International Review of Financial Analysis 86 (2023) 102533

Contents lists available at ScienceDirect

International Review of Financial Analysis

journal homepage: www.elsevier.com/locate/irfa

Sentiment and covariance characteristics

Vu Le Tran[1]

Nord University Business School, Norway

ARTICLE INFO

Keywords:
Characteristics
Sentiment model
Factor model
Risk

ABSTRACT

We propose a bridging model that connects risk-based factor models to sentiment models by using stock characteristics from the asset pricing literature. Investors use stock characteristics as information to form their biased view and hence creating mispricing in stock's price from its fundamental value. Characteristics also serve as the proxy for the covariance risk to a latent factor. The α from our factor model of mispricing ranges from 0.70% to 1.38% monthly after controlling for other common factors and mispricing measures. Well-known anomalies are only represented in either underpriced or overpriced stocks but not in all the cross-section.

1. Introduction

We construct a bridging model that connects a sentiment model, where investors disagree about the stock's expected future payoff, to a risk-based model. Stocks' characteristics enter our model with two roles. First, on the risk-based channel, they serve as proxies for the exposure to a latent factor, which is difficult to measure directly. The latent factor supposes to represent the aggregate-bias-belief about stocks' pay-off of all investors in the markets. The aggregate-bias-belief represents the common bias belief among investors in the markets about the stocks' return (or pay-off). Even though the belief is not right, if there are enough investors in the markets believe in it and trade stocks based on it, this common bias will impact the stocks' return. Second, on the sentiment-based channel, characteristics serve as information that investors use to form their own sentiment about the stocks' payoff. In our model, the mispricing return due to sentiment can also be seen as the return premium for the exposure to a latent factor.

Our model shows that the expected future return of a stock will negatively correlate with the investors' aggregate-bias-belief about its future payoff (or return). The aggregate-bias-belief about the stock's future return will immediately rise up or shoot down the stock's current price, thus creating some overpriced and underpriced stocks. In the next period, price correction occurs; therefore, the unbiased expected return of the next period is negatively correlated with the aggregate-bias-belief return. Using ex-ante information, the portfolio's α of going long[2] on under-valued stocks and going short[3] on over-valued stocks is statistically and economically significant from 0.7% to 1.38% on a monthly average, after controlling for other common factors and sentiment measures.

We perform empirical analyses to prove these above conclusions. To do that, on the sentiment view, we assume that investors use stock's characteristics as information to form their own biased belief about the stock's future payoff (or return). On the risk view, this assumption is equivalent to Kelly, Pruitt, and Su (2019)'s approach of using characteristics as proxies for latent factor exposure. We use a linear system and propose methods to estimate the importance of each characteristic ($\mathbf{f_t}$) to the mispricing return ($\boldsymbol{\theta}_{t+1}$) on the sentiment view. Each element in the vector $\mathbf{f_t}$ is the impact of each characteristic on the stock returns, and being also a long-short zero-net investment portfolio.

In the combined view, the mispricing return can be seen as the return premium from the exposure to a latent factor. This premium can be represented as a mimicking portfolio that combines different characteristics-based portfolios from $\mathbf{f_t}$. Thus, we impose a factors structure on the aggregate-bias-belief return and on the unbiased expected return.

Our approach shares the same spirit with Kozak, Nagel, and Santosh (2018)'s work, which proves that there is no significant difference between a sentiment model and a reduced-factors model. Daniel and Titman (1997) and Brennan, Chordia, and Subrahmanyam (1998) also suggest that one can test the behavioral effects on stock returns by investigating the returns associated with firm's characteristics[4] that are

E-mail address: vu.l.tran@nord.no.

[1] Many thanks for the research grant 309603 from the Norwegian Research Council and NorQuant AS, and for the comments from Stephen Dimmock, Ilan Cooper, Costas Xiouros, Sjur Westgaard, Petri Jylhä at seminars and conferences. The previous version of the paper is named "Mispricing Characteristics".

