Mimicking Pre-PEAD: The Predictive Power of Mispricing Scores [*]

--version 11 November 2025

Tiancheng Lan, Robert B. Durand, Yixuan Rui & John Gould[†]

Abstract

This study examines the relationship between cross-sectional mispricing at the end of the financial year and subsequent unexpected earnings for the following fiscal year. This relationship suggests that investors can mimic the pre-earnings announcement drift (pre-PEAD) first documented in Ball and Brown (1968). Further, the most underpriced (overpriced) stocks subsequently report significantly higher (lower) unexpected earnings over the next fiscal year and demonstrate. Our evidence suggests that investors with relatively concentrated portfolios are less likely to be able to exploit this relationship, as their limited diversification increases the probability of erroneous earnings prediction. We also find that the prediction relationship is stronger conditional on high institution or analyst coverage, and low investor sentiment.

Key words: Asset pricing, market anomalies, post-earnings announcement drift.

[*] For useful comments, we are grateful to the discussant Dr. Justin Nguyen and all the participants at the 2nd Western Australia's PhD Students Symposium in Accounting and Finance 2025, and participants at the 14th FIRN Annual Conference and 8th PhD Symposium 2025. We also thank seminar participants at Curtin University.

[†] Email: tiancheng.lan@curtin.edu.au (T. Lan, corresponding author), robert.durand@curtin.edu.au (R.B. Durand), yixuan.rui@curtin.edu.au (Y. Rui), j.gould@curtin.edu.au (J. Gould).

1. Introduction

Ball and Brown (1968) present seminal evidence that earnings are value-relevant but that the market does not fully and instantaneously incorporate this information as might be expected given the semi-strong version of efficient market hypothesis (Fama, 1970). Their pioneering study documents the post-earnings announcement drift (PEAD) that is stock prices tend to move in the direction of an earnings surprise for months after the announcement date. This finding has been extensively replicated and extended, both out-of-sample and across international markets, including in Ball and Brown's own 50-year retrospective (Ball and Brown, 2019).

In their original visual presentation[1], Ball and Brown (1968) plot cumulative abnormal returns around earnings announcements. While the PEAD has become the focal point of accounting anomalies and market efficiency research, the same figure also reveals a perhaps more striking price pattern. In the twelve months before the earnings announcement date (that is the period we will denote *t-1* to *t*), firms that have positive earnings surprises show strong upward abnormal returns, and those with negative surprises at time *t* show strong downward abnormal returns in the period *t-1* to *t,* where firms with negative surprises at time *t* show strong downward abnormal returns. This pattern, which we term pre-earnings announcement drift (pre-PEAD), implies that a significant portion of the eventual earnings surprise appears to be anticipated by the market before the official disclosure. Ball and Brown (1968), however, do not dwell on this aspect of their fame. Surprises, by definition, can not be known beforehand and, in Ball and Brown's experimental set up, these are *ex post* abnormal returns which can not be anticipated.

[1] We replicate the PEAD of Ball and Brown (1968) in Appendix A1

Pre-PEAD generates large abnormal returns when measured with the benefit of hindsight. Pre-PEAD is potentially unexploitable as investors would require perfect foresight about future unexpected earnings. However, the possibility remains that publicly available information might contain signals correlated with future earnings surprises, enabling investors to replicate part of the pre-PEAD pattern without such "god-like" power. For example, Ou and Penman (1989a, 1989b) and Penman and Zhu (2014) show such a relationship between available accounting ratios and future earnings levels and growth. Additionally, many of the accounting mispricing measures have been shown to carry predictive information about subsequent earnings performance. Sloan (1996) argues that the actual subsequent earnings of high (low) historical accruals are lower (higher) than investors' expectation which is the cause of accrual-based mispricing anomaly. Another example for anomaly-earnings relationship is the net-operating-income anomaly (Hirshleifer et al., 2004) which states a weaker subsequent earnings growth when the net operating income outstrips free cash flow over the same fiscal period. The research in this paper provides investors with potential to exploit current accounting ratio-based mispricing anomalies to achieve the pre-PEAD experience of Ball and Brown (1968).

We test two key hypotheses. First, portfolios formed on publicly available mispricing information at time *t-1* can forecast unexpected earnings of next fiscal year at time *t*. Second, these mispricing portfolios mimic the pre-PEAD experience of PEAD portfolios from *t-1* to *t*. We begin our work utilizing the mispricing measure of Stambaugh, Yu, and Yuan (*SYY*) (2012, 2015), which aggregates eleven well-documented mispricing anomalies. Combining multiple mispricing ranks into a composite mispricing score helps diversify the noise present in any single indicator and enhances the predictive accuracy of this metric. We also construct an aggregated accounting mispricing score (*ACCT*) that includes only six accounting-based components of *SYY.* These six anomalies are especially important because they share similar information characteristics with B&B's unexpected earnings, both being derived from annually released accounting data. All mispricing measures are calculated entirely from public accounting and market data available at the mispricing portfolio formation month-end. Thus, semi-strong efficiency suggests that any new information they provide should be discounted in prices.

Our first set of analyses sorts firms into quintiles based on historical mispricing scores at time *t-1* and examines the distribution of next year's (*t-1* to *t*) unexpected earnings, measured using Ball and Brown's (1968) definition. The results reveal a clear asymmetry: underpriced quintiles (highest mispricing scores) contain a greater proportion of firms with positive future earnings surprises, and overpriced quintiles (lowest mispricing scores) contain a greater proportion with negative surprises, indicating that historical mispricing scores are associated with future earnings outcomes that are not yet observable when the portfolios are formed. Our regression results of future unexpected earnings on the historical mispricing scores confirm the sort results. Controlling accounting and price-volume information, institutional ownership- and analyst recommendation-related variables, the coefficients on mispricing measures are positively significant. Stocks with higher mispricing score which is identified as more underpriced are more likely to deliver better-than-expected earnings in the subsequent fiscal year. The predictive relationship holds across multiple specifications, alternative definitions of unexpected earnings, and both in-sample and out-of-sample tests.

Having established the link between mispricing scores and future earnings surprises, we next consider the extent to which portfolios based on historical mispricing scores can replicate pre-PEAD returns. We first measure the perfect foresight pre-PEAD by forming portfolios on next year's unexpected earnings. It is an infeasible strategy in practice and therefore represents a "god-like" strategy, but one that potentially provides an upper bound on potential pre-PEAD returns and also echoes the spirit of Ball and Brown's original work. We then replicate the same holding period and portfolio construction procedure using historical mispricing scores instead of future earnings. The results show that portfolios based on mispricing scores earn returns in the same direction as those associated with the next fiscal year's earnings surprises and achieve approximately 40 to 60 percent of the returns from the foresight strategy. This indicates that a meaningful share of pre-PEAD can be replicated with information available at the start of the holding period.

When we test the return investors can achieve, we find that mispricing portfolios may generate returns with a direction against the initial mispricing signal if the future earnings do not confirm the initial holding position. Specifically, if the underpriced (overpriced) stock selected by an investor based on historical mispricing information further generates lower (higher) unexpected earnings in the next fiscal year, the return of the selected stock will be in an opposite way of the initial mispricing signal which is underpriced (overpriced) stocks will outperform (underperform). This finding indicates that investors need to predict earnings correctly to gain mispricing premium over the holding period. Using random stocks selection from mispricing portfolios, we investigate the prediction rate and future returns of stocks selections with different stock numbers. We find that selections with smaller number of stocks are associated with lower probability of achieving earnings-confirmed portfolios (a correct prediction) on both long and short sides. This difference is especially pronounced for overpriced portfolios, where incorrect predictions lead to unattractive return-risk profiles and weaker incentives to short sell.

We further find that the mispricing information carries more unexpected earnings for the next fiscal year if a stock has high institutional ownership, high analyst coverage, or when the prediction is followed by a period of low investor sentiment. Under these conditions, mispricing signals are more closely tied to fundamentals and less contaminated by arbitrage-risk-dominant noise, making the fundamental component of mispricing more dominant and therefore more effective in forecasting next-period earnings shocks, which are themselves fundamentally determined.

This study contributes to the mispricing anomalies and market efficiency literature in several ways. First, it links current mispricing to future material but unavailable financial information. Building on prior findings that historical accounting ratios are correlated with changes in future earnings growth (Ou and Penman, 1989a, 1989b; Penman and Zhu, 2014), we provide evidence that the cross-sectional mispricing based on historical accounting ratios captures information on future earnings surprises. Investors can utilize this link to form portfolios, and this mispricing portfolio can achieve 40 - 60 percent pre-PEAD experience which relies on perfect foresight on earnings. Second, we find that mispricing premium is dependent on the allocation of constituents' earnings. If the future earnings exhibit a different direction, mispricing premium does not exist. Extent to the second point, the third contribution of this study is that our random stock selection tests yield evidence that portfolios with low number of stocks (undiversified) are associated with lower chances of achieving a correct earnings prediction, which is more pronounced among overpriced selections. This finding is also consistent with the current mispricing pricing papers that overpriced stocks are more mispriced compared to underpriced stocks (Stambaugh, Yu and Yuan, 2012, 2015). This paper also shows that the predictive power of mispricing information for unexpected earnings is stronger conditional on a greater market efficiency as the mispricing information is more fundamental dominant.

We organize the rest of the paper as follows. Section 2 includes a brief literature review. We describe anomaly construction processes in section 3. Section 4 provides data collection and our empirical results. Section 5 concludes.

2. Literature review

Ball and Brown (1968) provide the first evidence that accounting earnings contain information relevant to equity valuation and that the market does not fully incorporate this information on or before the announcement date. It reveals a post-earnings announcement drift (PEAD), whereby stock prices continue to move in the direction of the earnings surprise after the information is publicly available. This evidence has been verified out of sample in international market using more recent data (Ball and Brown, 2019). Foster, Olsen and Shevlin (1984) replicate this finding using standardized unexpected earnings as a measure of earnings shocks and report that the long-short portfolios based on earnings over 60 days following the announcement yield an annualized abnormal return of approximately 25 percent. Bernard and Thomas (1989) attribute PEAD to investor underreaction, arguing that prices do not fully adjust to earnings-related information at the time of disclosure, leading to continued drift in the same direction. Consistent with the underreaction interpretation, subsequent studies document that PEAD is more pronounced among firms with greater information delay or trading friction, indicating that slower information diffusion and trade friction exacerbates the price drift after announcement (Haugen and Baker, 1996; Hou and Moskowitz, 2005; Ng, Rusticus and Verdi, 2008; Cao and Narayanamoorthy, 2012).

While the literature focuses on the returns following earnings announcement, the price drift leading up to the announcement, that is the pre-earnings announcement drift (pre-PEAD), remains less examined, despite generating much higher returns than PEAD experience does. Existing research on earnings prediction which exploits historical accounting ratios provides potential to capture a portion of pre-PEAD returns. Ou and Penman (1989a, 1989b) provide empirical evidence that accounting ratios can predict future earnings. By examining a large set of financial statement variables, they show that 34 out of 68 ratios are significantly correlated with next year’s earnings. Penman and Zhu (2014) further document that accounting ratios forecast forward earnings and growth in the same direction as they forecast returns. Accounting factors are also found impactful on earnings volatility, and the consideration of earnings volatility improves the prediction on both long- and short-term earnings (Dichev and Tang, 2009; Frankel and Litov, 2009).

Mispricing studies, especially those focusing on accounting anomalies, provide the return outcome of portfolios using publicly available accounting information. The accounting anomalies can be grouped by major accounting categories. The first category is profitability. Fama and French (2008) find that a higher ratio of operating profits to book equity is associated with superior stock performance. Similar profitability-related anomalies are reported using net income (Haugen and Baker, 1996), gross profit (Novy-Marx, 2013), earnings surprise (Ball and Brown, 1968; Foster, Olsen and Shevlin, 1984), sales growth (Abarbanell and Bushee, 1998). Beyond profitability, firm valuation is also linked to mispricing. Firm valuation proxies include earnings-to-price ratio (Basu, 1977), book-to-market ratio (Fama and French, 1992), dividend yield (Boudoukh et al., 2007). Investment-related ratios are another source of anomalies as documented by Titman, Wei, and Xie (2004); Cooper, Gulen and Schill (2008); Soliman (2008); and Richardson et al, (2005).

Anomalies other than accounting ratios have also been widely studied, including issuance-related mispricing (Ritter (1991), Loughran and Ritter (1995); Fama and French (2008)), momentum (Jegadeesh and Titman (1993), Grinblatt and Moskowitz (1999), and Blitz, Huij, and Martens (2011)), and distress anomalies (Ohlson (1980), Altman (1983), and Campbell, Hilscher, and Szilagyi (2008)). To consolidate these diverse sources of mispricing, Stambaugh, Yu, and Yuan (2012, 2015) select eleven anomalies which survives on Fama-French three factor model and aggregate these eleven anomalies into a mispricing score. This aggregation diversifies idiosyncratic noises from single anomaly signals and captures their common mispricing component.

3. Anomaly construction

This section outlines the construction of key anomalies in this study and the data collection process. Specifically, we focus on Ball and Brown (1968)'s measure of unexpected earnings, the mispricing scores developed by Stambaugh, Yu, and Yuan (2015), and an aggregated accounting-based mispricing score derived from six annual accounting ratios following Stambaugh, Yu, and Yuan's methodology.

3.1 Ball and Brown's unexpected earnings

Following Ball and Brown (2019), we utilize the primary annual earnings per share (*EPSPI)*[2] as our primary earnings measure, which captures net income including non-recurring items. Unexpected earnings are defined as the prediction error between the actual annual change in EPS and its predicted value based on the market model proposed by Ball and Brown (1968). By incorporating the proxy for market-wide earnings expectations, this measure of unexpected earnings is more likely to capture earnings-related price adjustments. To estimate unexpected earnings, we employ an OLS-based market model, where the change in a firm's EPS is regressed on the change in the average market EPS, computed as a share-weighted average of EPS across all firms in the market, excluding the firm under estimation:

$$\Delta EPS_{j,t-i} = \overline{a_{j,t}} + \overline{b_{j,t}} \Delta MKTEPS_{j,t-i} + \varepsilon_{j,t-i}, \tag{1}$$

the subscript i denotes the estimation window, which indicates that the estimation period begins at the earliest fiscal year-end with available EPS data up to the fiscal year-end immediately preceding year t. Following Ball and Brown (1968), our sample consists of stocks with at least 10-year consecutive EPS disclosures.

[2] COMPUSTAT reports four distinct measures of annual earnings per share (EPS). (1) EPSPI: Primary EPS, including extraordinary items. (2) EPSPX: Primary EPS, excluding extraordinary items. (3) EPSFI: Fully diluted EPS, including extraordinary items. (4) EPSFX: Fully diluted EPS, excluding extraordinary items.

The expected change in earnings for firm *j* in year *t* is estimated using the regression model (2) below, where the prediction is based on $\overline{a_{J,t}}$ and $\overline{b_{J,t}}$ derived from model (1):

$$E\left(\Delta EPS_{j,t}\right) = \overline{a_{J,t}} + \overline{b_{J,t}} \Delta MKTEPS_{j,t} \tag{2}$$

The unexpected earnings change is the difference between the actual change in EPS of firm *j* in year *t* and the expected earnings change:

$$\varepsilon_{j,t} = \Delta EPS_{j,t} - E\left(\Delta EPS_{j,t}\right) \tag{3}$$

3.2 Stambaugh, Yu, and Yuan's mispricing score

Stambaugh, Yu, and Yuan (2015) construct a cross-sectional ranking-based mispricing score to serve as a proxy of mispricing. Following their methodology, stocks are ranked within each cross-section based on eleven well-documented return anomalies that persist after controlling for the Fama-French three-factor model. These anomalies consist of six annual accounting-based anomalies and five anomalies with higher rebalancing frequencies (quarterly or monthly).

The six annual accounting-based anomalies are:

1. Ohlson O-score (Ohlson, 1980)
2. Total accruals (Sloan, 1996)
3. Net operating assets (Hirshleifer et al., 2004)
4. Gross profitability (Novy-Marx, 2013)
5. Asset growth (Cooper, Gulen, and Schill, 2008)
6. Investment-to-assets (Titman, Wei, and Xie, 2004; Xing, 2008)

The five high-frequency anomalies include:

7. Financial distress (Campbell, Hilscher, and Szilagyi, 2008)

8. Net stock issues (Ritter, 1991; Loughran and Ritter, 1995; Fama and French, 2008)
9. Composite equity issues (Daniel and Titman, 2006)
10. Momentum (Jegadeesh and Titman, 1993; Asness, 1995)
11. Return on assets (ROA) (Fama and French, 2006; Chen, Novy-Marx, and Zhang, 2010)

To ensure comparability across anomalies, each measure is first converted into cross-sectional ranks. Stocks with higher ranks are more inclined to generate positive subsequential abnormal returns (underpriced), where stocks with lower ranks (overpriced stocks) generate negative mispricing return in the next period. The final mispricing score for each stock is calculated as the cross-sectional average of all available ranks. Stocks with fewer than five available anomaly measures in a given cross-section are excluded from the computation to maintain robustness (Stambaugh, Yu and Yuan, 2012, 2015; Kumar, Motahari and Taffler, 2022).

3.3 An aggregated accounting-based mispricing score

Since six out of the eleven anomalies in Stambaugh, Yu, and Yuan's mispricing measure are based on annual accounting ratios, we also investigate the relationship between historical annual accounting mispricing and future earnings surprise as they share the similar information characteristics that are both annually released accounting ratios and released at the same time.

To achieve this, we follow the methodology outlined in Section 3.2, but we restrict the calculation to the six annual accounting-based anomalies. Specifically, for each firm, we compute the average ranks across these six anomalies at each cross-section, yielding an aggregated score based on annual accounting ratios. The accounting mispricing score (*ACCT*) of stocks with at least three of the anomaly variables are computed at each cross section.

4 Empirical results

4.1 Data collection

Our final sample during the period 1974 to 2024 comprises stocks with available unexpected earnings data and the mispricing scores. We obtain accounting data and quarterly reporting dates from COMPUSTAT to calculate unexpected earnings and construct the mispricing anomalies. Stock price data are sourced from CRSP. Risk factors for asset pricing models are retrieved from Kenneth French's website. Analyst- and institutional holding-related data is obtained from LSEG IBES and LSEG institutional holdings (13F).

We apply some selection criteria to our sample to ensure our measurements are clean and consistent. To allow for consistency in timing across different firms, we restrict our firms to those with a December fiscal year-end to predict earnings using the market model of Ball and Brown (1968). Also, firms must have 10-year consecutive EPS reporting to be considered. Consequently, the earliest fiscal year that the unexpected earnings are available in our sample is 1974. The accounting ratios used to compute the mispricing scores are available at the fourth-quarter reporting date of 1973, which generally occurs in the first quarter of 1974. We also require a minimum of 500 stocks in the cross-sectional analysis to ensure statistical power and stable coefficient estimates[3]. For regression-based tests, all variables are winsorized at the 1st and 99th percentiles of their cross-sectional distributions and standardized cross-sectionally to have mean of zero and standard deviation of one. Additionally, stocks trading below one dollar at the fourth-quarter reporting date are excluded to mitigate the impact of penny stocks.

[3] We impose a restriction of at least 500 stocks in each cross-section to ensure that the quintile portfolios formed on mispricing scores remain well diversified. This requirement guarantees a minimum of approximately 100 stocks per quintile, which in turn allows us to further split each quintile portfolio into two groups based on subsequent earnings outcomes while maintaining adequate diversification within each subgroup.

4.2 Descriptive statistics

Table 1 presents summary statistics for unexpected earnings, the individual anomaly measures used in constructing the mispricing scores, and the *SYY* mispricing score. As described in Sections 3.2, the anomaly indicators have been re-oriented so that higher (or lower) values consistently correspond to higher (or lower) future returns and winsorized at 1% and 99% level. Panel A reports that the Ball and Brown's unexpected earnings are on average slightly negative. A positive median suggests that more than half of firms generate positive unexpected earnings, which means the distribution of unexpected earnings on average is negatively skewed. The value at 25% (-0.694) and 75% (0.741) are symmetric, indicating that the negative skewness is driven by extreme values at left tail. The mean and median values of *SYY* mispricing score are close to 49.5 because it is the average of eleven cross sectional ranks ranged in 0 to 99. For each of eleven anomalies to construct *SYY* mispricing score, Panel A reports the summary statistics of the raw value. Panel B presents the correlation matrix across the unexpected earnings in the next fiscal year at time *t*, *SYY* mispricing scores at time *t-1* and eleven anomalies used to construct *SYY* mispricing scores. Seven of the eleven anomaly indicators, as well as the *SYY* mispricing score, are positively correlated with foresight unexpected earnings. Notably, five of the six annual accounting anomaly indicators show positive correlations with foresight unexpected earnings, where O-score is the exception, offering initial evidence of a relationship between stock mispricing and future earnings surprises.

