.028	0.033	0.039
obs.	144131	144131	144131	144131	144131	144131

The table presents results of Fama and MacBeth (1973) regressions in the out-of-sample period (2006Q3-2023Q2) using the 63-day buy-and-hold abnormal return after 10-K/10-Q filings, including size-adjusted return (SAR), market-adjusted return (MAR), and factor-adjusted returns (FF4, FF6, HMXZ5, and DHS3), as the dependent variables. The independent variables include the income statement buy probability (ISBP), standardized unexpected earnings (SUE), revenue surprise (RS), tax expense surprise (TES), earnings acceleration (EA), trend in gross profitability (TREND), firm size (SIZE), book-to-market ratio (BM), filing return (RET[−1, +1]), pre-filing return (RET[−30, −2]), earnings persistence (PERSIST), earnings volatility (VOL), gross profitability (GPA), operating profitability (OPA), total accruals (TA), operating accruals (OA), and asset growth (AG). All variables except for the six measures of buy-and-hold abnormal return are converted into scaled ranks ranging from −0.5 to 0.5 with a mean of zero. See Appendix for variable definitions. Newey and West (1987) *t*-statistics with three lags are reported in parentheses, and ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Table 5.

Income Statement Buy Probability, Earnings Growth, and Earnings Announcement Returns

	(1)	(2)
	SUE_{q+1}	$RET[-1, 1]_{q+1}$
Intercept	-0.056***	0.001
	(-4.785)	(0.345)
ISBP	0.073***	0.005***
	(8.319)	(4.592)
SUE	0.323***	-0.004***
	(38.463)	(-3.403)
RS	0.112***	0.007***
	(25.387)	(4.530)
TES	0.045***	0.002*
	(8.851)	(1.709)
EA	-0.093***	-0.000
	(-26.067)	(-0.144)
TREND	-0.025***	-0.002**
	(-4.777)	(-2.101)
RET[-1, +1]	0.031***	0.003***
	(11.742)	(2.960)
RET[-30, -2]	0.053***	-0.000
	(12.945)	(-0.303)
PERSIST	-0.005	-0.002**
	(-0.971)	(-2.018)
VOL	-0.065***	-0.003
	(-4.320)	(-1.488)
SIZE	0.010***	0.000
	(5.384)	(0.293)
BM	-0.041***	0.004**
	(-3.228)	(2.263)
GPA	-0.026***	-0.001
	(-3.831)	(-0.201)
OPA	-0.050***	0.002
	(-5.299)	(1.390)
OA	-0.048***	0.002
	(-6.105)	(1.174)
TA	-0.035***	0.000
	(-4.858)	(0.369)
AG	-0.038***	-0.002
	(-10.422)	(-1.557)
Adj. R^2	0.196	0.009
obs.	142,273	142,104

The table presents results of Fama and MacBeth (1973) regressions in the out-of-sample period (2006Q3-2023Q2) using one-quarter-ahead standardized unexpected earnings (SUE_{q+1}) or the three-day abnormal return around the next earnings announcement date ($RET[-1, 1]_{q+1}$) as the dependent variable. The independent variables include the income statement buy probability (ISBP), standardized unexpected earnings (SUE), revenue surprise (RS), tax expense surprise (TES), earnings acceleration (EA), trend in gross profitability (TREND), firm size (SIZE), book-to-market ratio (BM), filing return (RET[−1, +1]), pre-filing return (RET[−30, −2]), earnings persistence (PERSIST), earnings volatility (VOL), gross profitability (GPA), operating profitability (OPA), total accruals (TA), operating accruals (OA), and asset growth (AG). All variables except for $RET[-1, +1]_{q+1}$ are converted into scaled ranks ranging from −0.5 to 0.5 with a mean of zero. See Appendix for variable definitions. Newey and West (1987) *t*-statistics with three lags are reported in parentheses, and ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively

Table 6.
Test of Stock Market Efficiency

	3-day	Quarter-Long
γ_1	0.073***	0.073***
	(10.374)	(10.383)
γ_1^*	0.001	−0.083
	(0.029)	(−1.099)
β	0.054***	0.068***
	(25.876)	(9.908)
Market efficiency test ($\gamma_1 = \gamma_1^*$)		
Likelihood ratio statistic	6.853***	4.369**

The table presents the results of Mishkin test of whether stock prices behave as if investors underreact to the impact of income statement buy probability (ISBP) on earnings growth. We estimate nonlinear generalized least squares estimation of the following two equations in the out-of-sample period (2006Q3-2023Q2)

$$\text{Forecasting equation: } \mathrm{SUE}_{q+1} = \alpha + \gamma_1 \mathrm{ISBP}_q + \sum \gamma_c \mathrm{Controls}_q + \delta_{q+1}$$
$$\text{Pricing equation: } \mathrm{AR}_{q+1} = \beta(\mathrm{SUE}_{q+1} - \alpha^* - \gamma_1^* \mathrm{ISBP}_q - \sum \gamma_c^* \mathrm{Controls}_q) + \varepsilon_{q+1}.$$

The independent variables include ISBP, standardized unexpected earnings (SUE), revenue surprise (RS), tax expense surprise (TES), earnings acceleration (EA), trend in gross profitability (TREND), firm size (SIZE), book-to-market ratio (BM), filing return (RET[−1, +1]), pre-filing return (RET[−30, −2]), earnings persistence (PERSIST), earnings volatility (VOL), gross profitability (GPA), operating profitability (OPA), total accruals (TA), operating accruals (OA), and asset growth (AG). AR is the abnormal return from a 3-day window around quarter $q+1$'s earnings or the quarter-long window starting two days after the 10-K/10-Q filing date in quarter q and ending on one day after the next earnings announcement date. All variables except for the abnormal return AR are converted into scaled ranks ranging from −0.5 to 0.5 with a mean of zero. t-statistics based on the firm and quarter double-clustered standard errors are reported in parentheses. The likelihood ratio statistic for testing $\gamma_1 = \gamma_1^*$ is distributed asymptotically as $\mathcal{X}^2(1)$. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.