VILNIUS TECH
Vilnius Gediminas Technical University

Journal of Business Economics and Management
ISSN 1611-1699 / eISSN 2029-4433
2023 Volume 24 Issue 1: 37–53
https://doi.org/10.3846/jbem.2023.17832

ANALYSTS' DISAGREEMENT, SELF-SELECTION, AND STOCK RETURNS

Liang WU[1], Yunshen LONG[1*], Wenyue LI[2], Bingyan WU[1]

[1]*School of Economics, Sichuan University, Chengdu, China*
[2]*School of Finance, Southwestern University of Finance and Economics, Chengdu, China*

Received 25 November 2021; accepted 01 August 2022

Abstract. Two ex-ante variables are introduced to characterize the analysts' biased behavior, namely the analysts' disagreement and self-selection in analysts' earnings forecasts. The study investigates the impact of the analysts' disagreement and self-selection on the stock returns. A theoretical analysis derives how the stock returns are correlated with the two variables. There are two channels through which the stocks are priced according to the analysts' disagreement. The first one is the risk channel as the analysts' disagreement is associated with earnings uncertainty. The stock price will be discounted before the actual earnings announcement. The second one is the optimistic bias channel. The optimistic bias channel means that the stock is overpriced if the investors do not correct the analysts' bias. The self-selection is negatively correlated with the stock return through the optimistic bias channel as more self-selection means more optimistic bias as low forecasting values are not revealed. The empirical analysis using data from the Chinese stock market supports the theoretical conclusion.

Keywords: analysts' disagreement, self-selection, optimistic bias, stock returns, earning forecast, uncertainty.

JEL Classification: G12, G14.

Introduction

Analysts are important participants in capital markets. By collecting, interpreting and processing publicly disclosed information, analysts transmit information to other market participants (Gilson et al., 2001; Irvine et al., 2004), thereby reducing information asymmetry between investors and listed companies (Leuz, 2003). However, analysts are not always independent and objective, they tend to add bias to their true beliefs (Mola & Guidlin, 2009; Hu & Xia, 2017). Besides, analysts also have the so called "self-selection bias" in the sense that they may choose not to publish the earnings forecasts when their

*Corresponding author. E-mail: *longyunshen@scu.edu.cnt*

estimates are sufficiently low (McNichols & O'Brien, 1997). The existing literatures have a clear description of the motivation of the analysts' over-optimism. Jackson (2005), Hu and Xia (2017) believe that analysts can benefit from the commissions of recommended stock trading if they issue optimistic reports. Lin and McNichols (1998), Wang and Xu (2017) find that analysts can help brokerage firms to obtain refinancing underwriting business by reporting optimistic forecasts. Lim (2001), Zhao et al. (2013) think analysts' optimistic reports are catering to the management preferences, thus facilitating their access to private information of listed companies. Chu et al. (2019) find that asymmetric trading under short-selling and margin-trading program has prompted analysts to release optimistic forecasts based on their own salary and career development concerns.

The question this paper wants to study is whether investors can obtain excess returns by understanding the biased behavior of the analysts? To investigate the above question, this paper first checks whether the market has discovered the analysts' optimistic bias and corrected it when pricing the stock. If the market has not corrected the bias, the inefficiency would exist in the market in the sense that the stock price is overvalued following the analysts' overoptimistic forecasts and eventually would decline after the actual earnings are announced. Therefore, some pricing factors constructed according to the analysts' biased reports would have predicting power. In order to find out if this is the case, an empirical test is conducted by examining the relationship between the analysts' forecast bias and the stock's excess return before and after the announcement of the actual earnings. The existing literature that focuses on whether the market can correct the analysts' bias is quite limited. Zheng (2019) finds that managers can reduce analyst optimism by improving the quality of information disclosure and reducing earnings volatility. More literature examines whether new information is released and how the market is reacting when the earnings forecasts are released by analysts (Francis & Soffer, 1997; Li, 2012), and builds portfolios based on these effects (Stickel, 1991; Zhang et al., 2017). Palmon et al. (2020) examine the information conveyed by divergent analysts' recommendations. One related paper is You et al. (2013), which finds that analysts are likely to cater to the investors' prior beliefs, suggesting that it is difficult for investors to correct the analysts' biased forecasts.

This paper studies how the stock is over-priced by perceiving the analysts' biased earnings forecasts and which ex-ante variable can be used to identify the bias. Empirically, the analysts' optimistic bias cannot be observed ex-ante until the actual earnings are announced, which means that it cannot be used as a forecasting factor. Two ex-ante variables are introduced to characterize the analysts' behavior, namely the analysts' disagreement and self-selection in analysts' coverage. The existing literatures rarely analyze the pricing ability of the analysts' self-selection. The role of the analysts' disagreement in stock pricing is still an open question (Silva & Cerqueira, 2021). Varian (1985) proposes a risk theory, in which analysts' disagreement is regarded as risk, and the stock price will have a discount to compensate the uncertainty associated with the earnings. Consequently, there will be a positive return after the actual earnings are released. By the assumption of Miller (1977), Diether et al. (2002) suggest that the market is overpriced following the optimistic investors under the constraints of short selling. There

would be a negative return after the report date. In the empirical research, Doukas et al. (2006) concludes through empirical research that there is a significant positive correlation between the analysts' differences and stock returns. In contrast, Gharghori et al. (2011) have found a negative correlation between analysts' divergence and stock returns. Enormous disagreement among analysts means a discount rather than a premium to future yields. Tong et al. (2020) adopts a stochastic dominance approach to distinguish between the hypothesis.