[Insert Table 1 about here]

Table 2 reports the average portfolio characteristics obtained from univariate sorts based on the mispricing measure proposed by Stambaugh, Yu, and Yuan (*SYY*), as well as the interaction between mispricing quintile portfolios and future unexpected earnings. We ensure that all mispricing signals are constructed using information that is available at the time of portfolio formation. Specifically, firms release their financial statements at the fourth quarter reporting date. Assuming that by the end of the reporting month, both financial and market data have been disseminated and absorbed by the market. The *SYY* mispricing score is constructed using cross-sectional rankings, which requires the most updated information for all firms within each cross section. For firms with December fiscal year end which release their accounting information in February, their anomaly information can not be sorted until information from firms with later fiscal year ends than February becomes available. To ensure consistent data availability across firms, we calculate the mispricing score at the end of June each year ($t-1$). For instance, when predicting the unexpected earnings of fiscal year 2010 released in early 2011, we sort stocks and calculate their mispricing scores at the end of June 2010. It is important to note that, following Ball and Brown (1968), unexpected earnings for fiscal year 2010 are not yet observable at the time of portfolio formation (June 2010), and they become fully known when the firm releases its earnings next year (early 2011).

[Insert Table 2 about here]

Panel A displays the characteristics of the *SYY* mispricing quintile portfolios and the difference in each of characteristics between underpriced and overpriced portfolios. Compared to overpriced portfolio, underpriced stocks are larger in size, less volatile, less correlated with the market movement. Most importantly, it is evident that future unexpected earnings tend to increase as stocks become more underpriced. Specifically, the future unexpected earnings generated by underpriced stocks on average are statistically higher than overpriced stocks at 1% significance level. This relationship is further demonstrated in Panel B and C, which summarizes interactions between mispricing categories and future earnings surprises. Panel B and C report supporting statistics, including the number of stocks, average firm size, and average unexpected earnings for each interaction group. In Panel B, the average unexpected earnings for the overpricing combined with cross sectional low-earnings group are not only negative but also larger in magnitude compared to the overpricing combined with high-earnings group. Notably, out of 215 overpriced stocks, on average 125 are cross-sectional losers regarding unexpected earnings in the next fiscal year, where only 90 are winners, which is statistically significantly lower than the number of losers. Conversely within underpriced group, 122 out of 215 are associated with higher-than-median future earnings surprises, which are significantly higher than the number of losers. A similar pattern can be found in Panel C where unexpected earnings winners are defined as firms with greater-than-zero unexpected earnings. Overall, we find that overpriced stocks are more likely to generate below-median or negative unexpected earnings, whereas underpriced stocks tend to yield above-median unexpected or positive earnings in the following year. These results indicate that the current mispricing score is significantly associated with next year's unexpected earnings.

4.3 The relationship between historical mispricing and future unexpected earnings

To validate the predictive power of mispricing signals, it is essential that a firm's mispricing score is correlated with its subsequent earnings surprises. In other words, the distribution of future earnings shocks must be asymmetric across mispricing groups, specifically underpriced stocks should predominantly experience positive earnings shocks, while overpriced stocks should be more likely to incur negative shocks.

To empirically test this hypothesis, we employ OLS panel regressions using the following specification:

$$Unexp\Delta EPS_{j,t} = a + b_1 MISP_{j,t-1} + \sum_{i=1}^{n} \gamma_i control_{i,j,t-1} + \delta_{t-1} + \rho_j + \varepsilon_{j,t-1}, \quad (4)$$

where the dependent variable represents the unexpected earnings for fiscal *year t*. The independent variables consist of information for fiscal *year t-1*. Specifically, $MISP_{j,t-1}$ denotes the mispricing score *SYY*, $control_{i,j,t-1}$ denotes control variable *i*. The set of control variables for the baseline model includes the earnings per share for the previous fiscal year *t-1*, the natural logarithm of market capitalization, the book-to-market ratio, the investment-to-asset ratio, gross profit, and short-, medium-, and long-term past returns, all measured at the end of June in the year prior to the firm's earnings announcement year. The model controls for firm and fiscal year fixed effects to account for unobserved heterogeneity across firms and fiscal years. Standard errors are clustered at the firm and fiscal year levels to correct for potential cross-sectional and serial correlation in the residuals.

Table 3 presents the estimated regression coefficients for unexpected earnings. The first column of Table 3 shows the estimated coefficients of independent variables of our baseline model. The significantly positive coefficient of *SYY* mispricing score indicates a positive relationship between firm's mispricing in the past and unexpected earnings in the future. As all variables are standardized, the coefficient on the *SYY* mispricing variable can be interpreted directly: a one standard-deviation increase in the mispricing score is associated with a 0.152 standard-deviation increase in unexpected earnings per share and this relationship is statistically significant at 1% level with *t-value* of 11.220.

As the baseline model tests the predictive power of historical mispricing on future unexpected earnings, analyst- and institutional holding- related information may also add value to the prediction. The regression models for the second and third columns control for these by adding analyst recommendation- and institutional holding- related variables, where the model for the fourth column puts all control variables together. Across different regression models with different sets of control variables, the relationship between firm's mispricing in the past is always statistically significantly positive at 1% level suggesting that the earnings information carried by mispricing scores is not fully utilized by analysts and financial institutions.

[Insert Table 3 about here]

We further employ a logistic panel regression using the following specification (Equation (5)) to test whether the mispricing score available in June prior to the earnings announcement year helps investors to predict the direction of unexpected earnings which will be released in the following year. This test is motivated by the observation in Ball and Brown's (1968) original work that stocks with positive (negative) unexpected earnings in the fiscal year *t* is seen generating positive (negative) returns over the period *t-1* to *t*. Predicting the direction of future earnings is also potentially profitable while being methodologically simpler than predicting raw value of future earnings.

$$\ln\left(\frac{Pr\left(PosUE_{j,t}=1\right)}{1-Pr\left(PosUE_{j,t}=1\right)}\right)= a+b_1 MISP_{j,t-1}+\sum_{i=1}^{n}\gamma_i control_{i,j,t-1}$$

$$+\delta_{t-1}+\rho_{industry}+\varepsilon_{j,t-1}, \tag{5}$$

where, for firm *j* and fiscal year *t:* the binary positive unexpected earnings indicator variable $PosUE_{j,t}$, equals one if the Ball and Brown unexpected earnings of this firm for the fiscal year *t* is positive, and zero otherwise. All independent variables and fixed effects are the same as Equation (4). As Logistic model does not allow perfect prediction, we relax the fixed effect from previously fiscal year and firm level to fiscal year and industry level.[4]

[Insert Table 4 about here]

Table 4 report presents the estimated regression coefficients for unexpected earnings prediction model using Equation (5). Similar to Table 3, the first column reports the estimated coefficients for the independent variables in our baseline model. The coefficient on the mispricing variable is 0.318, statistically significant at the 1 percent level with *t-value* of 18.814. One-standard-

[4] In a binary outcome, some firms always take $PosUE_{j,t}=1$ or 0. The likelihood then can not estimate coefficients once firm fixed effect is added because it causes perfect prediction which is not allowed by logistic model. To avoid perfect prediction and keep sufficient variation within each fixed group, we apply industry fixed effect.

deviation increase in the *SYY* mispricing score available in June of year *t-1* raises the log-odds of positive unexpected earnings reported in the next year *t* by 0.318. Based on the empirical distribution of unexpected earnings, approximately 52% of firms on average report positive unexpected earnings, implying that a randomly selected firm has a 52 percent probability of a positive earnings shock. Given this benchmark, a 0.318 increase in log-odds corresponds to an estimated probability of roughly 60 percent for positive unexpected earnings under the baseline model.[5] The second and third columns include additional controls related to analyst coverage and institutional ownership, respectively, while the fourth column includes all control variables simultaneously. Across all specifications, the results remain robust, indicating that the information about future earnings embedded in the *SYY* mispricing score is not fully incorporated by analysts or institutional investors.

4.4 The impact of the timing of mispricing score calculation

In the previous tests, the mispricing score was calculated by sorting stocks at the end of June to ensure that each cross section included the most recently released accounting information for all firms. However, when computing the mispricing score for a single firm, it is feasible to update the score immediately after the firm releases its financial statements in its announcement month. For example, a firm with a February fiscal year-end can be ranked at the end of February based on its newly released accounting information. The challenge, however, arises in the cross-sectional context: when sorting all firms in February 2011 using accounting ratios from fiscal year 2010, some firms with later fiscal year-ends will not yet have released their fiscal-year-2010 data, so their most recent information would still pertain to fiscal year 2009. The trade-off, therefore, is between timeliness and data consistency. Sorting in the announcement

[5] If the benchmark probability P_0 is 0.52, $odds_0 = \frac{0.52}{1-0.52} = 1.083$. After a one-standard-deviation increase in the *SYY* mispricing score, $odds_1 = 1.083 \times e^{0.318} = 1.483$. The new probability becomes $P_1 = \frac{1.483}{1+1.483} = 0.60$

month provides investors with an earlier signal of a firm's mispricing level and opportunity to predict future earnings surprises in advance, but it reduces cross-sectional comparability due to differences in reporting times across firms. It is therefore interesting to examine whether the mispricing score measured in the announcement month, when a firm releases its accounting information, can predict unexpected earnings in the subsequent fiscal year.

In the following tests, we repeat the OLS and logistic regression analyses conducted in the previous section, with the key difference that all independent variables, including the *SYY* mispricing scores and control variables, are measured at the end of each firm's fourth-quarter announcement month of fiscal year *t-1*. This timing ensures that all explanatory variables are based on information available twelve months before the earnings for fiscal year *t* are released.

Table 5 presents the estimated regression coefficients from Equation (4), where both the mispricing score and control variables are measured at the end of each firm's fourth-quarter announcement month of fiscal year *t-1*, reflecting the earliest time point at which firm-level accounting information becomes available for calculating the mispricing score and further for predicting unexpected earnings for the next fiscal year.

[Insert Table 5 about here]

The first column of Table 5 shows that when each stock is sorted in its own announcement month, which is earlier than the common June sorting month used in previous tests and at the cost of reduced cross-sectional comparability, the current *SYY* mispricing score still significantly predicts the unexpected earnings of the following fiscal year. Compared with the results in Table 3, however, the earlier sorting leads to less accurate cross-sectional alignment especially for firms with earlier fourth-quarter announcement month, resulting in a roughly 50 percent reduction in the economic magnitude of the baseline coefficient. Nevertheless, the coefficient remains statistically significant at the 1% level, suggesting that even when measured earlier, the *SYY* mispricing score continues to capture meaningful forward-looking information about firms' future earnings performance. For the regression models that include control variables related to analyst recommendations and institutional holdings, the relationship between historical mispricing and future unexpected earnings remains significantly positive, although the corresponding coefficients are roughly half the magnitude of those reported in Table 4, where the mispricing score and control variables are measured later at the end of June. Similarly, Table 6 presents the results from the logistic regression model based on Equation (5), where each firm's independent variables are measured at the end of its announcement month. The results are consistent with earlier findings: firms that were more underpriced in the past are more likely to report positive unexpected earnings in the subsequent fiscal year. However, the predictive power of the mispricing score is weaker compared to that of the score calculated later at the end of June, as reported in Table 4.

[Insert Table 6 about here]

4.5 Accounting mispricing and future unexpected earnings

The *SYY* mispricing score is constructed from six annual accounting anomalies and five anomalies based on monthly or quarterly information. As shown in the correlation matrix panel of Table 1, five of the six annual accounting anomaly indicators are positively correlated with unexpected earnings in the following fiscal year. This finding motivates an examination of whether accounting-based mispricing alone can predict future earnings surprises. We construct an aggregated *accounting mispricing score (ACCT)* that includes only the six accounting-based components of the *SYY* measure. These six anomalies are particularly relevant because they share similar information characteristics with Ball and Brown's (1968) unexpected earnings, both being derived from annually released accounting data. The *ACCT* mispricing score is calculated entirely from public accounting and market data available at the mispricing portfolio formation month-end which is either at the end of fourth-quarter announcement month (*RD-Q4*) or at the later end-of-June date (*June*). Table 7 reports the estimated regression coefficients from both OLS and logistic panel regressions of Equation (4) and (5), respectively, with control variables for the baseline model in which the *SYY* mispricing is replaced with the accounting mispricing measure (*ACCT*). Consistent with the findings for the *SYY* mispricing score, the accounting-based mispricing measure is also positively associated with future unexpected earnings. Firms that are more underpriced tend to generate higher (or positive) unexpected earnings in the subsequent fiscal year, a relationship that holds across both the OLS and logistic regression models as well as under both sorting timings (*RD-Q4* and *June*).

[Insert Table 7 about here]

The coefficient on the *ACCT* mispricing score is smaller than that on the *SYY* measure reported in previous tests, as the *ACCT* specification omits the five non-accounting anomalies contained in *SYY*. This reduction suggests that the excluded monthly and quarterly components provide additional predictive information about future earnings beyond that captured by annual accounting data alone. Unlike the earlier finding that moving the timing of the *SYY* mispricing calculation from June to the fourth-quarter announcement month slightly weakens its predictive power, the *ACCT* measure exhibits stronger predictive ability when constructed at the end of the fourth-quarter announcement month (*RD-Q4*) rather than at the later June date. This pattern likely reflects the fact that, between *RD-Q4* and June, more up-to-date market information such as momentum becomes incorporated into prices, thereby reducing the predictive power of accounting-based mispricing. Consistent with this interpretation, the coefficient on the control variable *Ret[–12, –2]* increases substantially when the portfolio formation date is moved to June.

4.6 Out-of-sample prediction on the unexpected earnings for the next fiscal year

To further assess the robustness of the relationship between mispricing scores and future earnings shocks, we conduct an out-of-sample test to evaluate the predictive power of the mispricing indicator. We evaluate the classification accuracy of stocks by comparing the predicted earnings sign with the actual direction of unexpected earnings of the next fiscal year. Specifically, for each cross section, we compute the estimated earnings using control variables of the baseline model of Equation (4), with variable loadings obtained as the average of the coefficients from the previous five cross sections. This methodology ensures that all the information employed in forecasting next year's unexpected earnings is publicly available. The resulting table reports, on an average fiscal-year basis, the number of stocks classified into each category with diagonal entries denoting correct classifications (i.e., the estimated earnings sign matches the actual unexpected earnings sign) and off-diagonal entries representing misclassifications.

At each cross section *t,* firms are classified into positive (+) and negative (-) signals based on predictions, where variable loadings are estimated using cross-sectional regression over the past five years. Realized unexpected earnings in year *t* are similarly categorized into positive and negative outcomes. Table 8 shows that, on average, 567 stocks experience positive earnings shocks, slightly exceeding the number of firms with negative shocks. This asymmetry is consistent with Table 1, which reports a positive median Ball and Brown (1968) unexpected earnings measure. Under a random or neutral selection process, the expected probability of selecting a firm with a positive (negative) earnings shock is 52.36% (47.64%). In contrast, when incorporating the correlation between historical mispricing and subsequent earnings outcomes, the model improves predictive accuracy to 60.95% for positive earnings shocks and 55.60% for negative ones. These gains indicate that historical mispricing information enhances investors' ability to forecast the direction of future earnings surprises relative to an unbiased benchmark which is the chance generated by random selection. The final column of Table 8 reports the statistical significance of the difference between the model's predictions and the cross-sectional binomial distribution for each fiscal year. The results show that the use of *SYY* mispricing significantly increases the likelihood of correctly identifying the direction of future earnings shocks, and that the accounting-based *ACCT* mispricing measure yields qualitatively similar predictive improvements.

[Insert Table 8 about here]

4.7 Mimic B&B's pre-PEAD using mispricing signals

Prior analysis of this paper finds that mispricing signals derived from historical mispricing indicators are correlated with subsequent earnings shocks. Firms identified as underpriced tend to deliver more favourable unexpected earnings in the next fiscal year. As B&B portfolios require "god-like" perfect foresight on the unexpected earnings will be released 12 months later, it is not feasible to construct in practice. In contrast, mispricing portfolios rely solely on observable ex-ante signals and are therefore implementable regarding information availability. Our previous finding provides us with an opportunity to mimic the pre-PEAD returns of unexpected earnings before it becomes publicly available, as B&B's visual presentation shows an extremely large return over the pre-PEAD period. Motivated by the opportunity to achieve such incredibly attractive returns, the following tests investigate whether portfolios formed using mispricing scores can mimic the pre-PEAD experience of B&B earnings price drift.

We begin this investigation by visualizing return drift over the pre-PEAD (–12 to 0) and PEAD (1 to 12) periods using cumulative abnormal return plots. Figures 1 plots four-factor-adjusted cumulative abnormal returns of B&B portfolios and mispricing portfolios across a 24-month window from –12 to 12 months.[6] Portfolios are formed using both median-split (Panel A) and quintile sorts[7] (Panel B) on three mispricing signals including (1) accounting mispricing score *(ACCT)* based on six annual accounting ratios*,* (2) mispricing score of *SYY* which we only obtain information at the beginning of the pre-PEAD period*,* and (3) a version of the SYY mispricing score that is updated monthly throughout the full period *(SYYM)*. All mispricing information was released at time -12 to ensure the information availability where earnings are released at time 0, which means B&B portfolios remain "god-like" over the pre-PEAD period but exhibit the power of mortals during PEAD period.

[Insert Figure 1 about here]

When portfolios are formed using a median split sort (Panel A), the B&B portfolios exhibit a pronounced return drift during the pre-PEAD period. Compared to all mispricing portfolios, the "god-like" B&B strategy generates the highest cumulative abnormal returns across short, long, and long-short positions. Among the mispricing signals, the *ACCT* portfolio shows the weakest performance on the short side, whereas *SYYM* generates the most negative drift. The *SYY* portfolio lies between these two. This pattern suggests that incorporating additional anomaly indicators into accounting-based signals and refreshing information timely, as in the *SYYM* score, enhances the ability to capture future negative returns in overpriced firms.

In the long position return plot (the second figure of Panel A), the B&B portfolio again delivers the highest pre-PEAD cumulative return, reflecting the advantage of perfect foresight. Underpriced portfolios show similar upward drift patterns but at lower magnitudes. Taken

[6] The four-factor-adjusted monthly return abnormal returns are calculated with respect to the Fama-French-Carhart four-factor asset pricing model. For each stock-month, the Fama-French-Carhart four-factor asset pricing factor loadings are estimated out-of-sample for the [-60, -1] months.

[7] The monthly returns of portfolios sorted by historical mispricing scores (*ACCT* and *SYY*) and foresight B&B earnings are reported in Appendix A2. The monthly returns of B&B portfolios are not realizable in practice due to the unavailability of earnings information at the time. However, these returns represent the information of perfect foresight on future unexpected earnings.

together, the B&B long-short portfolio achieves the largest cumulative return over the pre-PEAD period, followed sequentially by those based on *SYY*, *SYYM*, and *ACCT*.

During the PEAD window, the return drift of long-short portfolios persists following the earnings announcement. This finding aligns with prior literature on post-earnings announcement drift. Among the mispricing strategies, *SYYM* exhibits stronger profitability in the latter part of the PEAD period, likely due to its monthly information updates, which provide more timely adjustments in portfolio holdings in response to newly available information.

Panel B, which presents results based on quintile-sorted portfolios, exhibits similar but more pronounced patterns due to the narrower cross-sectional grouping. A particularly notable finding is that the *SYYM* overpriced portfolio outperforms the B&B overpriced portfolio during the pre-PEAD period. This result is important because it demonstrates that under certain conditions, specifically when mispricing is measured with higher granularity and updated dynamically, feasible mispricing strategies can exceed the return performance of an idealized "god-like" earnings foresight strategy. The improved precision from quintile sorting, coupled with the incorporation of additional anomalies and monthly updates, allows the *SYYM* strategy to better identify, exploit overpriced securities and exceed the information intensity of B&B's earnings foresight. However, despite this strength on the short side, the underpriced portfolios constructed from *SYYM* and other mispricing signals underperform their B&B counterparts. As a result, the overall long-short mispricing portfolios do not surpass the B&B long-short returns in the pre-PEAD window.