[2] Going long means buying and holding a stock or a portfolio of stocks. The investors will benefit when stocks' price increases.

[3] Going short means borrowing a stock or a portfolio of stocks and then selling them right away to get cash. Investors have to return the stock (or a portfolio of stocks) at the agreed time in the future. Therefore, when stocks price goes down, investors can buy them back cheaper and get profit from the difference.

[4] In this study, characteristics are the book-to-market ratio, firm size, the stock price, the dividend yield, and lagged returns.

https://doi.org/10.1016/j.irfa.2023.102533
Received 30 August 2022; Received in revised form 1 December 2022; Accepted 10 January 2023
Available online 18 January 2023
1057-5219/

orthogonal (or not related) to other risk factors' returns.[5] Kelly et al. (2019) also use characteristics as instruments for the exposure to latent factors.

We perform the estimation every month and take the average of the estimates. We only found a few characteristics (4 to 11) that are statistically and economically significant to the unbiased expected return. These most important characteristics are related to size, momentum, liquidity, and volatility. This result corroborates with the recent line of research (see, for example, Green, Hand, & Zhang, 2017; Harvey, 2017; Harvey, Liu, & Zhu, 2016; Hou, Xue, & Zhang, 2015, 2020; Kelly et al., 2019; Kozak et al., 2018). Although in each cited paper above, the number of significant characteristics, factors is different. They share a common point that only a few characteristics, factors, or anomalies have significant impacts on assets' returns. The most significant characteristics in these papers are also related to size, price momentum, liquidity, and volatility. In our empirical analyses, 11 of 101 characteristics offer an annual return premium greater than 3%. When we raise the bar to 5%, there are only 4/101.

Our work contributes to the literature in several ways. First, the paper continues the theoretically discussion of mispricing in Daniel, Hirshleifer, and Subrahmanyam (2001) and Kozak et al. (2018). However, our paper is different from theirs. Indeed, while Daniel et al. (2001) use investors' overconfidence as the main source of mispricing, Kozak et al. (2018) mention the source of mispricing is sentiment, we extend that view by stressing that the biased view about the future pay-off is the main source of mispricing. This approach helps us to construct a bridge between the sentiment pricing model and the risk-based factor model. Indeed, we show that the mispricing return can also be presented as a return premium to the exposure to a latent risk factor. This is the main contribution of this paper.

For the second contribution, while Daniel et al. (2001), Kozak et al. (2018) only focus on building a theoretical model with sentimental investors who trade with arbitrageurs, we perform both theoretical and empirical work to find out how large the mispricing return is. Moreover, while Kozak et al. (2018) used principal components (PCs) from the covariance matrix of *portfolios*' returns as factors to explain assets' returns, we use factors based on characteristics from *individual stocks*. Estimating the PCs from the covariance matrix of individual stocks' returns are inefficient due to the large dimension of the covariance matrix and the unbalanced panel data.[6] Hence, estimating PCs from portfolio returns is more favorable. However, using portfolios as test assets, we may omit some information at the individual stock level. Furthermore, factors from PCs do not give us a clear idea of which forces are behind these factors. Our approach would thus help us in two ways (i) knowing which characteristics impact the asset return, and (ii) directly using individual stocks but not portfolios as test assets to save information at the individual stock level.

Third, empirically, our work is also aligned with Avramov, Chordia, Jostova, and Philipov (2019), Baker and Wurgler (2006), Liu, Stambaugh, and Yuan (2018), Stambaugh, Yu, and Yuan (2012, 2015), Stambaugh and Yuan (2017). Baker and Wurgler (2006) build sentiment indexes based on 6 underlying proxies, the other studies uses 11 prominent anomalies to build mispricing measures, and mispricing factors.[7] The literature on sentiment and stock return is massive and continues to growth.[8] While we share a common result with previous work on the impact of sentiment on the mispricing phenomenon, and on the cross-sectional stocks' returns, our work contributes to the literature by adding another layer to these cited works by modeling which mechanism can lead to this mispricing. We also find that some well-known anomalies and risk premia, such as β, idiosyncratic volatility, book-to-market, momentum in the last 12 months, maximum return, investment, and liquidity volatility, only exist in overpriced stocks. In underpriced stocks, the premia change sign (β, idiosyncratic volatility, maximum return, investment, and liquidity volatility) or disappear (book to market, momentum in the last 12 months). We show that the size premium and the liquidity premium only exist in underpriced stocks. These premia change signs in overpriced stocks. Our result corroborates the result of the premia on size (Baker & Wurgler, 2006), on the idiosyncratic volatility anomaly (Stambaugh et al., 2015), and on the β anomaly (Liu et al., 2018) when conditioning on mispricing (or sentiment). In addition, we contribute to the literature by enlarging these studies by investigating more anomalies conditioning on mispricing.