In this paper, two theoretical models are employed to identify the pricing function of the analysts' disagreement and self-selection. The first model is based on the analysts' decision making framework developed by Scherbina (2006), some adjustments are made to its utility function to derive the expression for the analysts' optimistic bias as a function of the analysts' disagreement and self-selection. The second model is to derive how and through which channel the stocks are priced according to the analysts' disagreement and self-selection. The theoretical analysis indicates that the self-selection is negatively correlated with the stock returns due to its contribution to the overall upward bias of the analysts' average earnings forecasts. There are two channels through which the stocks are priced according to the analysts' disagreement. The first one is through the risk channel due to the fact that the earnings uncertainty is implied in the analysts' disagreement. The stock price will be discounted before the release of the earnings. The second one is the optimistic bias channel. Concerned with their reputation, analysts tend to report optimistic forecasts when the earnings are uncertain so that analysts can hide their true intentions. The optimistic bias channel means that the stock is overpriced which yields negative correlation between the analysts' disagreement and stock returns. The seemingly contradictory results in the literatures reflect the two opposite pricing mechanism of the analysts' disagreement. Empirical experiments are designed to verify the two channels by selecting different control variables.

The contribution of this paper is mainly reflected in the following aspects. Firstly, a theoretic model is developed to analyze the market returns as a response of the analysts' biased behavior, which serves as the basis for the empirical studies. Two opposite channels are identified through which the analysts' disagreement is priced and provide an explanation why contradictory empirical results are found in the literature. Secondly, this paper reports some valuable empirical findings by using the data from the Chinese markets: the analysts' forecast bias is negatively correlated with the excess stock returns, indicating that the market fail to correct the analyst' forecast bias; Analysts' self-selection bias, disagreement and are both negatively correlated with the excess stock return; After controlling the analysts' forecast bias, analysts' disagreement becomes positively correlated with the excess returns, which means that analyst disagreement contains two pricing factors, one is a risk factor for earnings uncertainty and the other is the optimistic bias.

The rest of this paper is organized as follows: the first section presents the theoretical model. The second section puts forward the research hypothesis and research design. The third part gives the empirical analysis. The fourth part further discusses the result. Finally, we conclude this study.

1. Theoretical models

1.1. Optimistic bias model

The first theoretical model is based on the Scherbina (2006)'s framework. The model reveals how the optimistic bias is generated in the analysts' earnings forecasts, which serves as the basis for the stock pricing model in the next section. Assume that all analysts have the same public signal s_0 about the earnings per-share (EPS) that satisfies the normal distribution with a mean of the real EPS denoted by e and a standard deviation σ_0, i.e. $s_0 \sim \mathrm{N}\left(e,\sigma_0^2\right)$. Apart from this, each analyst i has a private signal s_i, which is also a random number with normal distribution $s_i \sim \mathrm{N}\left(e,\sigma_s^2\right)$. The mean of the private signal is assumed to equal to the real EPS.

Because the real EPS is unknown, the analysts predict the earnings based on the public signals and their own private signals. The expectations and variances of the prediction is given as

$$\begin{cases} E_i\left[e\right] = \dfrac{\sigma_s^2}{\sigma_0^2+\sigma_s^2}s_0 + \dfrac{\sigma_0^2}{\sigma_0^2+\sigma_s^2}s_i \\ \sigma^2 = \dfrac{\sigma_s^2\sigma_0^2}{\sigma_0^2+\sigma_s^2} \end{cases}, \tag{1}$$

where $E_i\left[e\right]$ and σ indicates the expectation and the standard deviation of the analysts' forecasts, respectively. Since σ^2 is the same for all analysts, the subscript *i* is dropped. Suppose that the standard deviation of the analyst's private signal is proportional to that of the public signal, i.e. $\dfrac{\sigma_s}{\sigma_0} = v$. Then $\sigma^2 = u^2\sigma_0^2, u^2 = \dfrac{v^2}{1+v^2}$. The variance of the public signal σ_0^2 measures the earnings uncertainty. σ^2 measures how much the analysts is uncertain about her/his prediction, which is proportional to the earnings uncertainty σ_0^2.

The analyst's utility is a function of the earnings forecasts f,

$$U\left(f\right) = f - E_i\left[e\right] - \frac{\alpha}{\sigma_0^2}\left(f - E_i\left[e\right]\right)^2 - C\left|_{f<s_0-k\sigma_0}\right., \tag{2}$$

where $C\left|_{f<s_0-k\sigma_0}\right.$ indicates that if the inequality is satisfied, a penalty C will be applied. The utility function of the analysts is composed of three parts, a utility gain from releasing optimistic forecasts, a penalty of publishing biased forecasts due to reputation concern, and a constant penalty C if the analyst's prediction is lower than k standard deviations of the average value of the public signal. The magnitude of the penalty is inversely proportional to the square of the earnings uncertainty σ_0, different from inversely proportional to σ_0 in Scherbina (2006). There are two considerations for the modification, the first one is to align with the dimension of the penalty term $\left(f-EPS\right)^2$, the other is to guarantee that in the next section the price is proportional to the variance which is a common result in finance literature.

In Eq. (3), one can see that the analysts would issue the reports in scenario one and scenario two. Instead of reporting their actual prediction values $E_i\left[e\right]$, they add an optimistic bias $\sigma_0^2/\left(2\alpha\right)$ to their predictions in scenario one. While in scenario two, analysts add something greater than $\sigma_0^2/\left(2\alpha\right)$. In both cases, $\sigma_0^2/\left(2\alpha\right)$ is the lower bound

of the bias. Scenario three is the analyst's self-selection case. Analysts choose not to issue the reports when their forecasts are sufficiently low. The probability of scenario three is $Q = \Pr[E_i\left[e\right] < a]$ and the expected value of the analysts' forecasts which are not reported is $z = E[E_i\left[e\right] | E_i\left[e\right] < a]$. One can see that the value of z is less than the expected value $\overline{f}$ reported by analysts in scenario one or two.