We then replicate the analysis of Figure 1 in Figure 2, but with an additional MCAP filter of the bottom tercile in each month applies to the investable universe. As our objective is to evaluate whether current mispricing portfolios mimic the B&B foresight portfolios, it is important to investigate if mispricing returns are simply artifacts of or driven by arbitrage risk or liquidity frictions (Stambaugh, Yu and Yuan, 2015; Hou, Xue and Zhang, 2020). Our earlier results show that the *SYYM* strategy can even outperform B&B portfolios in short positions under quintile sorts, raising the concern that such profitability is concentrated in small, illiquid stocks. Consistent with this concern, Campbell, Hilscher, and Szilagyi (2008) find that the average size of financial distress (one anomaly included in *SYYM*) anomaly's short position is extremely low. Momentum strategy's short position also exhibits similar characteristics. Therefore, we re-estimate the drift patterns after excluding firms in the bottom tercile of market capitalization in each month, in order to ensure that the observed performance is not driven disproportionately by small-cap stocks and to better assess the feasibility of implementing these strategies in practice.

[Insert Figure 2 about here]

Table 9 reports the differences between cumulative abnormal returns of B&B foresight portfolios and those of mispricing portfolios. T-test is applied to test the significance of difference through three horizons including the first six months of pre-PEAD (-12 to -6), the full pre-PEAD (-12 to 0), and the combined pre-PEAD and PEAD period (-12 to 12). Similar to the visual presentation, median split and quintile sorting are applied.

A consistent feature of the pre-PEAD results is that the difference between B&B and mispricing long-short portfolios are positively significant in most cases. Under median-split sorting, B&B's long-short portfolios exceed that of *ACCT* by 3.96% and *SYY* by 2.91% respectively over the full pre-PEAD period, where quintile sorting results in similar results that is 3.83% for *ACCT* and 3.27% for *SYY*. As the median (quintile) long-short portfolio of B&B generates 6.37% (9.15%), *ACCT* and *SYY* mimic 37.84% (58.14%) and 54.43% (64.26%) of the B&B pre-PEAD experience. Results in short legs during pre-PEAD show different price pattern. Although B&B significantly outperforms *ACCT's* short position over pre-PEAD, it fails to significantly beat the overpricing stocks under *SYY* and *SYYM* specifications. *SYYM's* short position even rivals B&B over pre-PEAD window. Such finding suggests that the incorporation of additional anomalies such as stock issuances-related and momentum and updating information over the holding period enhances mispricing measure's ability to capture subsequent negative returns in overpriced group. This asymmetry between long and short legs also suggests that shorting overpriced stocks using monthly refreshed mispricing signals could be an effective and implementable strategy to achieve even beat perfect foresight on accounting earnings.

B&B loses its advantage in the PEAD window as differences between B&B and mispricing portfolios shrink and lose significance. It reflects that B&B portfolios derive their "god-like" advantage from perfect foresight of earnings shocks which disappears once those earnings are announced, whereas mispricing portfolios by contrast continue to grow post announcements. Interestingly, the quintile *SYYM* long-short portfolio even outperforms B&B during the PEAD window. The reported negative difference indicates that B&B lags mispricing once mispricing refresh information and update portfolio holdings monthly. This result suggests that foresight dominates prior to announcements but dynamic mispricing strategies come back and surpass foresight portfolios once the "god-like" power becomes publicly accessible. The success of perfect foresight over pre-PEAD period is what the EMH expects.

As our result shows that perfect foresight portfolios can be rivalled by mispricing portfolios in certain specifications, the implementability of mispricing portfolios is worth noting, especially the outperforming strategy requires monthly updated holdings and short selling. Panel B demonstrates the results when the smallest tercile of firms in each month is removed. The overall patterns hold, which indicates that the outperformance of mispricing strategies especially the short position is not simply an artefact of small-cap illiquid stocks but reflect implementable return predictabilities.

[Insert Table 9 about here]

4.8 Why investors do not immediately trade on mispricing signals

As the previous sections demonstrate that mispricing signals are effective in predicting future earnings shocks and somehow replicate the returns of B&B perfect foresight portfolios over the pre-PEAD window, according to EMH, investors are expected to arbitrage away the associated abnormal returns of mispricing portfolios. Existing literature has documented that the abnormal return is significant, therefore it is worth asking whether all investors exploit his relationship for profit and who do not utilize it if not.

As mispricing signal is correlated with future earnings shocks, the profitability of mispricing strategies could be tied to the realization of subsequent earnings. As shown in Figure 3, the abnormal returns of mispricing portfolios are only meaningful when ex post earnings confirm the initial mispricing signal that is bad earnings shocks for overpriced stocks or good earnings shocks for underpriced stocks. If the realized earnings shocks do not materialize as what expected given by mispricing score, the portfolio delivers opposite outcomes. It indicates investors may need to not only recognize mispricing signal based on the current information but also correctly select stocks which will generate confirmative earnings.

[Insert Figure 3 about here]

The ability to correctly select stocks is not uniform across all investors. An investor who only holds one or a few stocks faces a binary outcome which is either the selection is correct, generating high abnormal returns or it does not, leading to big losses. If so, investment in mispriced portfolios becomes a risky game. By contrast, an investor capable of holding a greater number of stocks effectively averages the future earnings across firm-level variation. For example, assume that at a given cross-section the average predictions in Table 8 suggest there are 520 stocks with high mispricing predictions (anticipating good earnings) and 562 stocks with negative predictions (anticipating bad earnings), with 317 and 312 stocks, respectively, eventually confirming these predictions. Based on this empirical allocation of greater-than-zero predicted value, an investor limited to a portfolio of nine stocks drawn from the 520 high-prediction group has only a 75.32% probability of having five or more stocks that ultimately experience a good earnings shock (this selection process being "without replacement"). In contrast, an investor who can hold 90 stocks in the same group has an 87.64% probability of obtaining 50 or more stocks with confirmed good earnings shocks, and if the portfolio size increases to 900 stocks, the probability of securing at least 500 confirmed stocks rises to 99.96%[8].

The lower likelihood of achieving the desired proportion of earnings-confirmed stocks in less diversified portfolios reduces the willingness of these investors to trade solely on the basis of earnings predictions. Consequently, the initial stock prices do not fully reflect the information linking current accounting ratios to future earnings shocks. This incomplete incorporation of the predictive signal contributes to the persistence of the premium, as it compensates investors for the additional stock selection risk they incur when they cannot diversify as effectively.

To rigorously evaluate our hypothesis, we implement a simulation-based analysis that examines how the number of stocks in portfolio, serving as a proxy for investor scale, affects the accuracy of predictions based on historical accounting mispricing scores (*ACCT*). Specifically, at each cross section prior to a firm's reporting date for fiscal year *t*, stocks are randomly selected from the set defined by the *ACCT* scores. In our simulation, we assume that 2,000 investors are actively seeking investment opportunities within each mispriced stock group. Each investor randomly selects a portfolio containing a variable number of stocks, ranging from one up to the total number of available stocks in that cross section, and holds

[8] The calculation for the distribution-based probability is shown in Appendix A3

these stocks for the subsequent 12 months. At each reporting date, all 2,000 investors update their selections based on the latest available information.

For each portfolio constructed, we define three indicators of prediction accuracy. The first indicator is based on the count of earnings-confirmed stocks, that is a prediction is considered correct if the number of stocks whose earnings shocks confirm the mispricing signal exceeds the number of stocks that do not. The second indicator employs the average Ball and Brown unexpected earnings of the portfolio, assigning a value of one if the average is positive for underpriced stock selections (or negative for overpriced ones). The third criteria is based on the portfolio median of unexpected earnings. In addition, we classify these simulated investors according to the size of their portfolios: portfolios holding between 1 and 10 stocks are taken as a proxy for retail investors indicating that they can only hold undiversified portfolios; those with 10 to 50 stocks represent mutual or hedge funds; portfolios with 50 to 100 stocks approximate large mutual funds; and portfolios with more than 100 stocks stand in for index funds. Across the data sample, each mispricing portfolio is represented by 100,000 selections (derived from 50 fiscal years multiplied by 2,000 selections per year).

Table 10 summarizes the simulation results. In Panel A, where correct prediction is defined by the greater number of earnings-confirmed stocks relative to non-confirmed stocks, we observe that for portfolios comprising fewer than ten stocks, there are 4,511 overpriced portfolios and 4,608 underpriced portfolios, with 72.64% and 59.64% chance to correctly predict earnings, respectively. When investors are allowed to hold more stocks in their portfolio (10 to 50), the correction rate for overpriced and underpriced stocks becomes 79.54% and 82.89% respectively, which are significantly higher than the probability of portfolios only consists of 0 to 10 stocks. Panel B and C report the result for different identification of correct prediction. Interestingly, the least diversified portfolios seem better at utilizing underpriced stocks than overpriced stocks when considering mean and median earnings shocks of portfolios. Similarly, holding more stocks results in a significantly higher chance of correctly predicting earnings shocks on the portfolio level. Consequently, if such investors construct a long-short portfolio, the joint probability that both the long and short positions are correctly predicted is approximately 43%, a scenario corresponding to the highest profitability. In contrast, these less diversified investors face 46% chance of achieving a correct prediction on only one side (either long or short), and an 11% chance of incorrect predictions on both sides. Notably, as portfolio size increases, the likelihood of achieving correct predictions markedly improves. For instance, when investors hold more than 100 stocks per selection, 88.73% of the cases exhibit correct predictions on both long and short positions, with only 0.20% of cases resulting in incorrect predictions on both sides.

[Insert Table 10 about here]

One potential critique of the earnings-confirmation metric is that in less diversified portfolios, the impact of an additional confirmed stock might be disproportionately large relative to diversified portfolios. To address this issue, Panel B and C of Table 10 presents an alternative definition of correct prediction: a mispricing signal is deemed correct if the average earnings shock or the median of the portfolio aligns in the expected direction (positive for underpriced portfolios and negative for overpriced portfolios). The results remain consistent with those of Panel A. For portfolios with fewer than ten stocks, the probability that both the long and short sides are correctly predicted is 44.17% (for portfolio's average unexpected earnings) and 39.84% (for portfolio's median of unexpected earnings), with an 11.05% and 12.68% chance that both sides are incorrect. Increasing the portfolio size to include more stocks raises the correct prediction rate to 62.94% and 56.07% and reduces the probability of incorrect predictions to 3.35% and 2.84%. We, therefore, conclude that the number of stocks in portfolio are associated with a significantly higher likelihood of accurately capturing the predictive power of historical accounting mispricing signals. This finding suggests that investor scale plays a critical role in the degree to which market participants can effectively exploit the relationship between current accounting ratios and future earnings shocks.

Table 11 presents the time-series average monthly returns of all equally weighted portfolios adjusted for the Fama-French three-factor model and the Carhart momentum factor over a 12-month holding period for each stock selection, stratified by whether the portfolio's average earnings shock confirms the mispricing signal. In the case of overpriced portfolios, the risk-adjusted premium becomes increasingly negative as the number of stocks in the portfolio increases. For instance, portfolios comprising fewer than ten stocks yield a monthly return of -0.070% which fails to pass 10% significance level across time, whereas those with more than 100 stocks deliver a monthly return of -0.111% achieving a 10% significance. By contrast, underpriced portfolios exhibit similar returns across different portfolio diversification with 1% significance level, suggesting a greater degree of market efficiency for stocks identified as underpriced.

[Insert Table 11 about here]

Recognizing that return alone may not fully capture investor preference, we further compute the return-to-risk ratio for each group, which is the ratio of the time-series average of adjusted return to the time-series averaged standard deviation of returns, with the ratio's sign adjusted to ensure that higher values indicate more attractive outcomes. Within overpriced portfolios, those containing fewer than ten stocks have the lowest return-risk ratio (4.268), and this ratio increases as the number of stocks grows. This pattern reflects both higher average returns and lower return volatility in larger portfolios, supporting our hypothesis that smaller investors, who typically hold fewer stocks, require higher compensation for the risk of arbitraging accounting mispricing.

The second half of Table 11 examines the performance metrics including returns, risk, and return-risk ratios conditional on whether the portfolio's earnings shocks confirm the mispricing signal. For overpriced portfolios with fewer than ten stocks, selections associated with positive earnings shocks yield an average return of 0.173% at 5% significance, despite the portfolio being classified as overpriced. This outcome results in a negative sign-adjusted return-risk ratio, indicating that such a payoff is unattractive to investors. When combined with the probability data reported in Table 10, these results suggest that less diversified investors are less inclined to hold portfolios based on overpriced stocks, leading to a weaker price correction in this segment. A similar negative return-risk relationship is observed for underpriced portfolios when earnings shocks do not confirm the mispricing signal, again indicating low investor preference. Furthermore, as the number of stocks in the portfolio increases, the return-risk ratios generally become positive.

Comparing the return to risk ratio of overpriced and underpriced portfolio with the smallest number of stocks, investors are less willing to trade overpriced stocks than underpriced stocks as the ratio of overpriced stocks conditional on non-confirmative earnings is more negative, suggesting that investors feel less motivated to short these stocks because of the significantly positive returns once earnings are not negatively shocking. It potentially explains why overpriced stocks are less arbitraged and generating more significant returns in general than underpriced stocks are.

In summary, our findings indicate that investors can improve the accuracy of earnings shock predictions and the performance of *ACCT*-based portfolios by holding a larger number of stocks. Portfolios with a smaller number of stocks are associated with lower investor preference and necessitate a higher risk premium. This lower propensity to trade on accounting-based predictions among small investors contributes to an incomplete initial price adjustment relative to the fundamental values implied by the relationship between current accounting ratios and future earnings shocks. Notably, this phenomenon is more pronounced for overpriced stocks, suggesting that the future earnings premium is more likely to manifest in these securities.

We further run a panel regression of the correct rate of each stock selection on average *ACCT* score of portfolios, log value of the number stock in the portfolio and the interaction. The equation is as follows:

$$Correct_{p,t} = a + b_1 ACCT_{p,t-1} * Log(No.stocks)_{i,t-1} + b_2 ACCT_{p,t-1}$$

$$+ b_3 Log(No.stocks)_{p,t-1} + controls_{p,t-1} + \delta_{t-1} + \rho_{diversification} + \varepsilon_{p,t}\ , \qquad (6)$$

where the correct $Correct_{p,t}$ is defined as the ratio of earnings-confirmed stocks to the total number of stocks in a given selection. In the panel regression, we control for the portfolio average of firm size, book-to-market ratio, investment-to-asset ratio, gross profitability, and historical returns, while including fixed effects for fiscal year and diversification level (i.e., the number of stocks). Standard errors are clustered by diversification level. Table 12 reports the estimated coefficients from this regression. Notably, the interaction term between the *ACCT* score and the number of stocks in the portfolio is positively correlated with the correct rate, indicating that investors who base their trading decisions on *ACCT* scores and hold more stocks are more likely to have a higher proportion of earnings-confirmed stocks in their portfolios. Furthermore, the coefficient on the *ACCT* score is negative for overpriced groups and positive for underpriced groups. In overpriced portfolios, a higher correct rate implies a larger proportion of stocks with adverse earnings shocks. Given that the *ACCT* score is positively correlated with earnings shocks, higher *ACCT* scores in overpriced groups are associated with a diminished prediction accuracy which means they yield more good earnings shocks than expected. These results support our hypothesis that smaller investors, who typically hold fewer stocks, are less inclined to trade on accounting mispricing signals. This lack of motivation among small investors results in lower levels of arbitrage, thereby allowing the mispricing premium to persist in the market.

[Insert Table 12 about here]

4.9 The effect of market efficiency on the mispricing information

Our previous analysis shows that the historical mispricing score contains information about future earnings shocks. It is an information element that can be viewed as fundamentally driven, as these earnings outcomes occur within the mispricing portfolio's holding period. However, earnings-related information represents only part of what the mispricing score captures. The remaining component (non-earnings-related information) reflect factors such as arbitrage risk (Stambaugh, Yu and Yuan, 2015), managerial mis-valuation-motived equity issuance (Ritter, 1991; Loughran and Ritter, 1995), and disposition effect driving momentum. Conceptually, the overall mispricing signal can be viewed as a basket containing both earnings-related (benchmark) and non-earnings-related (floating) information, where the benchmark provides price changes and floating provides friction. If we think about momentum, Jegadeesh and Titman (1993) find that quarterly announced earnings play important role in the momentum premium, while Grinblatt and Han (2005) emphasize the role of the disposition effect as a behavioral source of delayed price adjustment. Put momentum and its explanations in our setting, quarterly earnings are the benchmark information providing sufficient room for price movement, whereas the disposition effect (floating component) induces delayed reaction. The predictive power of the benchmark component with respect to future earnings shocks does not violate the EMH, as it merely compensates investors for bearing the risk of forecasting future fundamentals, but the information intensity of floating component may vary across different levels of market efficiency. In more efficient markets, the floating part fades, leaving the fundamental part dominant. As a result, the overall mispricing score should more closely reflect earnings-related information, leading to a stronger observed relationship between mispricing score and subsequent earnings shocks. In this section, the proxies we select for market efficiency are analyst coverage (Engelberg, Mclean and Pontiff, 2020; Chen et al., 2023), institutional coverage (D'Avolio, 2002; Collins, Gong and Hribar 2003; Calluzzo, Moneta and Topaloglu, 2019; Gao and Wang, 2022) and investor sentiment (Stambaugh, Yu and Yuan, 2012).

Institutional and analyst coverage

Institutional ownership facilitates equity lending, making stocks with high institutional holdings more "short-able" and thereby enhancing market efficiency (D'Avolio, 2002). Institutional investors also actively engage in arbitrage transactions based on anomaly signals, which further mitigates mispricing (Collins, Gong and Hribar 2003; Calluzza, Moneta, Topaloglu, 2019; Gao and Wang, 2022). We investigate the relationship between current mispricing information and unexpected earnings for the next fiscal year for stocks with higher percentage of equity held by institutional investors. For each firm's fourth quarter announcement month, we classify stocks with the same fiscal year into high- and low-institutional-holding groups. Institutional holdings are measured using the most recent quarterly data from IBES, and firms with institutional ownership above the sample median are defined as having high institutional coverage. Similarly, Chen et al. (2023) show that analyst coverage improves market efficiency by reducing investor's underreaction to information. Motivated by this finding, we also investigate the different predictive power of mispricing information between firms with high and low analyst coverage. For each firm's fourth-quarter announcement month, we classify stocks within the same fiscal year into high- and low-coverage groups. Analyst coverage is proxied by the number of recommendations issued before the fourth quarter announcement month of year *t-1*. Missing values are treated as zero. Firms with a number of recommendations above the cross-section median are defined as having high analyst coverage. We then estimate the following panel regression model to test whether the predictive relationship between mispricing and subsequent earnings shocks differs across the two groups.

$$Unexp\Delta EPS_{j,t} = a + b_1 MISP_{j,t-1} + b_2 Coverage_{j,t-1} + b_3 MISP_{j,t-1} * Coverage_{j,t-1}$$

$$+ \sum_{i=1}^{n} \gamma_i control_{i,j,t-1} + \delta_{t-1} + \rho_j + \varepsilon_{j,t-1} \,, \tag{7}$$

where $Coverage_{j,t-1}$ is the dummy variable which equals one when the stock is identified as high institutional or analyst coverage. $MISP_{j,t-1} * Coverage_{j,t-1}$ is the interaction between stock's mispricing information and high coverage. Control variables include market capitalization, short-, medium- and long-term returns before the fourth quarter announcement month, book-to-market, investment-to-asset and gross profit ratios. It also controls fiscal year and firm fixed effects, and all standard errors are clustered by fiscal year and firm level.

Table 13 reports the correlation coefficients for all variables in Equation (7). For the *SYY* mispricing score, both high institutional ownership and high analyst coverage significantly enhance the predictive power of mispricing information for future unexpected earnings. This finding suggests that the *SYY* mispricing score becomes more closely associated with earnings shocks when the market is more efficient. The result aligns with our conjecture that improved market efficiency amplifies the benchmark (fundamental-related) component of mispricing information by compressing its floating (non-fundamental) part, thereby yielding a stronger link between mispricing and subsequent earnings shocks. In contrast, for the mispricing information (*ACCT*) derived solely from annual accounting ratios, the effect of high coverage is insignificant. As *ACCT* mispricing information is inherently dominated by accounting-based fundamental information, its informational content is unaffected by differences in market efficiency.