For the fourth contribution, we show that the gap in expected returns between underpriced and overpriced stocks generated from both ex-post and ex-ante information in our model cannot be explained by other mispricing factors such as the Stambaugh and Yuan (2017) mispricing factors, nor by the sentiment measure of Baker and Wurgler (2006), nor by common factors such as the Fama and French (2015) 5 factors plus the momentum factor. The premium is from 0.7% to 1.38% monthly after controlling for these factors. Thus, there will be more opportunities to use the spread premium from our model as a new complement mispricing factor to the previous factors in the literature.

Fifth, we discover that the aggregate-bias-belief return can have some very extreme values. However, it is interesting that the cross-sectional distribution of this mispricing return is quite symmetric and has a mean closes to zero. This would mean that on average, every month, almost half of the stocks are underpriced, and the other half are overpriced. Despite having extreme values in some crisis periods, the aggregate-bias-belief *market* return also has a small mean which cannot be different from zero over time. Hence, it is likely that the biased return on the *market* is corrected over time. It also means that the bias at the market-wide level is small in the long run.

The rest of the paper consists of five parts. In Section 2, we demonstrate our sentiment model. Section 3 provides the alternative risk-based explanation. In Section 4, we discuss the logic of using characteristics in both sentiment and risk-based channels and describe the way to estimate the mispricing part in the stocks' returns. In Section 4.3, we clarify the data we use. Section 5 presents empirical results. Finally, Section 6 concludes.

2. The sentiment model

Here, we demonstrate our model. To build a sentiment model, we need to define what is the sentiment component, and what is the rational component (risk-based component) that can impact the stocks' return. Therefore, we start by laying out the risk-based component in our model, which is the classic one to form the CAPM model. Then we improve the model by adding the sentiment-based component to the model. Later in Section 3, we also show that the sentiment part can also be represented as a latent risk factor.

Our model will start with the personal optimization problem of consumption and investment. Each person tries to optimize this problem

[5] The risk factors in that study are the first five (asymptotic) principal components of the cross-sectional stock returns and the Fama French three factors.

[6] Estimating the whole covariance matrix of stocks return is usually not feasible because of the large size of the cross-section of stocks relative to the short length of their available time series. To make matter worse, unbalanced panel data usually incurs missing data that make the process even more impossible.

[7] Stambaugh et al. (2012, 2015) use 11 anomalies to form mispricing measures. Liu et al. (2018), Stambaugh and Yuan (2017) build mispricing factors from these 11 anomalies. Avramov et al. (2019) uses 12 anomalies to build mispricing measures.

[8] We can name for a few recent examples (Al-Nasseri, Menla Ali, & Tucker, 2021; Baker, Wurgler, & Yuan, 2012; Benhabib, Liu, & Wang, 2016; Cortés, Duchin, & Sosyura, 2016; Dong & Gil-Bazo, 2020; Edmans, Fernandez-Perez, Garel, & Indriawan, 2022; Fang, Chung, Lu, Lee, & Wang, 2021; Gong, Zhang, Wang, & Wang, 2022; Islam, 2021; Jiang, Lee, Martin, & Zhou, 2019; Kim, Ryu, & Yang, 2021; Obaid & Pukthuanthong, 2022; Song & Yu, 2022).

using all available information about the firms that he invests in. The optimization problem takes into account the personal view of firms' future pay-offs. Then we aggregate all individual decisions to study the market reaction, the equilibrium price, and the return.