The expected value $\overline{f}$ reported by the analysts in scenario one or two is greater than the analysts' private estimate by at least $\sigma_0^2 / (2\alpha)$, which is deliberately added by the analysts. The true earnings can be approximately estimated as the summation of the following two parts:

$$\overline{e} \approx Qz + (1-Q)\left(\overline{f} - \sigma_0^2 / (2\alpha)\right). \tag{4}$$

Compared to the true value, the overall optimistic bias in earnings reports issued by the analysts can be expressed as

$$\overline{\Delta} \equiv \overline{f} - \overline{e} \approx \frac{(1-Q)\sigma_0^2}{2\alpha} + Q\left(\overline{f} - z\right) = A + B. \tag{5}$$

The overall optimistic bias in the outstanding forecasts has two components, one part $A = (1-Q)\sigma_0^2 / (2\alpha)$ that analysts deliberately add to their private estimate, and the other part $B = Q\left(\overline{f} - z\right)$ due to the self-selection in analysts' coverage. Analysts who receive a low private signal and choose to keep quiet will bias the mean of the reported forecasts up by $\overline{f} - z$. One can see that the optimal bias which analysts add to their estimates $\sigma_0^2 / (2\alpha)$ is increasing in the level of uncertainty of the public signal σ_0^2. The component of the bias caused by the self-selection in analyst coverage also increases in forecast uncertainty. Unreported low forecasts are likely to be lower relative to the true mean the more spread out the distribution, thus creating a positive relationship between the bias and the disagreement of the reported distribution. Therefore, the expected optimistic bias $\overline{\Delta}$ in an analyst's forecast is increasing in the level of uncertainty σ_0^2.

In the empirical study, the sample mean and sample variance of the analysts' earnings forecasts are used to approximate $\overline{f}$ and σ^2 (therefore σ_0^2 since there exists a proportional relationship between them). The self-selection is estimated based on the change in the number of the analysts' reports for two consecutive quarters. For example, if the number of analysts' reports change from n for the last quarter to m for this quarter (m < n), the self-selection can be expressed as

$$B \equiv Q\left(\overline{f} - z\right) = \frac{\left(\overline{f} - z\right)(n-m)}{n} = \overline{f} - \left(m\overline{f} + (n-m)z\right) / n. \tag{6}$$

1.2. Stock pricing model

How should the stock be priced based on the analysts' biased reports? Assume that the investors take each analyst's report as independent and objective, which means that the analyst's optimistic bias is not corrected by the market. The investors regard the sample mean $\overline{f}$ and variance σ^2 of the analysts' reports as the estimate of the true mean and the true variance of

the unknown *EPS*. If the investors value the stock after the release of earnings announcement based on the price-earnings ratio *N*, the stock price would be $P_1 = N \times EPS, EPS \sim \mathrm{N}\left(\overline{f},\ \sigma^2\right)$.

Before the announcement of the actual earnings, the investors hold y shares of stock at price P_0. Because the time is too short, the time value of the money is ignored, the change of investors' wealth after the announcement is $yP_1 - yP_0$. Assume that the investors' utility is the mean-variance utility, i.e.,

$$U(y) = y\left(E\left[P_1\right] - P_0\right) - \frac{\beta}{2} y^2 Var\left(P_1\right) = y\left(N\overline{f} - P_0\right) - \frac{\beta}{2} y^2 N^2 \sigma^2, \tag{7}$$

where is the risk aversion of investors. After maximizing this utility, the investors' demand function is given as

$$y = \frac{N\overline{f} - P_0}{\beta N^2 \sigma^2}. \tag{8}$$

When the market clears, the demand of the investors equals the total supply *M* of the stock, and the stock price before the earnings announcement can be calculated as

$$P_0 = N\overline{f} - \beta MN^2 \sigma^2. \tag{9}$$

After the market knows the real earnings EPS, the stock price at that time is $N \times EPS$. If $\overline{e}$ in Eq. (4) is assumed as the unbiased estimate of the *EPS*, then $P_1 = N\overline{e}$. The stock return can be given as

$$r = \frac{P_1 - P_0}{P_0} = \frac{N\left(\overline{e} - \overline{f}\right) + \beta MN^2 \sigma^2}{P_0} = \frac{-NA - NB + \beta MN^2 \sigma^2}{N\overline{f}\left(1 - \frac{\beta MN}{\overline{f}} \sigma^2\right)}. \tag{10}$$

The denominator in Eq. (10) can be approximated as

$$\frac{1}{\left(1 - \frac{\beta MN}{\overline{f}} \sigma^2\right)} \approx 1 + \frac{\beta MN}{\overline{f}} \sigma^2 + O\left(\sigma^4\right). \tag{11}$$

Substitute formula (11) into formula (10) to obtain

$$r \approx (-C_1 + C_2)\sigma^2 - B / \overline{f}, \tag{12}$$

where $C_1 = \frac{A}{\overline{f}\sigma^2}$, $C_2 = \frac{\beta MNv^2}{\overline{f}}$. There are two terms in the stock return that are proportional to the analysts' disagreement σ^2, namely, $-C_1\sigma^2$ and $C_2\sigma^2$, which are opposite in sign. Knowing that $C_1\sigma^2 = A$, the first term reflects the contribution from the optimistic bias deliberately added by the analysts. If the bias is not corrected by the market, P_0 is overvalued based on the analysts' optimistic reports. It creates a negative correlation between the stock return and the analysts' disagreement σ^2. $C_2\sigma^2$ is the risk discount due to the earnings uncertainty (implied in the analysts' disagreement, since $\sigma^2 = u^2\sigma_0^2$), which makes a positive correlation between the stock return and the analysts' disagreement (Varian, 1985). Two

channels are identified that the stock return can be influenced by the analysts' disagreement, i.e., the optimistic bias channel and the risk channel. They have opposite effects on the stock price, which might provide an explanation why contradictory empirical results are found in the literature. On the other hand, the self-selection B has a positive contribution to the analysts' forecast bias as shown in Eq. (5), therefore the stock return is negatively correlated with the self-selection B. Figure 1 summarizes the mechanism of the analyst earnings forecasts on the stock return.