[Insert Table 13 about here]

Investor sentiment

Stambaugh, Yu and Yuan (2012) find that the mispricing long-short portfolio generates higher premium in the months following high sentiment period. This evidence implies that market efficiency deteriorates during high-sentiment periods: stocks become more heavily mispriced, particularly on the overpriced side, and the subsequent correction of the mispricing becomes larger in the following month. Following our argument regarding information volatility, we expect that earnings prediction of mispricing information after a high-sentiment (and thus less efficient) period should be less accurate than those after a period of greater market efficiency. We utilize Baker and Wurgler (2007) investor sentiment index as a proxy for market efficiency. For each firm at each fourth quarter announcement date of each fiscal year, we calculate the average BW sentiment over the past *j* months prior to the announcement month plus the announcement month – that is (*j*+*1*)-month period from announcement month minus *j* to the announcement month, where higher value in the average sentiment score indicates the firm's most recent information environment is less efficient over the period.

We then estimate the following panel regression model to examine whether the predictive relationship between mispricing and subsequent earnings shocks varies conditional on BW sentiment as follows

$$Unexp\Delta EPS_{j,t} = a + b_1 MISP_{j,t-1} + b_2 BW_{j,t-1_{rdq-j,rdq}} + b_3 MISP_{j,t-1} * BW_{j,t-1_{rdq-j,rdq}}$$

$$+ \sum_{i=1}^{n} \gamma_i control_{i,j,t-1} + \delta_{t-1} + \rho_j + \varepsilon_{j,t-1}\ , \qquad (8)$$

where $BW_{j,t-1_{rdq-j,rdq}}$ is the average BW sentiment over the past *j* months plus the announcement month, and *j* takes value of 0, 2, 5 and 11 for the consideration of investor sentiment over the most recent 1-, 3-, 6-, and 12-month window, respectively. The interaction term $MISP_{j,t-1} * BW_{j,t-1_{rdq-j,rdq}}$ captures how the relationship between stock's mispricing information and subsequent earnings shocks varies with market sentiment. All other model specifications follow Equation (7). The results reported in Table 14 indicate that short- and medium-term sentiment (i.e. 1-, 3- or 6-month period) does not significantly alter the predictive relationship between mispricing information and earnings surprise. However, when the sentiment is measured over a 12-month horizon, higher sentiment significantly weakens the predictive power of the mispricing information on future earnings. The evidence is consistent with our conjecture that the prediction relationship of mispricing information on unexpected earnings of the next fiscal year is stronger when the prediction is formed following a period with lower sentiment and enhanced efficiency.

[Insert Table 14 about here]

5. Conclusion

This study builds on the seminal visual presentation by Ball and Brown (1968), as well as the established literature on post-earnings announcement drift (PEAD) and its less-explored pre-earnings announcement drift (pre-PEAD). While PEAD is understood as evidence of market anomaly after earnings surprises, pre-PEAD reflects a potential for huge abnormal returns prior to the announcement which is what EMH expects, contingent on the accurate anticipation of earnings outcomes. If earnings are the only information under the circumstances, pre-PEAD returns require perfect foresight, an unrealistic "god-like" power in real-world investment. However, drawing from the earnings prediction literature, we join the discussion by examining whether portfolios formed by publicly available mispricing information, including anomalies derived from annual accounting ratios and accounting-price-volume combined information can asymmetrically capture different subsequent earnings surprises.

Empirically, we find a relationship that firms exhibit stronger underpricing signals measured via available mispricing scores tend to deliver more favorable earnings surprises in the subsequent reporting period. This predictive power of mispricing scores on future earnings holds across multiple settings, including alternative timing conventions for mispricing measurement (either at the end of June or earlier at the fourth-quarter announcement month of the prior fiscal year), as well as both in-sample and out-of-sample tests. From investor's application perspective, we find that portfolios formed by the available mispricing information mimic pre-PEAD returns, without the unrealistic requirement of perfect foresight. However, the return of mispricing portfolios is dependent on real future earnings surprise. It means that mispricing premium is high when the initial mispricing signal is confirmed by subsequent earnings announcement, suggesting that investor's motivation to trade stocks based on mispricing scores depends on their chance to have a correct earnings allocation in their portfolio. Through a random stock selection experiment, we observe that investors with highly concentrated portfolios (holding a small number of stocks) are significantly less inclined to act on accounting-based mispricing signals. This is because concentrated portfolios face a higher probability of erroneous earnings forecasts on the portfolio level, and thus a higher risk and lower return of mispricing strategies. Conversely, diversified portfolios can better predict the portfolio averaged earnings, making the subsequent returns more alpha-like.

Using institutional ownership, analyst coverage and investor sentiment as proxies for market efficiency, we also find that the predictive power of mispricing information on the unexpected earnings of the next fiscal year is stronger when the market is more efficient. Our interpretation is that greater market efficiency reduces the influence of non-fundamental (arbitrage-driven) components embedded in mispricing measures and increases the relative weight of fundamental-related (earnings) information. As a result, the mispricing signal becomes more closely aligned with true earnings-related fundamentals, thereby strengthening its ability to predict subsequent earnings shocks.

Reference

Altman, E. I. (1983). Why Businesses Fail. *Journal of Business Strategy*, 3(4), 15-21.

Abarbanell, J. S., & Bushee, B. J. (1998). Abnormal Returns to A Fundamental Analysis Strategy. *Accounting Review*, 19-45.

Asness, C. S. (1995). The Power of Past Stock Returns to Explain Future Stock Returns. *Working Paper.*

Ball, R., & Brown, P. (1968) An Empirical Evaluation of Accounting Income Numbers. *Journal of Accounting Research*, 6(2), 159-178.

Ball, R., & Brown, P. (2019). Ball and Brown (1968) After Fifty Years. *Pacific-Basin Finance Journal*, 53, 410-431.

Basu, S. (1977). Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis. *Journal of Finance*, 32(3), 663-682.

Bernard, V. L., & Thomas, J. K. (1989). Post-Earnings-Announcement Drift: Delayed Price Response or Risk Premium? *Journal of Accounting Research*, 27, 1-36.

Blitz, D., Huij, J., & Martens, M. (2011). Residual Momentum. *Journal of Empirical Finance*, 18(3), 506-521.

Boudoukh, J., Michaely, R., Richardson, M., & Roberts, M. R. (2007). On The Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing. *Journal of Finance*, 62(2), 877-915.

Calluzzo, P., Moneta, F., & Topaloglu, S. (2019). When Anomalies Are Publicized Broadly, Do Institutions Trade Accordingly? *Management Science*, 65(10), 4555-4574.

Campbell, J. Y., Hilscher, J., & Szilagyi, J. (2008). In Search of Distress Risk. *Journal of Finance*, 63(6), 2899-2939.

Cao, S. S., & Narayanamoorthy, G. S. (2012). Earnings Volatility, Post-Earnings Announcement Drift, and Trading Frictions. *Journal of Accounting Research*, 50(1), 41-74.

Chen, L., Novy-Marx, R., & Zhang, L. (2011). An Alternative Three-Factor Model. *Working Paper.*

Chen, X., He, W., Tao, L., & Yu, J. (2023). Attention and Underreaction-Related Anomalies. *Management Science*, 69(1), 636-659.

Collins, D. W., Gong, G., & Hribar, P. (2003). Investor Sophistication and The Mispricing of Accruals. *Review of Accounting Studies*, 8(2), 251-276.

Cooper, M. J., Gulen, H., & Schill, M. J. (2008). Asset Growth and The Cross-Section of Stock Returns. *Journal of Finance*, 63(4), 1609-1651.

Daniel, K., & Titman, S. (2006). Market Reactions to Tangible and Intangible Information. *Journal of Finance*, 61(4), 1605-1643.

D'Avolio, G. (2002). The Market for Borrowing Stock. *Journal of Financial Economics*, *66*(2-3), 271-306.

Dichev, I. D., & Tang, V. W. (2009). Earnings Volatility and Earnings Predictability. *Journal of Accounting and Economics*, 47(1-2), 160-181.

Engelberg, J., McLean, R. D., & Pontiff, J. (2020). Analysts and Anomalies. *Journal of Accounting and Economics*, 69(1), 101249.

Fama, E. F. (1970). Efficient Capital Markets. *Journal of Finance*, 25(2), 383-417.

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *Journal of Finance*, 47(2), 427-465.

Fama, E. F., & French, K. R. (1996). Multifactor Explanations of Asset Pricing Anomalies. *Journal of Finance*, 51(1), 55-84.

Fama, E. F., & French, K. R. (2006). Profitability, Investment and Average Returns. *Journal of Financial Economics*, 82(3), 491-518.

Fama, E. F., & French, K. R. (2008). Dissecting Anomalies. *Journal of Finance*, 63(4), 1653-1678.

Fama, E. F., & French, K. R. (2015). A Five-Factor Asset Pricing Model. *Journal of Financial Economics*, 116(1), 1-22.

Foster, G., Olsen, C., & Shevlin, T. (1984). Earnings Releases, Anomalies, And the Behavior of Security Returns. *Accounting Review*, 574-603.

Frankel, R., & Litov, L. (2009). Earnings Persistence. *Journal of Accounting and Economics*, 47(1-2), 182-190.

Gao, X., & Wang, Y. (2023). Mining The Short Side: Institutional Investors and Stock Market Anomalies. *Journal of Financial and Quantitative Analysis*, *58*(1), 392-418.

Grinblatt, M., & Han, B. (2005). Prospect Theory, Mental Accounting, and Momentum. *Journal of Financial Economics*, 78(2), 311-339.

Haugen, R. A., & Baker, N. L. (1996). Commonality in the Determinants of Expected Stock Returns. *Journal of Financial Economics*,*41*(3), 401-439.

Hirshleifer, D., Hou, K., Teoh, S. H., & Zhang, Y. (2004). Do Investors Overvalue Firms with Bloated Balance Sheets? *Journal of Accounting and Economics*, 38, 297-331.

Hou, K., & Moskowitz, T. J. (2005). Market Frictions, Price Delay, And the Cross-Section of Expected Returns. *The Review of Financial Studies*, 18(3), 981-1020.

Jegadeesh, N., & Titman, S. (1993). Returns To Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance*, 48(1), 65-91.

Kaniel, R., Liu, S., Saar, G., & Titman, S. (2012). Individual Investor Trading and Return Patterns Around Earnings Announcements. *Journal of Finance*, 67(2), 639-680.

Loughran, T., & Ritter, J. R. (1995). The New Issues Puzzle. *Journal of Finance*, 50(1), 23-51.

Lyandres, E., Sun, L., & Zhang, L. (2008). The New Issues Puzzle: Testing the Investment-Based Explanation. *Review of Financial Studies*,21(6), 2825-2855.

Moskowitz, T. J., & Grinblatt, M. (1999). Do Industries Explain Momentum? *Journal of Finance*, 54(4), 1249-1290.

Ng, J., Rusticus, T. O., & Verdi, R. S. (2008). Implications of Transaction Costs for The Post–Earnings Announcement Drift. *Journal of Accounting Research*, 46(3), 661-696.

Novy-Marx, R. (2013). The Other Side of Value: The Gross Profitability Premium. *Journal of Financial Economics*, 108(1), 1-28.

Ohlson, J. A. (1980). Financial Ratios and The Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*.

Ou, J. A., & Penman, S. H. (1989a). Accounting Measurement, Price-Earnings Ratio, and The Information Content of Security Prices. *Journal of Accounting Research*, *27*, 111-144.

Ou, J. A., & Penman, S. H. (1989b). Financial Statement Analysis and The Prediction of Stock Returns. *Journal of Accounting and Economics*, 11(4), 295-329.

Penman, S. H., & Zhu, J. L. (2014). Accounting Anomalies, Risk, and Return. *Accounting Review*, 89(5), 1835-1866.

Richardson, S. A., Sloan, R. G., Soliman, M. T., & Tuna, I. (2005). Accrual Reliability, Earnings Persistence and Stock Prices. *Journal of Accounting and Economics*, 39(3), 437-485.

Ritter, J. R. (1991). The Long-Run Performance of Initial Public Offerings. *Journal of Finance*, 46(1), 3-27.

Sloan, R. G. (1996). Do Stock Prices Fully Reflect Information in Accruals and Cash Flows About Future Earnings? *Accounting Review*, 289-315.

Soliman, M. T. (2008). The Use of DuPont Analysis by Market Participants. *Accounting Review*, 83(3), 823-853.

Stambaugh, R. F., Yu, J., & Yuan, Y. (2012). The Short of It: Investor Sentiment and Anomalies. *Journal of Financial Economics*, *104*(2), 288-302.

Stambaugh, R. F., Yu, J., & Yuan, Y. (2015). Arbitrage Asymmetry and The Idiosyncratic Volatility Puzzle. *Journal of Finance*, 70(5), 1903-1948.

Titman, S., Wei, K. J., & Xie, F. (2004). Capital Investments and Stock Returns. *Journal of Financial and Quantitative Analysis*, 39(4), 677-700.

Welch, I., & Goyal, A. (2008). A Comprehensive Look at the Empirical Performance of Equity Premium Prediction. *Review of Financial Studies*, *21*(4), 1455-1508.

Xing, Y. (2008). Interpreting The Value Effect Through The Q-Theory: An Empirical Investigation. *Review of Financial Studies*, 21(4), 1767-1795.

Table 1 Summary statistics

This table reports the summary statistics of foresight unexpected change in EPS for the fiscal year *t* (Ball and Brown, 1968), the Stambaugh, Yu and Yuan's mispricing scores (*SYY*), and 11 anomalies to construct mispricing scores (Stambaugh et al., 2012, 2015) at *t-1*. Each of 11 anomaly indicators for each single stock are available at the reporting date of the fourth quarter of fiscal year *t-1.* As the *SYY* mispricing measure is an aggregated cross-sectional rank among these 11 anomalies, it requires anomaly indicators of all stocks at cross sections. Therefore, portfolio formation month of all 11 anomalies and *SYY* mispricing score is the end of June to ensure the cross-sectional data availability. The Ball and Brown's unexpected earnings are a foresight accounting ratio for the next fiscal year *t*, and it is unknown when the mispricing related anomalies are available. *Panel A* demonstrates statistical outcomes for each anomaly indicator. *Panel B* reports correlations between each pair of these anomaly indicators. All outcomes presented in this table are time-series averages of cross-sectional statistic result. The sample includes firms with a December fiscal year-end and sufficient accounting data to calculate Ball and Brown's earnings surprise and *SYY* measure over the period from 1974 to 2024.

Panel A: Cross-Sectional Summary Statistics (Averaged Over Time)

	B&B's unexpected earnings (*t*)	*SYY* mispricing score *(t-1)*	Financial distress *(t-1)*	O-score *(t-1)*	NSI *(t-1)*	CEI *(t-1)*	Accruals *(t-1)*	NOA *(t-1)*	Momentum *(t-1)*	Gross profitability *(t-1)*	Asset growth *(t-1)*	Return on asset *(t-1)*	Investment to asset *(t-1)*
Mean	-0.049	49.453	4.530	7.709	-0.297	0.553	0.036	-0.628	0.139	0.317	-0.107	0.006	-0.067
Std Dev	3.264	12.718	2.351	2.022	0.975	1.003	0.067	0.261	0.398	0.236	0.260	0.031	0.125
25th	-0.694	40.776	3.538	6.809	-0.246	0.082	0.005	-0.758	-0.099	0.159	-0.149	0.001	-0.103
Median	0.049	49.913	5.241	7.732	-0.051	0.329	0.036	-0.641	0.084	0.288	-0.061	0.01	-0.052
75th	0.741	58.624	6.207	8.856	0.021	0.71	0.068	-0.497	0.293	0.434	0.009	0.019	-0.013

Panel B: Cross-Sectional Correlation Matrix (Averaged Over Time)

	B&B's unexpected earnings (*t*)	*SYY* mispricing score *(t-1)*	Financial distress *(t-1)*	O-score *(t-1)*	NSI *(t-1)*	CEI *(t-1)*	Accruals *(t-1)*	NOA *(t-1)*	Momentum *(t-1)*	Gross profitability *(t-1)*	Asset growth *(t-1)*	Return on asset *(t-1)*	Investment to asset *(t-1)*
B&B's unexp. EPS (*t*)	1												
SYY	0.085	1											
Financial distress	-0.020	0.102	1										
O-score	-0.104	0.432	0.107	1									
NSI	0.004	0.436	0.060	0.223	1								
CEI	-0.003	0.349	0.022	0.147	0.287	1							
Accruals	0.066	0.294	-0.036	-0.040	0.027	-0.051	1						
NOA	0.077	0.385	-0.083	-0.190	0.067	0.008	0.174	1					
Momentum	0.079	0.320	0.021	0.088	0.048	0.215	0.002	0.006	1				
GP	0.001	0.410	-0.011	0.274	0.234	0.177	-0.036	0.034	0.062	1			
AG	0.101	0.404	-0.06	-0.091	0.174	-0.068	0.249	0.465	-0.057	0.031	1		
ROA	0.061	0.436	0.131	0.606	0.23	0.225	-0.083	-0.156	0.189	0.328	-0.082	1	
IA	0.089	0.442	-0.061	-0.084	0.114	-0.031	0.194	0.501	0.012	0.031	0.638	-0.089	1

Table 2 Inter-portfolio summary statistics: historical *SYY* mispricing quintile portfolios and foresight unexpected EPS portfolios

This table reports statistical outcomes of *SYY* quintile portfolios, Ball and Brown's unexpected earnings portfolios and intersections based on their foresight unexpected earnings and mispricing scores. Stocks are independently ranked and sorted at each cross section based on their historical mispricing scores *SYY* at *t-1* into quintile portfolios and unexpected earnings at *t* into two groups. Stocks are identified as overpriced (underpriced) when the mispricing score is in the bottom (top) quintile. The foresight Ball and Brown's unexpected earnings into two portfolios, where stocks are identified as high (low) earnings stocks when the unexpected earnings are above (below) the cross-sectional median or positive (negative). The sample includes firms with a December fiscal year-end and sufficient accounting data to calculate Ball and Brown's earnings surprise and *SYY* measures over the period from 1974 to 2024. We calculate the time series t-statistics and denote statistical significance with asterisks. *Significance at the 10% level; **significance at the 5% level; ***significance at the 1% level.