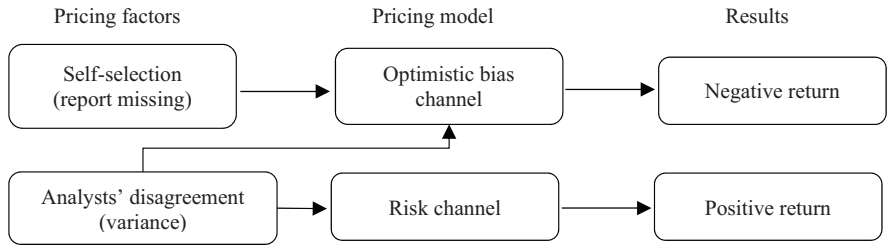

Figure 1. The mechanism of the analyst's earnings forecast on stock return

Because the analysts' disagreement σ^2, the analyst self-selection B and the analysts' overall optimistic bias Δ can be observed, different combinations of variables can be selected to identify the two channels of the analysts' disagreement on the stock return. The first regression analysis is to identify the risk channel by including the analyst's overall optimistic bias Δ and the analysts' disagreement σ^2 as the explanatory variables. Combing Eq. (5) and Eq. (12), the following relationship can be derived,

$$r \approx -\Delta / \overline{f} + C_2 \sigma^2. \tag{13}$$

It is expected to see that the analysts' overall optimistic bias Δ is negatively correlated with the stock return, and the analysts' disagreement is positively correlated with the stock return when Δ is present in the equation.

The second regression analysis is to identify the negative contribution of the self-selection B to the stock return. Since B increases with the forecast uncertainty, the negative correlation of the self-selection B with the stock return implies that the analysts' disagreement has a negative impact on the stock return through the optimistic bias channel. The regression uses the analysts' disagreement σ^2 and the self-selection B according to Eq. (12). Meanwhile, whether σ^2 is positively or negatively correlated with the stock return depends on the relative strength between the analysts' forecast bias effect C_1 and the risk effect C_2.

2. Research hypothesis

Analysts tend to publish overoptimistic forecasts. If the market does not correct the analysts' bias, the stock price would be overvalued following the analysts' optimistic forecasts and eventually would decline after the actual earnings are announced. The same is true with the analysts' self-selection bias. The investors may not notice that the analysts selectively keep silent. Therefore, this article proposes the first research hypothesis.

Hypothesis 1a: The market fails to correct the analyst's overall optimistic bias. The overall optimistic bias is negatively correlated with the excess stock return around the earnings announcement date.

Hypothesis 1b: The market fails to correct the analyst's self-selection bias. The self-selection is negatively correlated with the excess return around the stock earnings announcement date.

It should be mentioned that, the analysts' overall optimistic bias is an ex-post variable, which means that it can only be measured after the earnings are announced. Investors cannot use it to predict the stock price.

Analysts' disagreement reflects the uncertainty of the company's earnings. From the theoretical analysis, one can see the impact of the earnings uncertainty on the stock return through two channels. The first channel is the risk channel. The stock price is discounted due to the risk of the earnings implied in the analysts' disagreement, and there would be a positive correlation between the analysts' disagreement and the stock's excess return around the earnings announcement date. The second channel is the optimistic bias channel. As discussed in the theoretical analysis, the bias deliberately added by the analysts and the self-selection in analysts' coverage both increase with the earnings uncertainty. If the market prices the stock according to the analysts' overoptimistic forecasts, the stock is overvalued according to hypothesis 1. There would be a negative correlation between the analysts' disagreement and the stock's excess return. As shown in Eq. (13), the risk channel can be identified by including the analyst's overall optimistic bias Δ as a control variable and the analysts' disagreement σ^2 as the explanatory variables. This article proposes hypothesis 2.

Hypothesis 2a: The impact of the analysts' disagreement alone on the excess returns is uncertain depending on the relative strength of the risk channel and the optimistic bias channel.

Hypothesis 2b: After controlling for the analysts' overall optimistic bias, the analysts' disagreement is positively correlated with the stock excess return.

3. Empirical research

3.1. Research data sources

The A-share data from the Chinese stock market from 2010 to 2017 is used for the empirical analysis. In this paper, the analysts' earnings forecast data and all control variables are obtained from the CSMAR database, the stock price data is obtained from the WIND database, and the Fama-French three-factor data is obtained from the RESSET database. The following pre-processing is conducted: (1) remove the forecasts with missing values; (2) remove the samples when quarterly forecasts are less than 3; (4) remove financial firms; In addition, in order to ensure that the research conclusion is not affected by extreme values, this paper adopts a tail-shrinking process of up or down 1% for all continuous variables. After the preprocessing, we are left with 16230 data samples which covers 2009 firms and 568 analysts.

3.2. Definition of variables

In this section, the key variables are defined, which include the analysts' overall forecast bias, the analysts' disagreement, the self-selection in analysts' coverage, cumulative excess return, and some control variables.

The analysts' overall forecast bias is defined as the difference between the consensus analysts' forecast and the actual annual earnings, which is then normalized by dividing the absolute value of the actual EPS. For a robust check, the bias is normalized by dividing the net asset per share (Scherbina, 2006). It should be noted that, analysts always predict the company's annual earnings in each quarter. The variance of the analysts' forecasts decreases when the time approaches the end of the year. The bias is further normalized by dividing the square root of the number of quarters to the end of the year.

For a robust consideration, both median and mean of the analysts' forecast are used to represent the consensus analysts' opinion. $AEPS_{iT}$ denotes the actual annual EPS of company i in year T. $FEPS_{ijTt}$ denotes the company i's earnings forecast from analyst j in quarter t of year T. The analysts' overall forecast bias of company i in quarter t of year T can be expressed as

$$\Delta_{mean} = \frac{mean\left(FEPS_{ijTt}\right) - AEPS_{iT}}{abs\left(AEPS_{iT}\right)\sqrt{\tau}}; \tag{14}$$

$$\Delta_{median} = \frac{median\left(FEPS_{ijTt}\right) - AEPS_{iT}}{abs\left(AEPS_{iT}\right)\sqrt{\tau}}, \tag{15}$$

where τ is the number of quarters till the end of the year.

The analysts' disagreement is defined according to Diether et al. (2002) as

$$Disagreement = \frac{Var\left(FEPS_{ijTt}\right)}{\left(AEPS_{iT}\right)^2 \tau}, \tag{16}$$

where Var stands for the variance. The normalization is applied by dividing the square of the actual EPS and the number of quarters till the end of the year. According to the theoretical analysis, the analysts' disagreement is proportional to the company's earnings uncertainty.

Assume that the number of forecasts for company i is m in the t-th quarter of year T, changing from n in the t-1-th quarter. If m < n, it means that some analysts stop reporting since their forecasts are sufficiently low. We use one cent below the lowest value L of the analysts' forecast to approximate z. According to Eq. (6), the self-selection can be expressed as

$$B = \begin{cases} FEPS_{iTt} - \dfrac{(n-m)\times(L-0.01) + m\times FEPS_{iTt}}{n\times\left|AEPS_{iT}\right|\times\sqrt{\tau}}, \; if \; n > m \\ 0, \; if \; n \le m \end{cases}, \tag{17}$$

where the same normalization is applied as the analysts' overall forecast bias.

This paper uses the Fama and French (1993)'s three factor model to compute the excess return as

$$\varepsilon_{it} = R_{it} - \left(\alpha_{it} + \beta_{1i} r_{1t} + \beta_{2i} r_{2t} + \beta_{3i} r_{3t}\right), \tag{18}$$

where R_{it} is the raw return of stock i on date t, ε_{it} is the excess return. r_{1t}, r_{2t}, and r_{3t}, respectively, represents one of the returns of the Fama-French three factor portfolios, namely, market portfolio, market capitalization portfolio, and book-to market portfolio. The param-

eters α and βs are estimated using a 400-day rolling window and the excess return is computed for the subsequent month. In this manner, the parameters are gradually adapted and do not cause large disruption to the excess returns. The excess returns of three (or five) consecutive trading days from –1 (–2) to +1 (+2) are then accumulated as the cumulative excess return around the announcement date

$$CAR_{it} = \sum_{\tau=-1,-2}^{1,2} \varepsilon_{it}. \tag{19}$$

The following control variables are included: (1) Trace, which is the logarithm of the number of reports released in each quarter; (2) Size, which is the logarithm of each company's total asset at the end of each quarter; (3) Bm, the ratio of the company's book value to its market value; (4) Lev, which is the company's asset-liability ratio; (5) Institution, which is the ratio of institutional investors' holdings; (6) Top10, which is the ratio of the top 10 shareholders' holdings; (7) Turnover, which is the average daily turnover in each quarter; (8) Pe, which is price earnings ratio; (9) Geps, which is the growth rate of net assets per share; (10) β_1, β_2, β_3, which are obtained from the Fama-French's three-factor regression.

Table 1 lists the descriptive statistics of the variables used in the empirical analysis. The analysts' overall optimistic bias (either Δ_mean or Δ_mean) is positive, which indicates that there is a significant optimistic bias in analysts' forecasts. Among all the valid forecasts, 78.31% over forecast the earnings.

Table 1. Descriptive statistics for variables

Variables	Variables Discription	Mean	Median	25%	75%	S.D.
Disagreement	Analysts' disagreement	0.16	0.005	0.001	0.023	0.78
B	Self-selection bias	0.016	0.000	0.00	0.00	0.046
Trace	Analysts' total reports	8.72	7.000	4.00	11.00	6.18
Size	Company's Asset	22.46	22.23	21.39	23.29	1.41
Institution	Institutional Investors' holdings	0.19	0.136	0.068	0.26	0.17
Top10	Top 10 Shareholders' holdings	0.46	0.456	0.31	0.61	0.20
Geps	Growth rate of earnings per share	0.31	0.103	–0.20	0.39	1.45
Lev	Company's leverage	0.42	0.418	0.248	0.59	0.21
Turnover	Turnover rate	0.016	0.012	0.006	0.021	0.013
Pe	Price earnings ratio	50.25	34.16	19.95	58.06	79.76
Bm	Book-to-market value	0.96	0.544	0.31	1.11	1.12
β_1	Market factor	0.93	0.966	0.78	1.12	0.30
β_2	Market value factor	0.48	0.476	0.083	0.89	0.60
β_3	Book-to-market value factor	–0.38	–0.37	–0.82	0.064	0.69
CAR3	Cumulative excess return (3 days)	0.002	0.000	–0.02	0.025	0.05
CAR5	Cumulative excess return (5 days)	0.001	–0.003	–0.02	0.021	0.045
Δ_mean	Analysts' overall bias (mean)	0.36	0.10	0.003	0.35	0.87
Δ_median	Analysts' overall bias (median)	0.34	0.098	0.00	0.34	0.85

3.3. Empirical model

To test the hypothesis in this article, this paper uses the following regression models.

$$CAR_{it} = b_1 \Delta\_mean_{it} + \sum Control_{it}; \tag{20}$$

$$CAR_{it} = b'_1 \Delta\_median_{it} + \sum Control_{it}; \tag{21}$$

$$CAR_{it} = b_2 Disagreement_{it} + b_3 B_{it} + \sum Control_{it}; \tag{22}$$

$$CAR_{it} = b_1 \Delta\_mean_{it} + b_2 Disagreement_{it} + b_3 B_{it} + \sum Control_{it}; \tag{23}$$

$$CAR_{it} = b'_1 \Delta\_median_{it} + b_2 Disagreement_{it} + b_3 B_{it} + \sum Control_{it}. \tag{24}$$

Model (20) and (21) are used to test hypothesis 1a. It is expected to see negative b_1 and b'_1. By including B in model (22), (23) and (24), we want to test whether the self-selection is negatively correlated with the stock return. In model (22), the sign of b_2 is uncertain as suggested in hypothesis 2a. However, when the analysts' overall optimistic bias is controlled, it is expected to see positive b_2 in model (23) and (24).