Portfolio		No. obs	Ln MCAP	Idiosyncratic volatility	Market beta	Unexpected earnings *(t)* mean	Unexpected earnings *(t)* median	Unexpected earnings *(t)* Std.Dev.
		Panel A Mispricing score (*t-1*), *SYY* (Quintiles)						
Overpriced		215***	12.142***	0.027***	1.131***	-0.428***	-0.091***	2.914***
2		216***	12.576***	0.022***	1.004***	-0.073	0.005	3.024***
3		215***	12.903***	0.020***	0.939***	0.010	0.065***	3.022***
4		215***	13.299***	0.017***	0.897***	0.170***	0.114***	2.947***
Underpriced		215***	13.834***	0.015***	0.819***	0.269***	0.184***	3.101***
Diff. (Under - Over)		0.000	1.692***	-0.011***	-0.312***	0.697***	0.275***	0.187**
Mispricing Score *(t)*	Unexpected earnings *(t)*	Panel B *SYY* (*t-1*), (Quintiles) * B&B's unexpected earnings (*t*) (Cutoff: Median)						
Overpriced	Low	125***	12.104***	0.027***	1.139***	-1.723***	-0.715***	2.674***
	High	90***	12.205***	0.027***	1.119***	1.348***	0.623***	2.102***
	High-Low	-35***	0.101**	0.000	-0.020	3.071***	1.338***	-0.572***
2	Low	113***	12.523***	0.022***	0.997***	-1.64***	-0.712***	2.552***
	High	102***	12.636***	0.022***	1.012***	1.653***	0.749***	2.426***
3	Low	105***	12.843***	0.02***	0.928***	-1.675***	-0.722***	2.611***
	High	110***	12.96***	0.02***	0.951***	1.626***	0.752***	2.386***
4	Low	101***	13.283***	0.017***	0.89***	-1.597***	-0.732***	2.369***
	High	115***	13.305***	0.017***	0.902***	1.716***	0.824***	2.42***
Underpriced	Low	93***	13.752***	0.015***	0.812***	-1.669***	-0.795***	2.531***
	High	122***	13.884***	0.015***	0.824***	1.779***	0.849***	2.539***
	High-Low	29***	0.132**	0.000	0.012	3.448***	1.644***	0.008
Mispricing Score *(t)*	Unexpected earnings *(t)*	Panel C *SYY* (*t-1*), (Quintiles) * B&B's unexpected earnings (*t*) (Cutoff: 0)						
Overpriced	Low	118***	12.109***	0.027***	1.146***	-1.817***	-0.780***	2.716***
	High	97***	12.184***	0.027***	1.114***	1.255***	0.547***	2.054***
	High-Low	-21**	0.075*	0.000	-0.032**	3.072***	1.328***	-0.662***
2	Low	108***	12.529***	0.022***	0.999***	-1.719***	-0.781***	2.585***
	High	108***	12.615***	0.022***	1.010***	1.571***	0.672***	2.396***
3	Low	100***	12.882***	0.02***	0.929***	-1.765***	-0.798***	2.65***
	High	115***	12.927***	0.02***	0.946***	1.565***	0.692***	2.363***
4	Low	96***	13.301***	0.017***	0.894***	-1.677***	-0.791***	2.4***
	High	119***	13.293***	0.017***	0.898***	1.648***	0.763***	2.399***
Underpriced	Low	89***	13.783***	0.015***	0.814***	-1.757***	-0.866***	2.574***
	High	126***	13.859***	0.015***	0.821***	1.728***	0.814***	2.525***
	High-Low	37***	0.075	0.000	0.007	3.485***	1.679***	-0.050

Table 3 OLS Panel regression: prediction on unexpected earnings using *SYY* mispricing

This table presents the results of OLS panel regressions examining the relationship between foresight unexpected earnings, as defined by Ball and Brown (1968), and the *SYY* mispricing score developed by Stambaugh, Yu and Yuan (2012 and 2015). The dependent variable, unexpected earnings for year *t*, is regressed on *SYY* mispricing score which is constructed by publicly available information at the end of June in the previous year and a set of control variables. For the first column, the baseline control variables include the earnings per share for the previous fiscal year, natural logarithm of market capitalization, past stock returns over short-, medium-, and long-term horizons, and accounting controls, including book-to-market, investment-to-assets, and gross profitability. The regression model for column two and three controls for analyst recommendation- and institutional holding-related variables, where the model for the fourth column put all control variables together. The *SYY* variable is constructed as the average cross-sectional rank of 11 anomalies used in the mispricing metrics of Stambaugh, Yu, and Yuan (2012, 2015). The sample period covers 1974 to 2024 for baseline model, 1980 to 2024 for institutional holding data, and 1994 to 2024 for analyst recommendation data. The table reports estimated coefficients along with clustered t-statistics on fiscal year and firm. *Significance at the 10% level; **significance at the 5% level; ***significance at the 1% level.

	Ball and Brown's raw value *(t)*			
	(1)	(2)	(3)	(4)
Predictors *(t-1)*	Baseline	Recommendation	Institutional holding	All
SYY mispricing	0.152***	0.165***	0.142***	0.164***
	(11.220)	(9.204)	(10.402)	(9.108)
Previous EPS	-0.612***	-0.625***	-0.622***	-0.626***
	(-25.426)	(-16.688)	(-24.052)	(-16.750)
MCAP	0.170***	0.314***	0.189***	0.325***
	(5.238)	(6.499)	(5.131)	(6.930)
Ret [-1,0]	0.047***	0.031***	0.043***	0.033***
	(6.175)	(2.853)	(5.712)	(3.113)
Ret [-12,-1]	0.057***	0.035**	0.054***	0.039***
	(6.441)	(2.516)	(6.102)	(2.807)
Ret [-60,-13]	0.008	0.004	0.016**	0.004
	(0.919)	(0.394)	(2.057)	(0.453)
BM	-0.136***	-0.167***	-0.144***	-0.168***
	(-8.512)	(-6.942)	(-8.603)	(-6.879)
Investment	0.030***	0.050***	0.033***	0.051***
	(3.167)	(4.412)	(3.326)	(4.662)
Profitability	0.012	0.011	0.020	0.009
	(0.824)	(0.587)	(1.363)	(0.468)
No. Rec		-0.099***		-0.098***
		(-3.508)		(-3.278)
Ave. Rec		-0.039***		-0.040***
		(-5.062)		(-5.163)
Std. Rec		-0.007		-0.007
		(-0.928)		(-0.939)
HHI Concentration			0.003	-0.019
			(0.230)	(-1.035)
Institutional holding			-0.010	-0.045*
			(-0.603)	(-1.906)
Change in institutional holding			-0.014**	-0.024***
			(-2.203)	(-2.954)
Adj. Squared R	0.248	0.246	0.257	0.248
No. Fiscal Years	51	31	45	31
No. Observations	52,602	29,013	46,445	28,759
Cross-Sectional Average No. Observations	1,031	936	1,032	928
FYEAR Fixed	Yes	Yes	Yes	Yes
Firm Fixed	Yes	Yes	Yes	Yes

Table 4 Logistic panel regression: prediction on unexpected earnings using *SYY* mispricing

This table presents the results of logistic panel regressions examining the relationship between foresight unexpected earnings, as defined by Ball and Brown (1968), and the *SYY* mispricing score developed by Stambaugh, Yu and Yuan (2012 and 2015). The dependent variable, which equals one when the unexpected earnings for year *t* is positive, otherwise zero. is regressed on *SYY* mispricing score which is constructed by publicly available information at the end of June in the previous year and a set of control variables. For the first column, the baseline control variables include the earnings per share for the previous fiscal year, natural logarithm of market capitalization, past stock returns over short-, medium-, and long-term horizons, and accounting controls, including book-to-market, investment-to-assets, and gross profitability. The regression model for column two and three controls for analyst recommendation- and institutional holding-related variables, where the model for the fourth column put all control variables together. The *SYY* variable is constructed as the average cross-sectional rank of 11 anomalies used in the mispricing metrics of Stambaugh, Yu, and Yuan (2012, 2015). The sample period covers 1974 to 2024 for baseline model, 1980 to 2024 for institutional holding data, and 1994 to 2024 for analyst recommendation data. The table reports estimated coefficients along with clustered t-statistics on fiscal year and firm. *Significance at the 10% level; **significance at the 5% level; ***significance at the 1% level.

	Dependent variable is 1 when firm's unexpected earnings for year *t* are positive			
	(1)	(2)	(3)	(4)
Predictors (*t-1*)	Baseline	Recommendation	Institutional holding	All
SYY mispricing	0.318***	0.335***	0.315***	0.334***
	(18.814)	(14.059)	(17.381)	(14.052)
Previous EPS	-0.637***	-0.629***	-0.646***	-0.629***
	(-32.707)	(-22.528)	(-29.942)	(-22.483)
MCAP	0.210***	0.275***	0.217***	0.293***
	(14.469)	(8.602)	(11.348)	(9.136)
Ret [-1,0]	0.114***	0.086***	0.112***	0.092***
	(10.411)	(5.475)	(9.378)	(5.712)
Ret [-12,-1]	0.168***	0.196***	0.175***	0.204***
	(11.163)	(8.672)	(10.686)	(8.871)
Ret [-60,-13]	-0.071***	-0.066***	-0.045***	-0.062***
	(-5.084)	(-3.622)	(-3.144)	(-3.379)
BM	-0.066***	-0.072***	-0.050***	-0.070***
	(-5.323)	(-3.960)	(-3.803)	(-3.849)
Investment	0.069***	0.087***	0.070***	0.089***
	(5.431)	(5.240)	(5.279)	(5.345)
Profitability	-0.065***	-0.074***	-0.073***	-0.070***
	(-4.565)	(-3.887)	(-4.858)	(-3.608)
No. Rec		-0.118***		-0.121***
		(-3.268)		(-3.322)
Ave. Rec		-0.059***		-0.061***
		(-4.285)		(-4.316)
Std. Rec		-0.012		-0.013
		(-0.668)		(-0.698)
HHI Concentration			0.035**	0.056**
			(2.371)	(2.211)
Institutional holding			-0.008	0.003
			(-0.552)	(0.134)
Change in institutional holding			-0.014	-0.036***
			(-1.341)	(-2.701)
Pseudo Squared R	0.067	0.071	0.067	0.071
No. Fiscal Years	51	31	45	31
No. Observations	52,602	29,013	46,445	28,759
Cross-Sectional Average No. Observations	1,031	936	1,032	928
FYEAR Fixed	Yes	Yes	Yes	Yes
Industry Fixed	Yes	Yes	Yes	Yes

Table 5 OLS Panel regression: prediction on unexpected earnings using *SYY* in announcement month

This table presents the results of OLS panel regressions examining the relationship between foresight unexpected earnings, as defined by Ball and Brown (1968), and the *SYY* mispricing score developed by Stambaugh, Yu and Yuan (2012 and 2015). The dependent variable, unexpected earnings for year *t*, is regressed on *SYY* mispricing score which is constructed by publicly available information at the end of announcement month of the previous fiscal year and a set of control variables. For the first column, the baseline control variables include the earnings per share for the previous fiscal year, natural logarithm of market capitalization, past stock returns over short-, medium-, and long-term horizons, and accounting controls, including book-to-market, investment-to-assets, and gross profitability. The regression model for column two and three controls for analyst recommendation- and institutional holding-related variables, where the model for the fourth column put all control variables together. The *SYY* variable is constructed as the average cross-sectional rank of 11 anomalies used in the mispricing metrics of Stambaugh, Yu, and Yuan (2012, 2015). The sample period covers 1974 to 2024 for baseline model, 1980 to 2024 for institutional holding data, and 1994 to 2024 for analyst recommendation data. The table reports estimated coefficients along with clustered t-statistics on fiscal year and firm. *Significance at the 10% level; **significance at the 5% level; ***significance at the 1% level.

	Ball and Brown's raw value *(t)*			
	(1)	(2)	(3)	(4)
Predictors (*t-1*)	Baseline	Recommendation	Institutional holding	All
SYY mispricing	0.082***	0.088***	0.127***	0.089***
	(6.836)	(5.706)	(9.683)	(5.577)
Previous EPS	-0.602***	-0.616***	-0.443***	-0.616***
	(-24.454)	(-15.894)	(-11.857)	(-15.842)
MCAP	0.098***	0.215***	0.117***	0.218***
	(2.824)	(3.829)	(6.049)	(3.650)
Ret [-1,0]	0.054***	0.059***	0.049***	0.060***
	(9.119)	(5.670)	(6.542)	(5.696)
Ret [-12,-1]	0.021***	0.006	0.015	0.004
	(2.713)	(0.491)	(1.215)	(0.379)
Ret [-60,-13]	0.008	0.001	0.020*	0.000
	(0.785)	(0.120)	(1.719)	(0.001)
BM	-0.148***	-0.200***	-0.059***	-0.206***
	(-9.371)	(-8.880)	(-6.081)	(-8.680)
Investment	-0.003	0.021*	0.012	0.022*
	(-0.277)	(1.675)	(1.346)	(1.821)
Profitability	0.044***	0.044**	-0.013	0.040**
	(3.192)	(2.424)	(-1.095)	(1.984)
No. Rec.		-0.061**		-0.074***
		(-2.277)		(-2.763)
Ave. Rec.		-0.021***		-0.021***
		(-2.716)		(-2.790)
Std. Rec.		-0.000		-0.005
		(-0.054)		(-0.752)
HHI Concentration			0.016*	-0.023
			(1.910)	(-1.575)
Institutional holding			-0.004	-0.043**
			(-0.470)	(-2.221)
Change in institutional holding			-0.000	-0.017**
			(-0.010)	(-1.962)
Adj. Squared R	0.230	0.229	0.157	0.230
No. Fiscal Years	51	31	45	31
No. Observations	54,326	29,448	46,738	29,061
Cross-Sectional Average No. Observations	1,065	950	1,039	937
FYEAR Fixed	Yes	Yes	Yes	Yes
Firm Fixed	Yes	Yes	Yes	Yes

Table 6 Logistic panel regression: prediction on unexpected earnings using *SYY* in announcement month

This table presents the results of logistic panel regressions examining the relationship between foresight unexpected earnings, as defined by Ball and Brown (1968), and the *SYY* mispricing score developed by Stambaugh, Yu and Yuan (2012 and 2015). The dependent variable, which equals one when the unexpected earnings for year *t* is positive, otherwise zero. is regressed on *SYY* mispricing score which is constructed by publicly available information at the end of announcement month of the previous fiscal year and a set of control variables. For the first column, the baseline control variables include the earnings per share for the previous fiscal year, natural logarithm of market capitalization, past stock returns over short-, medium-, and long-term horizons, and accounting controls, including book-to-market, investment-to-assets, and gross profitability. The regression model for column two and three controls for analyst recommendation- and institutional holding-related variables, where the model for the fourth column put all control variables together. The *SYY* variable is constructed as the average cross-sectional rank of 11 anomalies used in the mispricing metrics of Stambaugh, Yu, and Yuan (2012, 2015). The sample period covers 1974 to 2024 for baseline model, 1980 to 2024 for institutional holding data, and 1994 to 2024 for analyst recommendation data. The table reports estimated coefficients along with clustered t-statistics on fiscal year and firm. *Significance at the 10% level; **significance at the 5% level; ***significance at the 1% level.

	Dependent variable is 1 when firm's unexpected earnings for year *t* are positive			
	(1)	(2)	(3)	(4)
Predictors (*t-1*)	Baseline	Recommendation	Institutional holding	All
SYY mispricing	0.172***	0.175***	0.168***	0.177***
	(11.057)	(7.979)	(9.943)	(8.034)
Previous EPS	-0.576***	-0.574***	-0.579***	-0.569***
	(-31.224)	(-21.681)	(-27.755)	(-21.317)
MCAP	0.215***	0.313***	0.225***	0.306***
	(15.065)	(9.580)	(11.833)	(9.207)
Ret [-1,0]	0.116***	0.140***	0.125***	0.142***
	(11.402)	(9.311)	(11.087)	(9.330)
Ret [-12,-1]	0.012	0.041**	0.014	0.036*
	(0.909)	(2.210)	(0.954)	(1.903)
Ret [-60,-13]	-0.059***	-0.063***	-0.042***	-0.065***
	(-4.410)	(-3.555)	(-2.971)	(-3.538)
BM	-0.082***	-0.097***	-0.070***	-0.102***
	(-6.629)	(-5.149)	(-5.254)	(-5.307)
Investment	-0.011	0.017	-0.005	0.019
	(-0.864)	(1.055)	(-0.352)	(1.161)
Profitability	-0.010	-0.010	-0.021	-0.010
	(-0.732)	(-0.541)	(-1.409)	(-0.500)
No. Rec.		-0.109***		-0.103***
		(-3.141)		(-2.870)
Ave. Rec.		-0.014		-0.013
		(-1.037)		(-0.993)
Std. Rec.		-0.006		-0.006
		(-0.304)		(-0.294)
HHI Concentration			0.031**	0.018
			(2.147)	(0.739)
Inst. holding			-0.020	-0.005
			(-1.282)	(-0.228)
Change in inst.			0.004	-0.010
			(0.447)	(-0.765)
Pseudo Squared R	0.049	0.053	0.050	0.052
No. Fiscal Years	51	31	45	31
No. Observations	54,326	29,448	46,738	29,061
Cross-Sectional Average No. Observations	1,065	950	1,039	937
FYEAR Fixed	Yes	Yes	Yes	Yes
Industry Fixed	Yes	Yes	Yes	Yes

Table 7 Panel regression: prediction on unexpected earnings using accounting mispricing signal *ACCT*

This table presents the results of OLS and logistic panel regressions examining the relationship between foresight unexpected earnings, as defined by Ball and Brown (1968), and an aggregated accounting mispricing score (*ACCT*). The dependent variable, unexpected earnings for year *t*, is regressed on *ACCT*, unexpected earnings for fiscal year *t*, and a set of control variables. The control variables include the logarithm of market capitalization, past stock returns over short-, medium-, and long-term horizons, and accounting controls, including book-to-market, investment-to-assets, and gross profitability. The *ACCT* variable is constructed as the average cross-sectional rank of six annual accounting ratios used in the mispricing metrics of Stambaugh, Yu, and Yuan (2012, 2015). The dependent variable, unexpected earnings, is measured in two different ways: (1) raw unexpected earnings and (2) a dummy variable equal to one if unexpected earnings are positive. We apply OLS regression for the raw unexpected earnings regression and logistic model when the dependent variable is binary. Under each model, the accounting mispricing score *ACCT* and all control variables are calculated on information available in June in the previous year of unexpected earnings (*June*) and in the fourth quarter announcement month of the previous fiscal year (*RD-Q4*). The sample period covers 1974 to 2024. The table reports estimated coefficients of OLS model (Logistic model) along with clustered t-statistics on fiscal year and firm (industry). *Significance at the 10% level; **significance at the 5% level; ***significance at the 1% level.

	OLS model		Logistic model	
	(1)	(2)	(3)	(4)
Predictors (*t-1*)	*June*	*RD-Q4*	*June*	*RD-Q4*
ACCT mispricing	0.021**	0.026***	0.076***	0.090***
	(2.248)	(2.640)	(4.725)	(5.558)
Previous EPS	-0.598***	-0.590***	-0.569***	-0.535***
	(-25.163)	(-24.273)	(-30.849)	(-30.417)
MCAP	0.184***	0.103***	0.274***	0.241***
	(5.646)	(2.950)	(19.558)	(17.046)
Ret [-1,0]	0.051***	0.056***	0.119***	0.119***
	(6.567)	(9.355)	(10.748)	(11.640)
Ret [-12,-2]	0.097***	0.040***	0.251***	0.054***
	(10.564)	(4.646)	(17.106)	(4.193)
Ret [-60,-13]	0.013	0.011	-0.060***	-0.049***
	(1.492)	(1.091)	(-4.308)	(-3.687)
BM	-0.139***	-0.150***	-0.065***	-0.080***
	(-8.534)	(-9.344)	(-5.293)	(-6.468)
Investment	-0.024**	-0.022**	-0.033**	-0.033**
	(-2.544)	(-2.101)	(-2.363)	(-2.285)
Profitability	0.062***	0.065***	0.024*	0.024*
	(4.478)	(4.839)	(1.704)	(1.654)
Adj. Squared R	0.241	0.228		
Pseudo Squared R			0.061	0.048
No. Fiscal Years	51	51	51	51
No. Observations	52,602	54,326	52,602	54,326
Cross-Sectional Average No. Observations	1,031	1,065	1,031	1,065
FYEAR Fixed	Yes	Yes	Yes	Yes
Cross-Sectional Fixed	Firm	Firm	Industry	Industry

Table 8 Out of sample prediction based on *ACCT* and *SYY* mispricing score in the announcement month of fiscal year *t* versus actual *B&B*'s unexpected earnings in year *t*

This table presents the classification of future unexpected earnings, as defined by Ball and Brown (1968), based on predictive signals that can be interpreted as *SYY* mispricing signal and accounting mispricing signal *ACCT*, the aggregated accounting mispricing score based on six annual accounting ratios incorporated in Stambaugh, Yu and Yuan's mispricing measure (2012, 2015). Firms are classified into positive (+) and negative (-) signals based on predictions, where variable loadings are estimated using cross-sectional regressions with mispricing score *SYY* or *ACCT* and control variables of the baseline model in previous tables over the past five years. The average coefficient of each independent variable is used in the next year to predict unexpected earnings. All information used to estimate correlation coefficient is available in the month of fourth quarter announcement of the previous fiscal year. Realized unexpected earnings in year *t* are similarly categorized into positive and negative outcomes. The table reports the number of observations in each category averaged across fiscal years, along with total counts. The diagonal values represent correct classifications, while off-diagonal values indicate misclassifications. The last two columns show the significance level that the proportion of correct prediction is different from the actual binomial distribution of unexpected earnings realized during the period. *Significance at the 10% level; **significance at the 5% level; ***significance at the 1% level.