3.4. Empirical analysis

Table 2 shows the correlation between each pair of variables. As suggested in the theoretical analysis, there is a significant positive correlation between the analysts' disagreement and the analysts' self-selection (B). The univariate analysis from the correlation indicates that disagreement and CAR3 (CAR5) is negatively correlated, B and CAR3 (CAR5) is negatively correlated, disagreement and CAR3 (CAR5) is negatively correlated.

Table 2. Correlation analysis

	Disagreement	B	Δ_mean	Δ_median	CAR3	CAR5
Disagreement	1.000	0.109***	0.631***	0.575***	–0.015*	–0.016**
B		1.000	0.142***	0.144***	–0.023***	–0.020**
Δ_mean			1.000	0.978***	–0.066***	–0.063***
Δ_median				1.000	–0.065***	–0.062***
CAR3					1.000***	0.892***
CAR5						1.000

This section examines whether the analysts' disagreement and self-selection can predict the stock returns. Table 3 reports the regression results according to the empirical models from Eq. (20) to Eq. (24). The first two regressions use the analysts' overall forecast bias (Δ_mean or Δ_median) as the explanatory variable to test hypothesis 1a. It shows that the overall optimistic bias is negatively correlated with the excess stock return, which means that the market has failed to discover the analysts' optimistic bias when pricing the stock. The stock is overvalued according to the analysts' overall optimistically biased earnings forecasts. Therefore, hypothesis 1a should not be rejected.

The third regression uses the analysts' disagreement and self-selection as explanatory variables to test hypothesis 1b and hypothesis 2a. The disagreement is not statistically significantly correlated with the excess return. From the theoretical analysis in Eq. (12), when both disagreement and self-selection are in the equation, there are two opposite channels from which the disagreement contributes to the excess return, namely, the optimistic bias channel and the risk channel. The coefficient of the analysts' disagreement is not significantly different from zero, which means that the contribution of the disagreement alone is uncertain depending on the relative strength of the two channels. Hence, hypothesis 2a should not be rejected. However, the analysts' self-selection is significantly negative, which means that the market fails to discover the analyst's self-selection bias. Hypothesis 1b should not be rejected. The fourth and fifth regression use the analysts' overall bias (Δ_mean or Δ_median) as a control variable to test hypothesis 2b. The analysts' self-selection is still consistently negative. The disagreement now becomes significantly positive, which illustrates that hypothesis 2b should not be rejected.

Table 3. Regression results

	CAR3				
	(1)	(2)	(3)	(4)	(5)
Δ_mean	–0.0038***				–0.0054***
	(–7.20)			(–8.01)	
Δ_median		–0.0038***			–0.0049***
		(–7.13)			(–7.56)
Disagreement			–0.0005	0.0031***	0.0024***
			(–0.91)	(4.25)	(3.48)
B			–0.0274***	–0.0196**	–0.0197**
			(–2.83)	(–2.02)	(–2.03)
Trace	0.0002**	0.0002**	0.0002**	0.0002**	0.0002**
	(2.14)	(2.17)	(2.40)	(2.01)	(2.07)
Size	–0.0019***	–0.0020***	–0.0017***	–0.0019***	–0.0018***
	(–3.09)	(–3.07)	(–2.69)	(–2.97)	(–2.94)
Institution	0.0068**	0.0068**	0.0073**	0.0069**	0.0069**
	(2.41)	(2.41)	(2.58)	(2.46)	(2.45)
Top10	0.0052**	0.0053**	0.0053**	0.0049*	0.0049*
	(2.00)	(2.00)	(2.01)	(1.86)	(1.89)
Geps	0.0013***	0.0013***	0.0016***	0.0013***	0.0013***
	(4.29)	(4.32)	(5.32)	(4.13)	(4.24)
Lev	0.0021	0.0020	0.0004	0.0022	0.0020
	(0.68)	(0.65)	(0.14)	(0.71)	(0.65)
Turnover	–0.0104	–0.0099	–0.0305	–0.0127	–0.0126
	(–0.23)	(–0.22)	(–0.69)	(–0.29)	(–0.28)

End of Table 3

	CAR3				
	(1)	(2)	(3)	(4)	(5)
Pe	-1.0×10^{-5}*	-1.1×10^{-5}*	-1.4×10^{-5}**	-1.0×10^{-5}*	−1.08e-05*
	(−1.77)	(−1.85)	(−2.40)	(−1.71)	(−1.84)
Bm	0.0016**	0.0016**	0.0015**	0.0015**	0.0015**
	(2.54)	(2.52)	(2.28)	(2.28)	(2.30)
Beta1	−0.0003	−0.0004	−0.0006	−0.0002	−0.0003
	(−0.21)	(−0.24)	(−0.37)	(−0.12)	(−0.17)
Beta2	−0.0026***	−0.0025***	−0.0027***	−0.0025***	−0.0025***
	(−2.63)	(−2.60)	(−2.81)	(−2.59)	(−2.58)
Beta3	0.0022***	0.0022***	0.0022***	0.0021**	0.0021**
	(2.69)	(2.67)	(2.60)	(2.54)	(2.53)
Year & Industry	Yes	Yes	Yes	Yes	Yes
R-squared	0.043	0.043	0.041	0.045	0.044
Number of samples	16230	16230	16230	16230	16230

So far, we have shown that statistically both the analysts' disagreement and self-selection can predict the stock returns. We will illustrate the economic significance of the two variables in two ways. First, we can check the coefficients of the variables and their average values to roughly estimate how much the return can be predicted. The coefficient of variable B is around −0.02 in regression Eq. (3)–(5). The average value of B is 0.016. Given that 75% of time B is zero, its average value would be 0.048 when B is not zero. Thus B would give a prediction of three-day price return of 0.00096, which is economically significant. Likewise, in regression Eq. (4), the coefficient of disagreement is 0.003, its average value is 0.16, which gives a 0.00048 prediction of price return.

Table 4 reports the cumulative excess returns at different quantile regions of the disagreement and self-selection B. For the disagreement variable, the difference of the excess cumulative return of the 0–25% region and the 75–100% region is 0.006 for CAR3 and 0.0078 for CAR5, which means that the investors can get the excess return of 0.006 in three days and 0.0078 in five days if they long the stocks belonging to 0–25% region and short stocks belonging to 75–100% region. The excess return of 0.006 in three days is very profitable (which corresponds to annualized return of 50% if 250 trading days are counted for a year). For B

Table 4. Cumulative excess returns at different quantiles of disagreement and B

Disagreement	CAR3	CAR5	B	CAR3	CAR5
0–25%	0.0049	0.0053	0–75%	0.0033	0.0023
25–50%	0.0024	0.0030	75–83%	0.003	0.0037
50–75%	−0.0012	−0.0021	83–92%	0.0012	0.0019
75–100%	−0.0011	−0.0025	92–100%	−0.0024	−0.0024

variable, 75% of B is zero as we know from the description statistics. We divide the top 25% into three regions. The top 92–100% region, the cumulative return is –0.0024, which is lower than the 75–83% by 0.0054. The variability of B is small, however, it has a large impact on the stock returns.

A robustness test is conducted by changing the window size to compute the betas, and by using the net asset per share as a normalizing factor in Eq. (14), (15), (16) and (17). In addition, the excess returns are computed by cumulating five consecutive days around the earnings announcement date (2 days before and 2 days after). To ensure the conciseness of the results, only the regression results of the main variables are listed. The results are shown in Table 5, which is consistent with previous conclusion[1].

Table 5. Robust test

	CAR5				
	(1)	(2)	(3)	(4)	(5)
Δ_mean	–0.0026***			–0.0038***	
	(–6.20)			(–6.94)	
Δ_median		–0.0026***			–0.0034***
		(–6.08)			(–6.48)
Disagreement			–0.0003	0.0022***	0.0017***
			(–0.73)	(3.72)	(3.02)
B			–0.0189**	–0.0135*	–0.0136*
			(–2.42)	(–1.71)	(–1.73)
R-squared	0.010	0.010	0.008	0.011	0.011
Year & Industry effects	Yes	Yes	Yes	Yes	Yes
Number of samples	16 230	16 230	16 230	16 230	16 230

Note: *, **, and *** denotes significance at the 10%, 5%, and 1% critical levels, respectively.

The results are reported in this paper by excluding the samples when the number of quarterly forecasts is less than 3. This paper has also tried to eliminate samples with less than 5 forecasts in each quarter to obtain 11559 samples. The results are completely in line with the reported results.

4. Discussion

This paper illustrates that the optimistic bias in analysts' forecasts has not been corrected by the investors. Investors tend to overprice the stocks based on the biased forecasts, which causes market inefficiency. Policies should be designed to encourage analysts not to float their

[1] We have made some heterogeneity tests. The environmental uncertainty such as the economic policy uncertainty slightly reduces the association of disagreement and B on the stock returns (the cross terms are marginally significant at 10% level). In addition, we divide the samples into large firms and small firms. The effect of the analysts' biased behavior on the stock return is more prominent in larger firms. One reason could be that the analysts forecasting data in larger firms may be less noisy since larger firms get more analysts coverage.

findings about the earnings and not to keep silent when their findings are not satisfying. Two ex-ante variables, namely analysts' disagreement and self-selection, are both associated with the optimistic bias. The two variables are shown to be able to predict short-term stock returns around the earning announcement date. Investors can exploit the market inefficiency and get abnormal returns if they long stocks with low analysts' disagreement/self-selection and short stocks with high analysts' disagreement/self-selection. It is implied in this paper that the mispricing in stock prices due to the biased analysts' reports can be corrected when the actual earnings are announced.

However, it should be noted that this article only focuses on the pricing factors in analysts' earnings forecasting behaviors. Analysts' other behaviors such as ratings, trading recommendations, etc, are not discussed in this paper. In addition, this paper focuses on the collective behavior of analysts. It would also be interesting to see the differences by following individual analysts. These topics are worth further in-depth discussion in the future.

Conclusions

In this paper, two ex-ante variables are introduced to characterize the analysts' earnings forecast behavior, namely the analysts' disagreement and self-selection in analysts' coverage. The main objective of this paper is to investigate the impact of the two variables on the stock return.

This paper first develops a theoretical analysis to derive how the stock returns are correlated with the analysts' disagreement and self-selection. The self-selection is negatively correlated with the stock return as more self-selection means that the average reported forecasts would be more optimistically biased as the low forecasting values are not revealed. The stock is overvalued and an eventual decline would be expected to follow if the stock is priced according to the optimistic reports. There are two channels through which the stocks are priced according to the analysts' disagreement. The first one is through the risk channel due to the fact that the earnings uncertainty is implied in the analysts' disagreement. The stock price will be discounted before the release of the actual earnings. The second one is the optimistic bias channel. The optimistic bias channel means that the stock is overpriced if the investors can not correct the analysts' bias and the stock is priced according to the optimistic reports. This paper empirically identifies the two opposite channels through which the analysts' disagreement is priced and provide an explanation why contradictory results are found in the literature.