		Realized *B&B*'s unexpected EPS *(t)*					
Prediction *(t)* Based on *(t-5)* to *(t-1)*		+	-	Total	Prob. (+)	Prob. (-)	Is the prediction significantly better than the cross-sectional binomial distribution? (p-value)
SYY	+	317	203	520	60.95%	39.05%	0.000***
	-	250	312	562	44.40%	55.60%	0.000***
	Total	567	516	1082	52.36%	47.64%	
ACCT	+	314	203	517	60.70%	39.30%	0.000***
	-	253	312	565	44.73%	55.27%	0.000***
	Total	567	516	1082	52.36%	47.64%	

Table 9 Extent to which mispricing portfolios capture pre-PEAD and PEAD returns

This table presents the difference between cumulative monthly abnormal returns of of perfect foresight portfolios on B&B's unexpected earnings which are released at time 0 and mispricing portfolios that was formed at time t-12 to the abnormal returns. Monthly returns are adjusted using the Fama-French-Carhart four-factor model. The mispricing information refers to three mispricing scores (1) accounting mispricing score *(ACCT),* (2) mispricing score of *SYY* which we only obtain information at the beginning of the pre-PEAD period, and (3) a version of the *SYY* mispricing score that is updated monthly throughout the full period *(SYYM).* All mispricing signals are available across the entire observation window *(-12, 12).* Panel A reports the relative performance of these mispricing portfolios over the first six month of pre-PEAD, the entire pre-PEAD period and over the full pre-PEAD and PEAD period combined. Panel B excludes stocks with market capitalisation in the bottom cross-sectional tercile, to address concerns related to liquidity and implementability.

Panel A the cumulative return of B&B portfolios – mispricing portfolios										
			B&B - *ACCT*			B&B - *SYY*			B&B - *SYY*M	
		Overpriced	Underpriced	Long-Short	Overpriced	Underpriced	Long-Short	Overpriced	Underpriced	Long-Short
Pre-PEAD (-12, -6)	Median Split	-1.04%**	0.80%***	1.84%***	-0.45%	0.83%***	1.27%*	-0.45%	0.84%***	1.29%*
	Quintile	-1.52%**	1.73%***	3.25%***	-0.88%	1.54%**	2.42%**	0.07%	1.82%***	1.75%**
Pre-PEAD (-12, 0)	Median Split	-2.01%***	1.95%***	3.96%***	-0.99%	1.92%***	2.91%**	-1.07%	1.95%***	3.02%***
	Quintile	-1.51%**	2.33%**	3.83%***	-0.75%	2.51%***	3.27%**	0.61%	2.86%***	2.25%
Pre-PEAD & PEAD (-12, 12)	Median Split	-0.84%	2.38%**	3.23%	-0.26%	2.84%**	3.10%	0.30%	2.76%***	2.46%
	Quintile	-0.25%	0.29%	0.54%	2.70%	0.97%	-1.73%	3.53%	1.07%	-2.46%

Panel B the cumulative return of B&B portfolios – mispricing portfolios, bottom tercile MCAP removed										
			B&B - *ACCT*			B&B - *SYY*			B&B - *SYY*M	
		Overpriced	Underpriced	Long-Short	Overpriced	Underpriced	Long-Short	Overpriced	Underpriced	Long-Short
Pre-PEAD (-12, -6)	Median Split	-0.99%**	1.28%***	2.27%***	-0.40%	1.30%***	1.71%*	-0.40%	1.32%***	1.72%*
	Quintile	-1.48%**	1.69%***	3.17%***	-0.84%	1.51%**	2.35%**	0.12%	1.78%***	1.66%*
Pre-PEAD (-12, 0)	Median Split	-1.94%***	2.30%***	4.24%***	-0.92%	2.27%***	3.19%**	-0.99%	2.29%***	3.28%**
	Quintile	-1.43%*	2.17%**	3.61%**	-0.67%	2.36%**	3.03%**	0.68%	2.68%***	2.00%
Pre-PEAD & PEAD (-12, 12)	Median Split	-0.76%	2.95%*	3.70%	-0.18%	3.40%**	3.58%	0.40%	3.32%**	2.93%
	Quintile	-0.11%	0.13%	0.24%	2.85%*	0.81%	-2.03%	3.63%	0.91%	-2.72%

Table 10 Stock selection: summary of correct and wrong predictions on *B&B*'s

This table summarizes the predictive performance of stock selection within the top and bottom quintile portfolios based on historical accounting mispricing scores (*ACCT*). For each fiscal year, a random number of stocks are selected from each group 2,000 times. The analysis investigates whether the selection of stocks within underpriced (overpriced) portfolios is associated with higher (lower) foresight unexpected earnings, as well as how the number of selected stocks influences the relationship between accounting mispricing scores and future unexpected earnings. The first panel defines a correct prediction as the case where, for underpriced (overpriced) stocks, the number of stocks with favorable earnings outcomes exceeds (falls below) the number with unfavorable outcomes. The second panel considers a prediction accurate if the average foresight unexpected earnings of the portfolio are positive for underpriced stocks or negative for overpriced stocks, where the third panel identifies accurate if the portfolio median is confirmative. Chi-square test shows how significantly the probability is different from that of benchmark group which is (0, 10] or (10,50]. The probabilities of profitable predictions are further categorized as follows: Profitable for both long and short portfolios: the probability that both overpriced and underpriced portfolios correctly predict Ball and Brown's (1968) earnings surprise. Profitable for only long or short portfolio: the probability that only one portfolio (long or short) correctly predicts the earnings surprise. Unprofitable for both portfolios: the probability that neither portfolio correctly predicts the earnings surprise. *Significance at the 10% level; **significance at the 5% level; ***significance at the 1% level.

No. stocks in portfolio	*ACCT* mispricing *(t-1)*	No. obs	Prob. of correct prediction	Is the likelihood significantly different from that of (0, 10]? (Chi-square test)	Is the likelihood significantly different from that of (10, 50]? (Chi-square test)	Prob. of profitable, both long and short portfolios	Prob. of profitable, only long or short portfolio	Prob. of unprofitable, neither long nor short portfolio
	Panel A: If number of stocks with confirmative B&B EPS is greater than that of stocks with non-confirmative B&B EPS							
(0, 10]	Overpriced	4,511	72.64%		-***	43.32%	45.64%	11.04%
	Underpriced	4,608	59.64%		-***			
(10, 50]	Overpriced	18,547	79.54%	+***		65.93%	30.57%	3.50%
	Underpriced	18,575	82.89%	+***				
(50, 100]	Overpriced	23,102	86.77%	+***	+***	80.94%	18.18%	0.89%
	Underpriced	23,088	93.27%	+***	+***			
(100, more]	Overpriced	53,840	90.72%	+***	+***	88.73%	11.06%	0.20%
	Underpriced	53,729	97.81%	+***	+***			
	Panel B: If average B&B EPS is confirmative							
(0, 10]	Overpriced	4,511	62.96%		-***	44.17%	44.78%	11.05%
	Underpriced	4,608	70.16%		-***			
(10, 50]	Overpriced	18,547	73.97%	+***		60.76%	34.59%	4.65%
	Underpriced	18,575	82.13%	+***				
(50, 100]	Overpriced	23,102	75.55%	+***	+***	66.90%	30.30%	2.80%
	Underpriced	23,088	88.55%	+***	+***			
(100, more]	Overpriced	53,840	71.26%	+***	-***	62.94%	33.70%	3.35%
	Underpriced	53,729	88.32%	+***	+***			
	Panel C: If median B&B EPS is confirmative							
(0, 10]	Overpriced	4,511	55.95%		-***	39.84%	47.48%	12.68%
	Underpriced	4,608	71.20%		-***			
(10, 50]	Overpriced	18,547	60.45%	+***		51.61%	42.61%	5.78%
	Underpriced	18,575	85.38%	+***				
(50, 100]	Overpriced	23,102	62.61%	+***	+***	57.22%	39.56%	3.22%
	Underpriced	23,088	91.39%	+***	+***			
(100, more]	Overpriced	53,840	61.17%	+***	+*	56.70%	40.46%	2.84%
	Underpriced	53,729	92.68%	+***	+***			

Table 11 Return outcomes of simulated stock selections

This table presents the monthly average, risk-adjusted returns achieved through stock selection within the top and bottom quintile portfolios based on historical accounting mispricing scores (*ACCT*). For each fiscal year, a random number of stocks is selected from each quintile 2,000 times. This experimental design assesses whether the selection of stocks from underpriced portfolios is associated with higher foresight unexpected earnings, and conversely, whether overpriced portfolios are associated with lower foresight unexpected earnings, as well as how varying the number of stocks selected impacts the relationship between mispricing scores and future earnings shocks. A prediction is deemed accurate if the average foresight unexpected earnings for the portfolio are positive for underpriced stocks and negative for overpriced stocks. Each portfolio is held from the month immediately following the reporting date of fiscal year *t-1* until the subsequent reporting month. The table reports three key metrics for portfolios with different numbers of stocks: (i) the time series average returns in percentage adjusted by Fama-French-Carhart four-factor model, (ii) the time series average standard deviation of adjusted returns, and (iii) the sign adjusted value of the return-to-standard-deviation ratio in percentage, which serves as a proxy for investor preference. The sign adjusted ratio is positive when the direction of realized returns is consistent with initial mispricing signal (positive returns for underpriced group and negative returns for overpriced group), negative otherwise. Conditional on the returns exhibiting the correct sign according to mispricing signals, a higher return-risk ratio implies a stronger preference for that portfolio and a lower or negative ratio suggests investors hesitate to take positions based on the initial mispricing signal. Returns are reported in percentage. *Significance at the 10% level; **significance at the 5% level; ***significance at the 1% level.

No. stocks in portfolio	*ACCT* mispricing *(t-1)*	Risk-adjusted returns by FF3&UMD (%)	Standard deviation of risk-adjusted returns	Sign adjusted ratio of risk adjusted return / standard deviation %	Are averaged unexpected earnings at *t* confirming the mispricing signal at *t-1*?	Risk-adjusted returns by FF3&UMD (%)	Standard deviation of risk-adjusted returns	Sign adjusted ratio of risk adjusted return / standard deviation %
(0, 10]	Overpriced	-0.070	0.016	4.268	0	0.173**	0.017	-10.485
					1	-0.235***	0.016	14.609
	Underpriced	0.408***	0.028	14.833	0	-0.120	0.028	-4.255
					1	0.622***	0.027	23.468
(10, 50]	Overpriced	-0.106	0.006	17.650	0	0.002	0.006	-0.254
					1	-0.141**	0.006	23.500
	Underpriced	0.439***	0.009	46.734	0	0.254*	0.010	25.400
					1	0.457***	0.009	49.097
(50, 100]	Overpriced	-0.116*	0.003	37.484	0	-0.050	0.003	14.829
					1	-0.129**	0.003	41.710
	Underpriced	0.434***	0.005	88.551	0	0.271*	0.005	55.306
					1	0.441***	0.005	89.898
(100, more]	Overpriced	-0.111*	0.001	79.286	0	-0.096	0.002	60.250
					1	-0.112*	0.002	74.533
	Underpriced	0.433***	0.002	188.435	0	0.217	0.003	84.077
					1	0.435***	0.002	197.957

Table 12 Panel regression: the impact of portfolio size on predictive performance of random stock selection
This table examines the predictive performance of stock selection within the top and bottom quintile portfolios based on historical accounting mispricing scores (*ACCT*) using panel regressions. For each fiscal year, a random number of stocks is selected from each group 2,000 times. For each selection, the average values of various stock characteristics are calculated, including accounting mispricing scores, log market capitalization, book-to-market ratio (BM), investment, profitability, medium-term past performance, and long-term past performance. Additionally, the correct prediction rate is computed as the proportion of stocks with favorable foresight Ball and Brown (*B&B*) earnings outcomes for next fiscal year relative to the total number of stocks in the selection. The table reports the coefficients from a panel regression of the correct prediction rate on independent variables, including the log of the number of stocks selected, the average *ACCT* scores, the interaction term between the log number of stocks selected and the average *ACCT* scores, and control variables. Fixed effects are employed cross sectionally at the level of group of stock number of each selection, and time serially at fiscal year level. Standard errors are clustered at the level of stock numbers of each selection. *Significance at the 10% level; **significance at the 5% level; ***significance at the 1% level.

	ACCT overpriced	*ACCT* underpriced
ACCT score * Log (No. stocks)	0.024***	0.012***
	(3.026)	(3.113)
ACCT score	-0.343***	0.400***
	(-12.672)	(16.852)
Log (No. stocks)	0.010***	-0.014**
	(4.570)	(-1.977)
Control variables		
Log (MCAP)	0.004***	0.016***
	(4.848)	(3.922)
BM	0.055***	-0.001
	(9.301)	(-0.725)
Investment	-0.002	-0.031***
	(-0.105)	(-13.137)
Profitability	0.077***	-0.092***
	(6.439)	(-27.495)
RET[-12, -1]	-0.012***	-0.010***
	(-11.285)	(-6.380)
RET[-60, -13]	0.006**	-0.019***
	(2.566)	(-4.705)
Constant	0.113***	0.110***
	(4.011)	(2.608)
No. Observations	100,000	100,000
Adj. R-square	0.009	0.030

Table 13 The impact of institutional and analyst coverage on the earnings prediction

This table presents the results of OLS panel regressions examining the impact of institutional and analyst coverage on the predictive relationship between foresight unexpected earnings, as defined by Ball and Brown (1968), and mispricing scores. We identify high institutional (analyst) coverage if a firm's institutional holding percentage (the number of recommendations) is higher than the cross-sectional median. For the number of recommendations, if IBES returns a missing value, we replace it with zero. The dependent variable, unexpected earnings for year *t*, is regressed on two mispricing scores *SYY* and *ACCT,* where *SYY* is the aggregated cross-sectional rank of 11 anomalies (Stambaugh, Yu and Yuan, 2012, 2015) and *ACCT* is the aggregated cross-sectional rank of six annual accounting-based anomalies out of *SYY's* 11 anomalies. All information used to construct mispricing scores is publicly available information at the end of announcement month of the previous fiscal. We investigate the impact of institutional (analyst) coverage on the earnings prediction by adding an interaction between mispricing scores and the dummy variable of high coverage. Control variables include the earnings per share for the previous fiscal year, natural logarithm of market capitalization, past stock returns over short-, medium-, and long-term horizons, and accounting controls, including book-to-market, investment-to-assets, and gross profitability. The sample period for institutional holding (analyst) data covers 1980 (1994) to 2024. The table reports estimated coefficients along with clustered t-statistics on fiscal year and firm. *Significance at the 10% level; **significance at the 5% level; ***significance at the 1% level.

	SYY mispricing score		*ACCT* mispricing score	
	(1)	(2)	(3)	(4)
	Institutional coverage	Analyst coverage	Institutional coverage	Analyst coverage
Mispricing	0.113***	0.108***	0.026***	0.024**
	(8.438)	(6.541)	(2.726)	(2.113)
High coverage -- institution	-0.026		-0.027	
	(-1.239)		(-1.249)	
Mispricing * High coverage -- institution	0.056***		-0.018	
	(3.494)		(-1.335)	
High coverage -- analyst		-0.052***		-0.067***
		(-2.667)		(-3.493)
Mispricing * High coverage -- analyst		0.094***		0.001
		(4.590)		(0.081)
Previous EPS	-0.623***	-0.628***	-0.610***	-0.612***
	(-24.021)	(-17.355)	(-23.772)	(-17.199)
MCAP	0.189***	0.241***	0.200***	0.258***
	(5.364)	(6.208)	(5.614)	(6.587)
Ret [-1,0]	0.043***	0.032***	0.046***	0.036***
	(5.624)	(3.570)	(6.012)	(3.802)
Ret [-12,-1]	0.053***	0.046***	0.088***	0.081***
	(5.936)	(3.894)	(10.158)	(7.131)
Ret [-60,-13]	0.016**	0.008	0.020**	0.010
	(2.024)	(1.128)	(2.562)	(1.312)
BM	-0.144***	-0.146***	-0.149***	-0.155***
	(-8.590)	(-7.793)	(-8.808)	(-7.948)
Investment	0.032***	0.050***	-0.018*	0.002
	(3.260)	(4.884)	(-1.753)	(0.228)
Profitability	0.019	0.022	0.068***	0.071***
	(1.293)	(1.456)	(4.892)	(4.803)
Adj. Squared R	0.257	0.247	0.250	0.238
No. Fiscal Years	45	31	45	31
No. Observations	46,738	33,810	46,738	33,810
Cross-Sectional Average No. Observations	1,039	1,090	1,039	1,090
FYEAR Fixed	Yes	Yes	Yes	Yes
Firm Fixed	Yes	Yes	Yes	Yes

Table 14 The impact of investor sentiment on the earnings prediction using mispricing scores

This table presents the results of OLS panel regressions examining the impact of investor sentiment on the predictive relationship between foresight unexpected earnings, as defined by Ball and Brown (1968), and mispricing scores. We calculate the average Baker and Wurgler (2006) investor sentiment (*BW*) over the past *j*-month period prior to and including the fourth quarter announcement month of the previous fiscal year. The dependent variable, unexpected earnings for year *t*, is regressed on two mispricing scores *SYY* and *ACCT,* where *SYY* is the aggregated cross-sectional rank of 11 anomalies (Stambaugh, Yu and Yuan, 2012, 2015) and *ACCT* is the aggregated cross-sectional rank of six annual accounting-based anomalies out of *SYY's* 11 anomalies. All information used to construct mispricing scores is publicly available information at the end of announcement month of the previous fiscal. We investigate the impact of investor sentiment on the earnings prediction by adding an interaction between mispricing scores and the average investor sentiment. Control variables include the earnings per share for the previous fiscal year, natural logarithm of market capitalization, past stock returns over short-, medium-, and long-term horizons, and accounting controls, including book-to-market, investment-to-assets, and gross profitability. The sample period covers 1974 to 2023. The table reports estimated coefficients along with clustered t-statistics on fiscal year and firm. *Significance at the 10% level; **significance at the 5% level; ***significance at the 1% level.

	j=0		*j=2*		*j=5*		*j=11*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	SYY	*ACCT*	*SYY*	*ACCT*	*SYY*	*ACCT*	*SYY*	*ACCT*
Mispricing	0.081***	0.026***	0.081***	0.026***	0.080***	0.026***	0.080***	0.025**
	(6.788)	(2.655)	(6.751)	(2.666)	(6.773)	(2.628)	(6.713)	(2.553)
$BW_{(rdq-j,\ rdq)}$	0.012	0.018	-0.016	-0.013	0.023	0.029	0.001	-0.007
	(0.194)	(0.308)	(-0.125)	(-0.107)	(0.175)	(0.228)	(0.003)	(-0.040)
Mispricing * $BW_{(rdq-j,\ rdq)}$	-0.017	-0.019*	-0.015	-0.018*	-0.016	-0.018*	-0.023**	-0.022**
	(-1.635)	(-1.934)	(-1.429)	(-1.786)	(-1.591)	(-1.854)	(-2.439)	(-2.391)
Previous EPS	-0.601***	-0.590***	-0.601***	-0.590***	-0.601***	-0.590***	-0.601***	-0.591***
	(-24.240)	(-24.074)	(-24.244)	(-24.082)	(-24.222)	(-24.067)	(-24.295)	(-24.145)
MCAP	0.098***	0.102***	0.097***	0.101***	0.098***	0.102***	0.097***	0.101***
	(2.721)	(2.820)	(2.712)	(2.815)	(2.721)	(2.829)	(2.705)	(2.819)
Ret [-1,0]	0.054***	0.056***	0.055***	0.056***	0.054***	0.056***	0.054***	0.056***
	(9.025)	(9.214)	(9.055)	(9.192)	(9.025)	(9.166)	(9.012)	(9.202)
Ret [-12,-1]	0.021***	0.039***	0.021***	0.040***	0.021***	0.040***	0.022***	0.040***
	(2.626)	(4.476)	(2.683)	(4.560)	(2.700)	(4.560)	(2.750)	(4.581)
Ret [-60,-13]	0.008	0.011	0.008	0.011	0.008	0.011	0.009	0.012
	(0.842)	(1.139)	(0.827)	(1.127)	(0.834)	(1.138)	(0.874)	(1.185)
BM	-0.152***	-0.155***	-0.152***	-0.155***	-0.152***	-0.155***	-0.153***	-0.155***
	(-9.475)	(-9.443)	(-9.480)	(-9.460)	(-9.485)	(-9.469)	(-9.554)	(-9.554)
Investment	-0.006	-0.025**	-0.006	-0.024**	-0.006	-0.025**	-0.006	-0.025**
	(-0.566)	(-2.436)	(-0.563)	(-2.427)	(-0.575)	(-2.440)	(-0.581)	(-2.423)
Profitability	0.041***	0.062***	0.041***	0.062***	0.041***	0.061***	0.041***	0.061***
	(2.930)	(4.466)	(2.937)	(4.463)	(2.944)	(4.453)	(2.990)	(4.459)
Adj. Squared R	0.231	0.229	0.231	0.229	0.231	0.229	0.231	0.229
No. Fiscal Years	50	50	50	50	50	50	50	50
No. Observations	52,609	52,609	52,609	52,609	52,609	52,609	52,609	52,609
Cross-Sectional Average No. Observations	1,052	1,052	1,052	1,052	1,052	1,052	1,052	1,052
FYEAR Fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm Fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Figure 1 The perfect foresight vs. mispricing drift

This figure reported value-weighted, four-factor adjusted cumulative returns over the pre-PEAD *(-12, 0)* and PEAD *(1, 12)* periods. The perfect foresight information refers to the unexpected earnings (Ball and Brown, 1968) which is measured by the difference between actual earnings reported at *time 0* and expected earnings estimated by market earnings model. The foresight information is not available over the pre-PEAD period, is released at *time 0*, and publicly available over the PEAD period. The mispricing information refers to three mispricing scores (1) accounting mispricing score *(ACCT),* (2) mispricing score of *SYY* which we only obtain information at the beginning of the pre-PEAD period*,* and (3) a version of the *SYY* mispricing score that is updated monthly throughout the full period *(SYYM).* All the information required by the three mispricing scores is available over the entire period *(-12, 12).*The monthly cumulative adjusted return on the portfolio of firms in the top or bottom group under cross-sectional median split (Panel a) and quintile sorts (Panel b) are plotted as the return that investors can achieve if they have "god-like" power (perfect foresight on earnings), which is compared to the drifts of the most underpriced or overpriced portfolios under cross-sectional median split (Panel a) and quintile sorts (Panel b) which is human power. In each panel, the drift comparison between short positions of "god-like" power ad human power is shown in the first figure, long positions second, and long-short difference is in the third figure.