References

Chu, J., Qin, X., & Fang, J. (2019). Margin-trading, short-selling and analysts' forecast optimism. *Management World*, *1*, 151–228. https://doi.org/10.19744/j.cnki.11-1235/f.2019.0011

Diether, K. B., Malloy, C. J., & Scherbina, A. (2002). Differences of opinion and the cross section of stock returns. *Journal of Finance*, *57*(5), 2113–2141. https://doi.org/10.1111/0022-1082.00490

Doukas, J. A., Kim, C., & Pantzalis, C. (2006). Divergence of opinion and equity returns under different states of earnings expectations. *Journal of Financial Markets*, *9*(3), 310–331. https://doi.org/10.1016/j.finmar.2006.02.002

Francis, J., & Soffer, L. (1997). The relative informativeness of analysts' stock recommendations and earnings forecast revisions. *Journal of Accounting Research*, *35*(2), 193–211. https://doi.org/10.2307/2491360

Gharghori, P., See, Q., & Veeraraghavan, M. (2011). Difference of opinion and the cross-section of equity returns: Australian evidence. *Pacific-Basin Finance Journal*, *19*(4), 435–446. https://doi.org/10.1016/j.pacfin.2011.03.004

Gilson, S. C., Healy, P. M., Noe, C. F., & Palepu, K. G. (2001). Analyst specialization and conglomerate stock breakups. *Journal of Accounting Research*, *39*(3), 565–582. https://doi.org/10.1111/1475-679X.00028

Hu, F., & Xia, Y. (2017). Analysts' commercial motivations and earnings forecast bias: Evidence from margin transactions. *Journal of Finance and Economics*, *43*(07), 45–56. https://doi.org/10.16538/j.cnki.jfe.2017.07.004

Irvine, P., Simko, P. J., & Nathan, S. (2004). Asset management and affiliated analysts' forecasts. *Financial Analysts Journal*, *60*(3), 67–78. https://doi.org/10.2469/faj.v60.n3.2622

Jackson, A. R. (2005). Trade generation, reputation, and sell-side analysts. *Journal of Finance*, *60*(2), 673–717. https://doi.org/10.1111/j.1540-6261.2005.00743.x

Leuz, C. (2003). ADRs, analysts, and accuracy: Does cross listing in the United States improve a firm's information environment and increase market value?. *Journal of Accounting Research*, *41*(2), 347–362. https://doi.org/10.1111/1475-679X.00107

Li, L. (2012). Can analyst's earning forecast be the best proxy for "earnings expectations"? *Nankai Management Review*, *15*(6), 44–50+84. https://doi.org/10.3969/j.issn.1008-3448.2012.06.006

Lim, T. (2001). Rationality and analyst forecast bias. *The Journal of Finance*, *56*(1), 369–385. https://doi.org/10.1111/0022-1082.00329

Lin, H. W., & Mcnichols, M. F. (1998). Underwriting relationships, analysts' earnings forecasts and investment recommendations. *Journal of Accounting and Economics*, *25*(1), 101–127. https://doi.org/10.1016/S0165-4101(98)00016-0

Mcnichols, M., & O'Brien, P. C. (1997). Self-selection and analyst coverage. *Journal of Accounting Research*, *35*, 167–199. https://doi.org/10.2307/2491460

Miller, E. M. (1977). Risk, uncertainty, and divergence of opinion. *Journal of Finance*, *32*(4), 1151–1168. https://doi.org/10.1111/j.1540-6261.1977.tb03317.x

Mola, S., & Guidolin, M. (2009). Affiliated mutual funds and analyst optimism. *Journal of Financial Economics*, *93*(1), 108–137. https://doi.org/10.1016/j.jfineco.2008.06.006

Palmon, D., Sarath, B., & Xin, H. C. (2020). Bold stock recommendations: Informative or worthless?. *Contemporary Accounting Research*, *37*(2), 773–801. https://doi.org/10.1111/1911-3846.12555

Scherbina, A. D. (2006). *Analyst disagreement, forecast bias and stock returns*. SSRN. https://doi.org/10.2139/ssrn.894381

Silva, D., & Cerqueira, A. (2021). Does investors' divergence of opinion affect stock mispricing?. *ACRN Journal of Finance and Risk Perspectives*, *10*(1), 1–24. https://doi.org/10.35944/jofrp.2021.10.1.001

Stickel, S. E. (1991). Common stock returns surrounding earnings forecast revisions: More puzzling evidence. *Accounting Review*, *66*, 402–416. https://www.jstor.org/stable/247762

Tong, S. K., Grundy, B. D., Chen, N., Huang, M., Kobayashi, T., & Titman, S. (2020). Dispersion in analysts earnings forecasts and market efficiency. *International Review of Finance*, *20*(1), 247–260. https://doi.org/10.1111/irfi.12204

Varian, H. R. (1985). Divergence of opinion in complete markets: A note. *Journal of Finance*, *40*(1), 309–317. https://doi.org/10.1111/j.1540-6261.1985.tb04951.x

Wang, H., & Xu, X. (2017). A study on the optimism of recommendations on the part of security analysts in the equity private placements. *Journal of Nanjing University of Audit*, *14*(03), 47–56.

You, J., Qiu, S., & Liu, C. (2013). "Changed face" phenomena of security analysts 'forecasting behaviors: A reputation game model and evidences. *Journal of Management Sciences in China*, *16*(6), 67–84. https://doi.org/10.3969/j.issn.1007-9807.2013.06.007

Zhang, R., Wang, R., & Wang, S. (2017). Analysts' revisions, fundamental analysis and future stock returns. *Financial Research*, *445*(7), 156–174. http://www.jryj.org.cn/CN/Y2017/V445/I7/156

Zhao, L., Li, Z., & Liu, J. (2013). The managers' preferences, the optimization in the evaluation of the investment level and the obtainment of the private information. *Management World*, *4*, 33–47. https://doi.org/10.19744/j.cnki.11-1235/f.2013.04.004

Zheng, S. (2019). Managerial ability and analyst optimism bias. *Modern Finance and Economics-Journal of Tianjin University of Finance and Economics*, *39*(12), 97–110. https://doi.org/10.19559/j.cnki.12-1387.2019.12.007