Panel A: Median-split portfolios

Figure 1 B&B Losers v.s. Overpriced Stocks (Median)

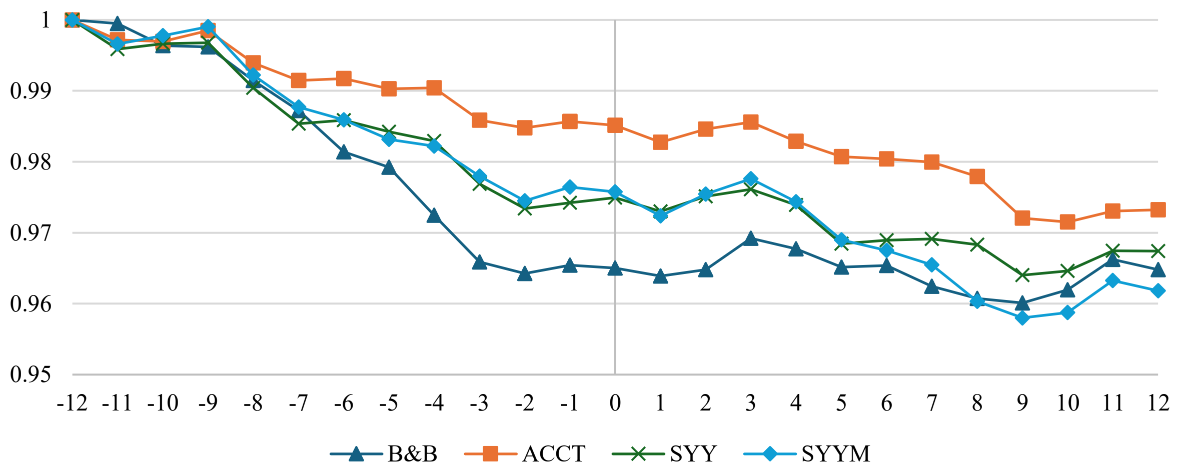


Figure 2 B&B Winners v.s. Underpriced Stocks (Median)

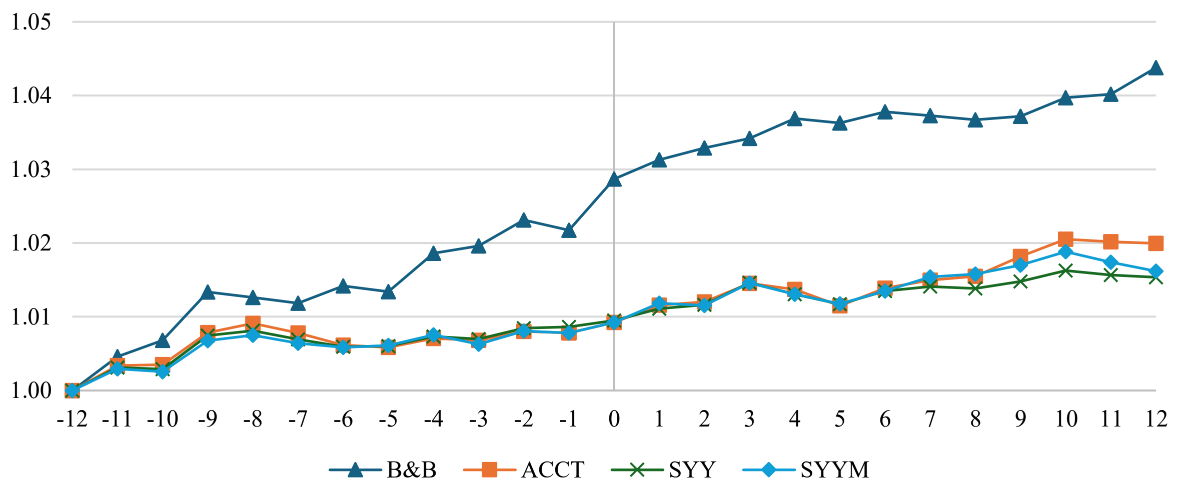
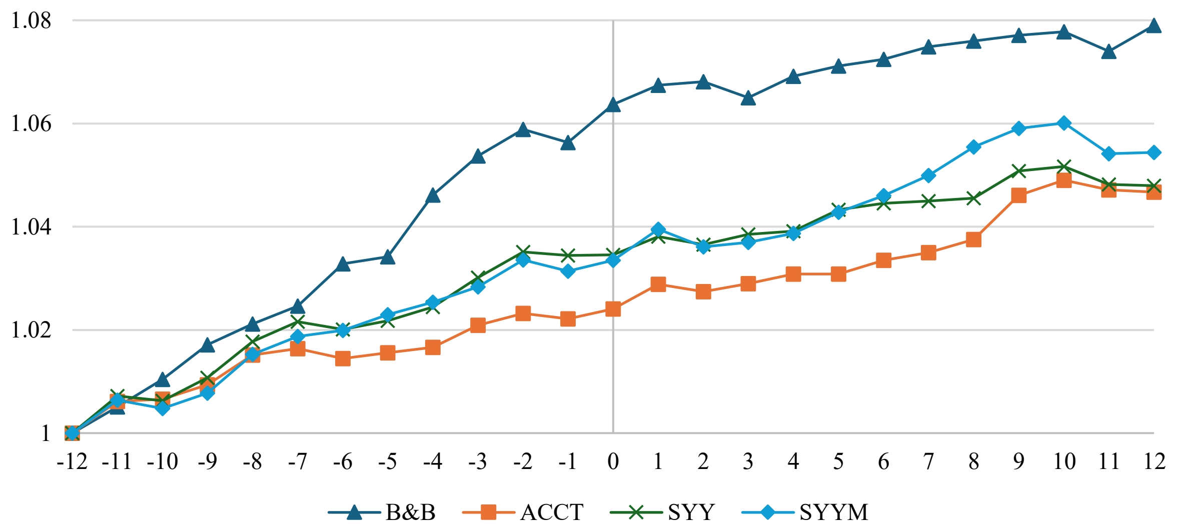
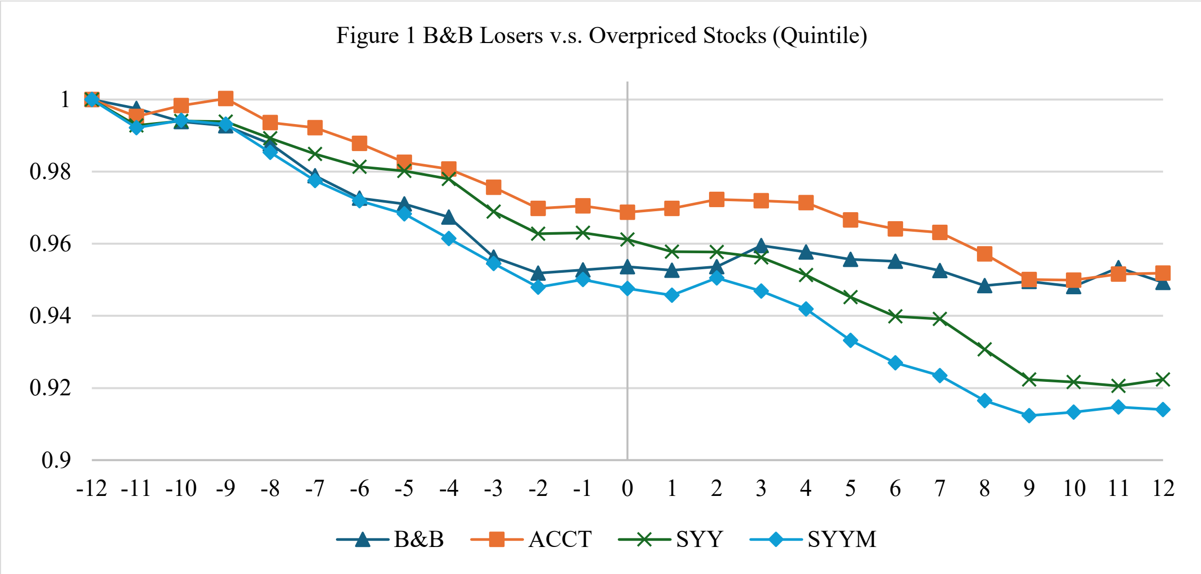
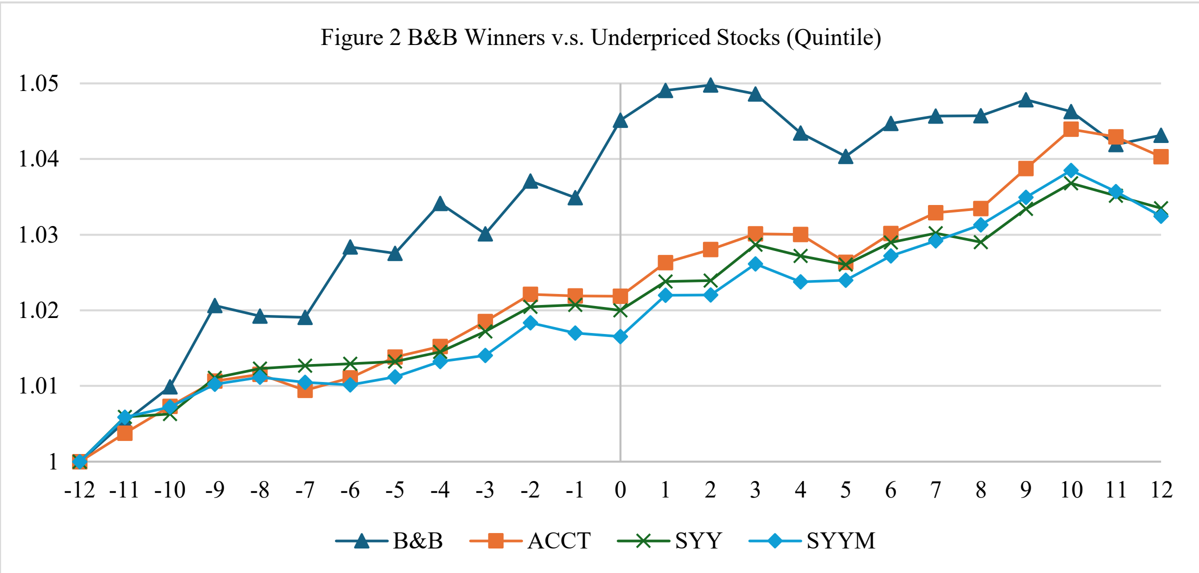


Figure 3 B&B WML Portfolio v.s. Mispricing Long-Short Portfolio (Median)



Panel B Quintile portfolios





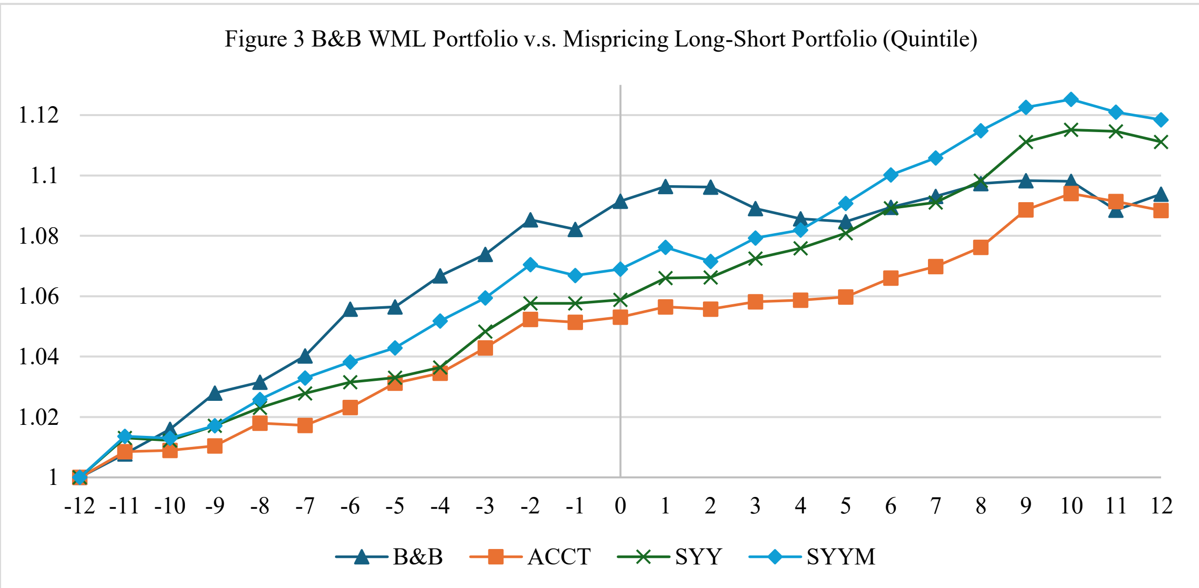


Figure 2 The perfect foresight vs. mispricing drift, size controlled

This figure reported value-weighted, four-factor adjusted cumulative returns over the pre-PEAD *(-12, 0)* and PEAD *(1, 12)* periods. The perfect foresight information refers to the unexpected earnings (Ball and Brown, 1968) which is measured by the difference between actual earnings reported at *time 0* and expected earnings estimated by market earnings model. The foresight information is not available over the pre-PEAD period, is released at *time 0*, and publicly available over the PEAD period. The mispricing information refers to three mispricing scores (1) accounting mispricing score *(ACCT),* (2) mispricing score of *SYY* which we only obtain information at the beginning of the pre-PEAD period*,* and (3) a version of the *SYY* mispricing score that is updated monthly throughout the full period *(SYYM).* All the information required by the three mispricing scores is available over the entire period *(-12, 12).*The monthly cumulative adjusted return on the portfolio of firms in the top or bottom group under cross-sectional median split (Panel a) and quintile sorts (Panel b) are plotted as the return that investors can achieve if they have "god-like" power (perfect foresight on earnings), which is compared to the drifts of the most underpriced or overpriced portfolios under cross-sectional median split (Panel a) and quintile sorts (Panel b) which is human power. In each panel, the drift comparison between short positions of "god-like" power ad human power is shown in the first figure, long positions second, and long-short difference is in the third figure. At each month, we remove stocks with the market capitalisation lower than the 33% cross-sectional threshold to avoid liquidity and transaction cost issues and improve the feasibility of human power returns.

Panel A: Median-split portfolios

Figure 1 B&B Losers v.s. Overpriced Stocks (Median)

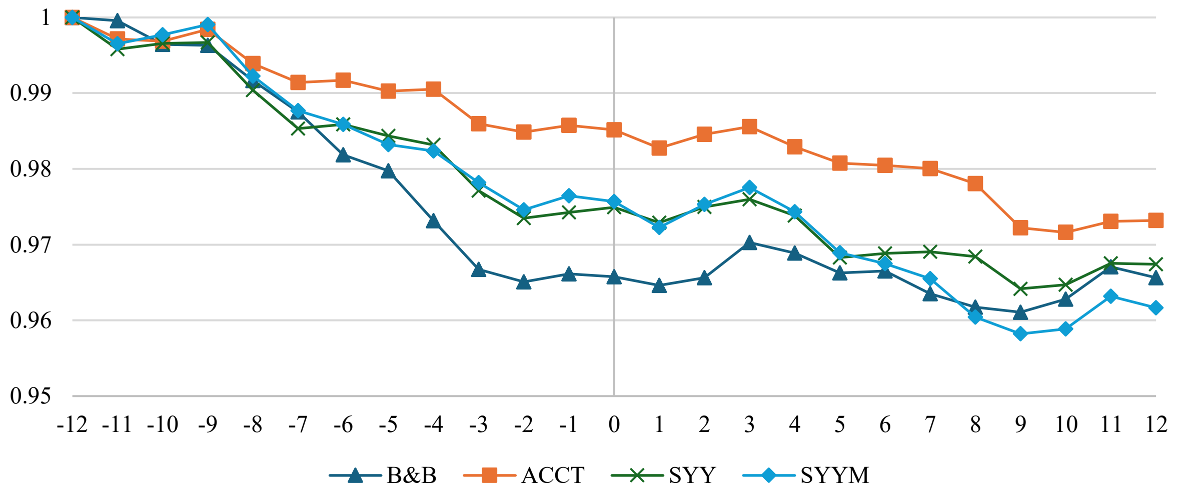


Figure 2 B&B Winners v.s. Underpriced Stocks (Median)

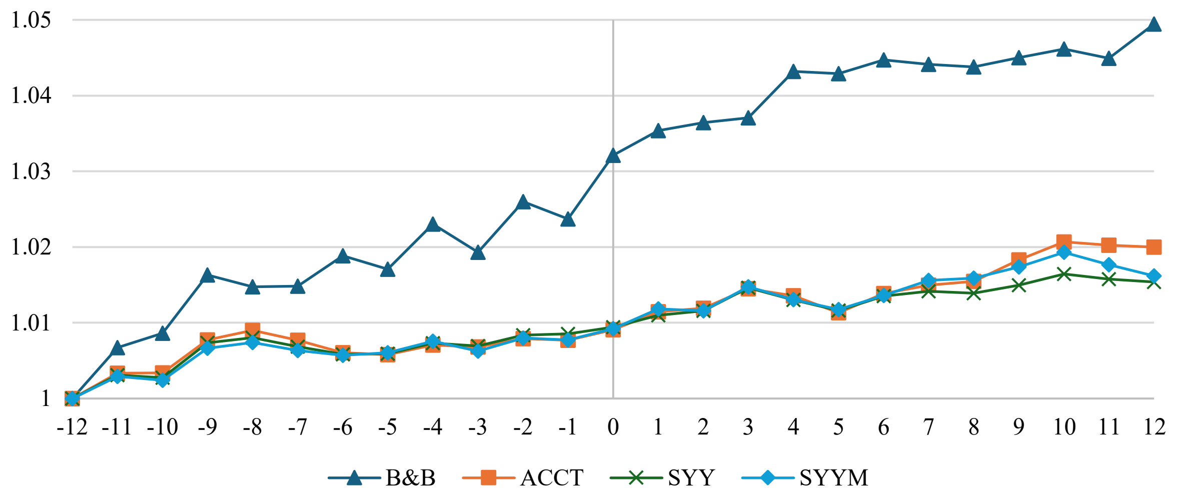
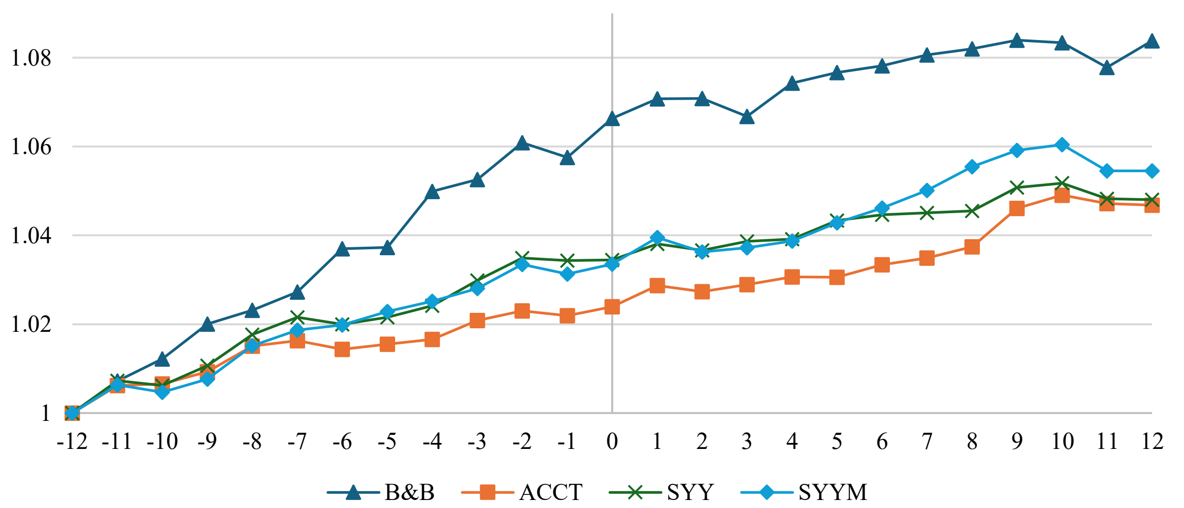


Figure 3 B&B WML Portfolio v.s. Mispricing Long-Short Portfolio (Median)



Panel B Quintile portfolios

Figure 1 B&B Losers v.s. Overpriced Stocks (Quintile)

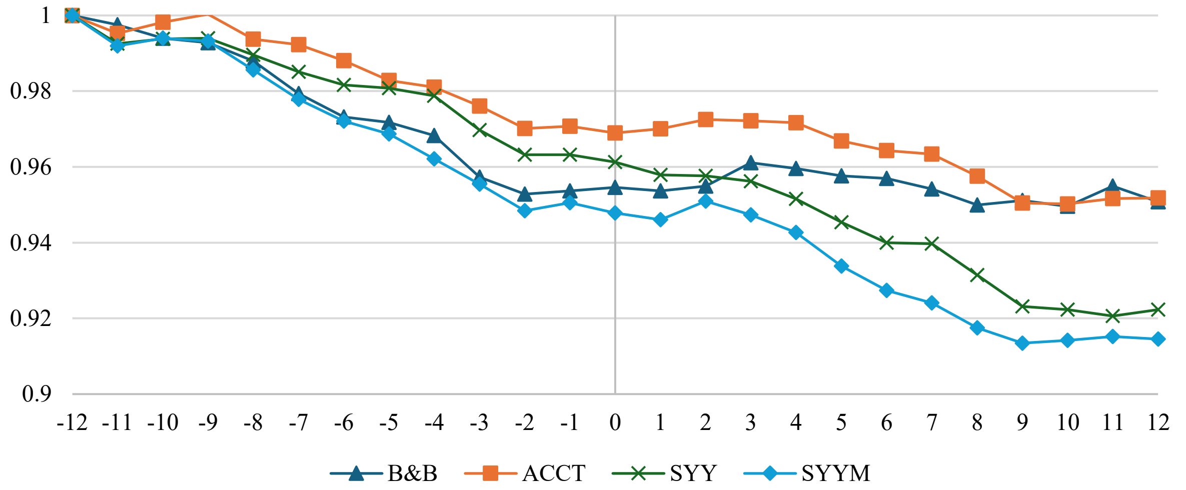


Figure 2 B&B Winners v.s. Underpriced Stocks (Quintile)

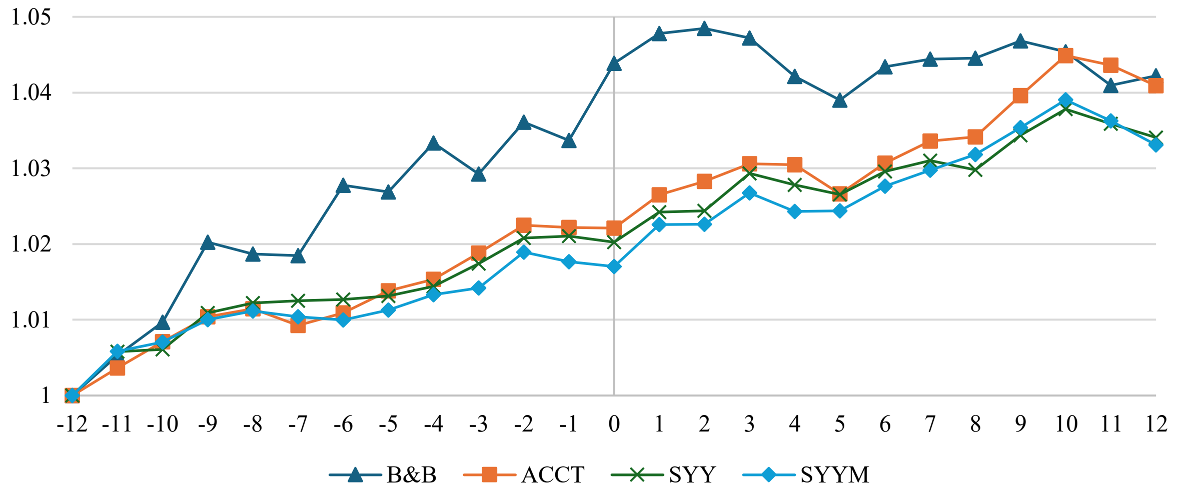


Figure 3 B&B WML Portfolio v.s. Mispricing Long-Short Portfolio (Quintile)

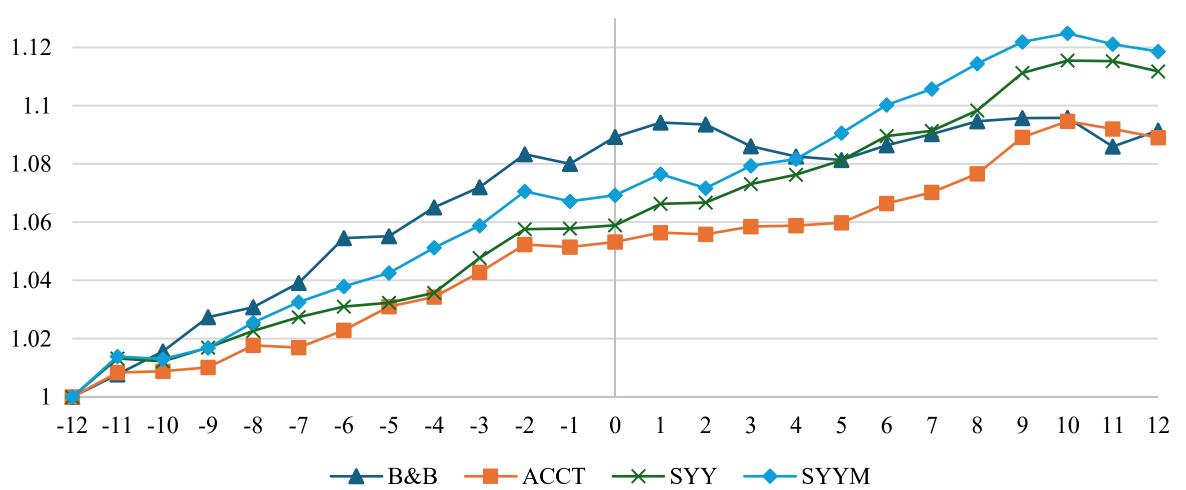
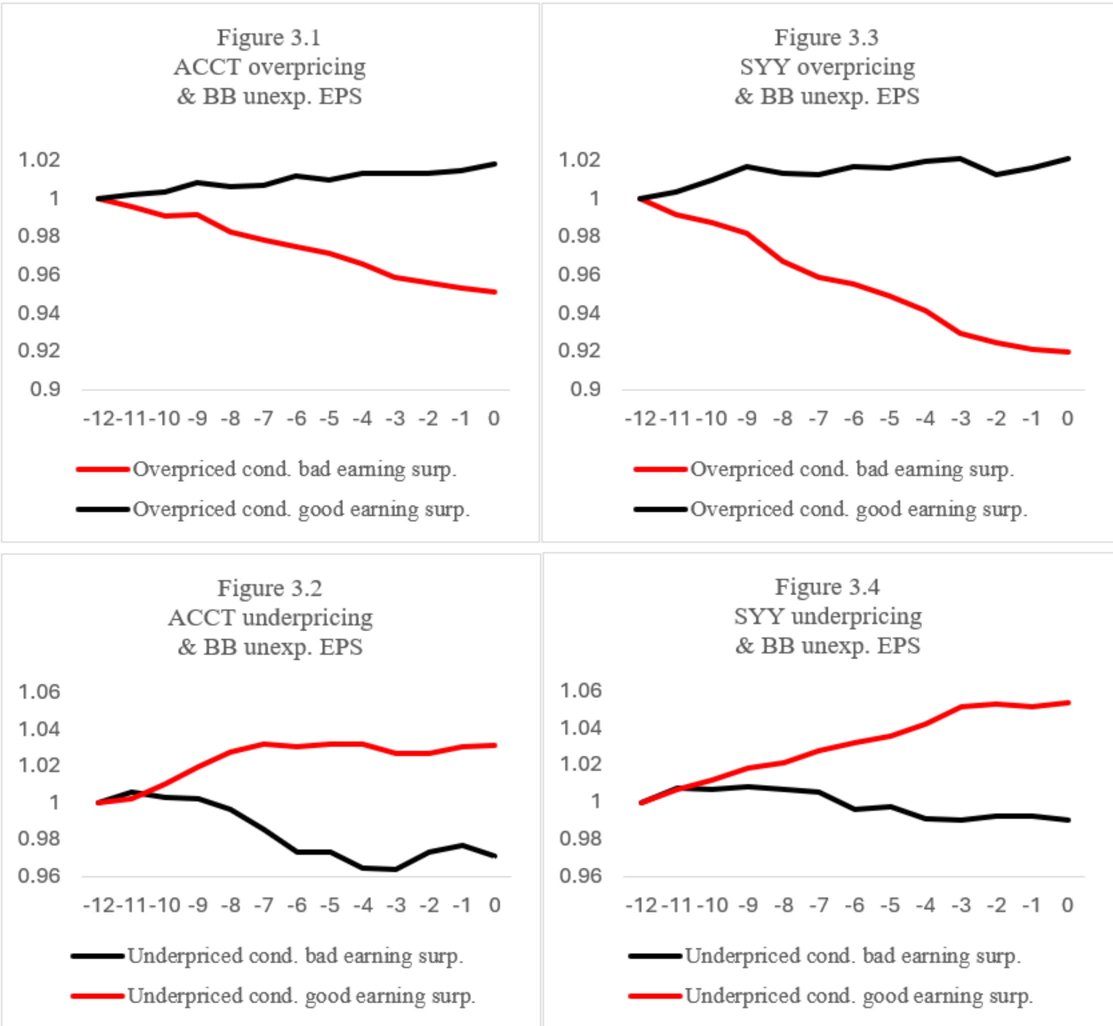


Figure 3 Mispricing groups conditional on future earnings shock

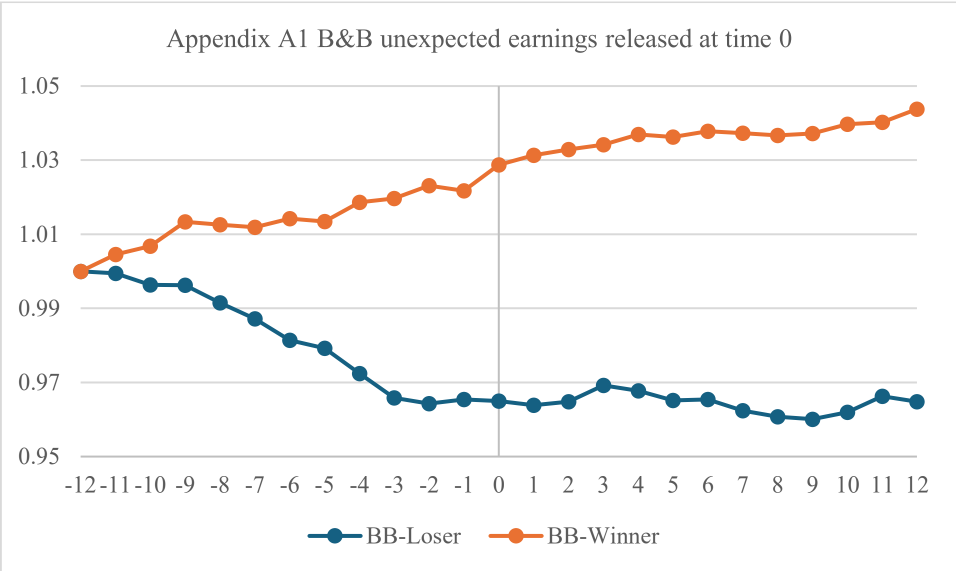
This figure reported value-weighted, four-factor adjusted cumulative returns prior to the earnings shocks. Earnings shocks are proxied by unexpected earnings (Ball and Brown, 1968) which is measured by the difference between actual earnings reported and expected earnings estimated by market earnings model. The cumulative adjusted return on the portfolio of firms those are overpriced and underpriced based on mispricing scores intersected with good or bad earnings shocks are plotted monthly from -12 months before the earnings reporting month to the earnings reporting month for each mispricing group. Red (Black) line represents the cumulative returns of mispricing portfolio conditional on confirmative (non-confirmative) earnings shocks, where confirmative earnings shocks refer to good unexpected earnings for underpriced stocks and bad unexpected earnings for overpriced stocks. Figure 3.1 and 3.2 reports the returns of accounting mispricing portfolios constructed by six annual accounting indicators, conditional on confirmative or non-confirmative earnings shocks, where Figure 3.3 and 3.4 reports the returns of mispricing portfolios constructed by eleven anomalies (*SYY*), conditional on confirmative or non-confirmative earnings shocks



Appendix

Appendix A1 Pre-PEAD and PEAD of *B&B*'s median split portfolio

This figure reported value-weighted, four-factor adjusted cumulative returns prior to and following favorable and unfavorable earnings shocks. Earnings shocks are proxied by unexpected earnings (Ball and Brown, 1968) which is measured by the difference between actual earnings reported and expected earnings estimated by market earnings model. The cumulative adjusted return on the portfolio of firms with higher or lower than median earnings shocks are plotted monthly from 12 months before the earnings reporting month to 12 months after.



Appendix A2 Average monthly returns over the holding period of portfolios based on two mispricing signals and Ball and Brown's unexpected earnings

This table reports the equal-weighted and value-weighted average monthly returns of bottom, top and top-minus-bottom quintile portfolios sorted based on two mispricing signals (Stambaugh, Yu and Yuan's mispricing score (*SYY*) and an aggregated accounting mispricing score (*ACCT*) based on six annual accounting ratios used in *SYY*) and the foresight unexpected earnings developed by Ball and Brown (1968). Mispricing signal refers to information that is available at the time the portfolio is constructed, where the B&B's unexpected earnings-related information is not available but released at the end of the holding period. At each fourth quarter reporting date of fiscal year *t-1*, stocks are independently sorted based on each power. The information on mispricing-based portfolios includes accounting and financial market information, which are publicly available to sort. The information on Ball and Brown's unexpected earnings is for firm's next fiscal year *t,* which will be available at the end of the holding period that is the reporting date of the next final quarter for the firm. The monthly returns of B&B portfolios are not realizable in practice due to the unavailability of earnings information at the time. However, these returns represent the information of perfect foresight on future unexpected earnings. The sample includes firms with a December fiscal year-end and sufficient accounting data to calculate Ball and Brown's earnings surprise, either *ACCT* or *SYY* mispricing measures over the period from 1974 to 2024. All T-statistics are adjusted employing Newey-West t-statistics. Returns are reported in percentage. *Significance at the 10% level; **significance at the 5% level; ***significance at the 1% level.

	Mispricing #1: *ACCT* mispricing *(t-1)*			Mispricing #2: *SYY* mispricing *(t-1)*			Perfect foresight: *B&B*'s unexpected EPS *(t)*		
	Overpriced	Underpriced	Underpriced-Overpriced	Overpriced	Underpriced	Underpriced-Overpriced	Bad earnings	Good earnings	Good-Bad
				Panel A Equal Weighted					
Excess return	0.541**	1.163***	0.622***	0.374	1.159***	0.785***	-0.111	1.752***	1.863***
	(2.445)	(4.553)	(5.431)	(1.488)	(5.676)	(7.067)	(-0.445)	(7.387)	(20.812)
FF3 adjusted	-0.276***	0.253***	0.529***	-0.539***	0.389***	0.928***	-1.055***	0.840***	1.895***
	(-3.962)	(2.978)	(5.464)	(-5.733)	(6.322)	(8.759)	(-13.337)	(11.260)	(21.941)
FF3 UMD adjusted	-0.177***	0.351***	0.528***	-0.319***	0.391***	0.710***	-0.870***	0.920***	1.790***
	(-2.611)	(4.086)	(5.314)	(-3.750)	(6.183)	(7.187)	(-11.971)	(10.814)	(18.333)
FF5 adjusted	-0.238***	0.303***	0.541***	-0.456***	0.225***	0.682***	-1.066***	0.727***	1.792***
	(-3.550)	(3.692)	(5.802)	(-4.546)	(3.954)	(6.568)	(-12.123)	(9.448)	(20.682)
				Panel B Value Weighted					
Excess return	0.428**	0.931***	0.503***	0.330	0.814***	0.484***	0.161	1.107***	0.946***
	(2.227)	(4.588)	(3.947)	(1.566)	(4.584)	(3.563)	(0.818)	(5.741)	(8.989)
FF3 adjusted	-0.194***	0.232**	0.426***	-0.399***	0.270***	0.670***	-0.492***	0.476***	0.968***
	(-2.812)	(2.255)	(3.275)	(-4.163)	(3.576)	(4.997)	(-6.151)	(6.419)	(9.054)
FF3 UMD adjusted	-0.188***	0.269**	0.457***	-0.270***	0.190***	0.461***	-0.454***	0.453***	0.907***
	(-2.756)	(2.531)	(3.502)	(-2.808)	(2.578)	(3.524)	(-5.629)	(6.284)	(8.603)
FF5 adjusted	-0.190***	0.078*	0.268**	-0.337***	0.026*	0.363***	-0.543***	0.332***	0.876***
	(-2.736)	(1.684)	(2.123)	(-3.257)	(1.682)	(2.681)	(-6.625)	(4.504)	(8.267)

Appendix A3 Calculations for the probability of correct selections based on the no-replacement game.

If we assume that at a specific cross section the future earnings shock and mispricing predictions follow the average level shown in the first panel of Table 8, there are 520 stocks with high predictions where 317 will be confirmed by good earnings shock.

We assume that investors hold *n* stocks, calculation of the probability to have *k* or more good earnings shock stocks is as follow:

$$p(X \geq k) = 1 - \sum_{i=0}^{k-1} \binom{n}{i} p^i \, (1-p)^{n-i}$$

where p is the empirical chance of a correct prediction of 60.95%. *n* is the number stocks investors hold in portfolio. *k* is the number of stocks confirmed by future earnings in the selected portfolio.

For the probability of the scenario (1) select *n*=9 stocks and get *k*=5 confirmed by future earnings:

$$p(X \geq 5) = 1 - \sum_{i=0}^{5-1} \binom{9}{i} 60.95\%^i \, (1 - 60.95\%)^{9-i} = 75.32\%$$

For the probability of the scenario (2) select *n*=90 stocks and get *k*=50 confirmed by future earnings:

$$P(X \geq 50) = 1 - \sum_{i=0}^{50-1} \binom{90}{i} 60.95\%^i \, (1 - 60.95\%)^{90-i} = 87.64\%$$

Using similar calculation process, for the probability of the scenario (3) select *n*=900 stocks and get *k*=500 confirmed by future earnings:

$$P(X \geq 500) = 1 - \sum_{i=0}^{500-1} \binom{900}{i} 60.95\%^i \, (1 - 60.95\%)^{900-i} = 99.96\%$$