For the sake of clarifying our notation, we will present a scalar in normal font with uppercase or lowercase characters, for example (a or A), a vector with lowercase characters in bold ($\mathbf{a}$), and a matrix with capital letters in bold ($\mathbf{A}$). Consider an economy that consists of I individuals. These individuals live only two periods, t and $t+1$. They were born in period t with individual wealth $\mathbf{v_t} = (v_{1,t}, \ldots, v_{i,t}, \ldots, v_{I,t})^T$ a constant absolute risk aversion, where T is the transpose operator. There are S securities to be traded on the market with shares outstanding as $\mathbf{s_t} = (s_{1,t}, s_{2,t}, \ldots, s_{S,t})^T$, price $\mathbf{p_t}$, $\mathbf{p_{t+1}}$, and dividend $\mathbf{d_t}, \mathbf{d_{t+1}}$.

In each time period t, individual i invests one part of his wealth in a portfolio of securities $\mathbf{x_{i,t}} = (x_{1,t}, x_{2,t}, \ldots, x_{s,t}, \ldots, x_{S,t})^T$ and another in a risk-free asset with a gross rate of return R_f. Individual i has a constant absolute risk aversion γ_i. He tries to maximize his expected utility function: $-E_{i,t}[e^{-\gamma_i \cdot v_{i,t+1}}]$, where $v_{i,t+1}$ is his wealth level in period $t+1$, which is also his final consumption, and $E_{i,t}[.]$ is his personal expectation at time t.

By investing in both risk-free and risky assets and assuming that his final wealth is normally distributed, maximizing $-E_{i,t}[e^{-\gamma_i \cdot v_{i,t+1}}]$ is equivalent to maximizing the Lagrangian function L_i:

$$\max_{\mathbf{x_{i,t}}} L_i = (\mathbf{x_{i,t}})^{\mathbf{T}} E_{i,t}[\mathbf{p_{t+1}} + \mathbf{d_{t+1}} - R_f \cdot \mathbf{p_t}] - \frac{\gamma_i}{2}(\mathbf{x_{i,t}})^{\mathbf{T}} \mathbf{\Omega}_t (\mathbf{x_{i,t}}) \tag{1}$$

where $\mathbf{\Omega}_t$ is the covariance matrix of $\mathbf{p_{t+1}} + \mathbf{d_{t+1}}$ estimated at time t, and R_f is the gross risk free rate. The gradient of L is then equal to

$$\nabla_{\mathbf{x_{i,t}}} L^i = E_{i,t}[\mathbf{p_{t+1}} + \mathbf{d_{t+1}} - R_f \cdot \mathbf{p_t}] - \gamma_i \mathbf{\Omega_t} \mathbf{x_{i,t}} \tag{2}$$

Solving for the first-order condition $\nabla_{\mathbf{x_{i,t}}} L^i = 0$, we can have the optimal portfolio for person i at time t, $\mathbf{x^*_{i,t}}$, as follows:[9]

$$\mathbf{x^*_{i,t}} = \frac{1}{\gamma_i} \mathbf{\Omega_t^{-1}} E_{i,t}[\mathbf{p_{t+1}} + \mathbf{d_{t+1}} - R_f \cdot \mathbf{p_t}] \tag{3}$$

This is the optimal case when individual i can possibly get $\mathbf{x^*_{i,t}}$. We herein assume that the covariance matrix is the common knowledge that everyone agrees on. The only difference is the personal expectations of $E_{i,t}[\mathbf{p_{t+1}} + \mathbf{d_{t+1}} - R_f \cdot \mathbf{p_t}]$ and the personal risk aversion level. These differences lead to a different portfolio allocation for each person.

In our model settings, we assume that the personal expectation of $E_{i,t}[\mathbf{p_{t+1}} + \mathbf{d_{t+1}} - R_f \cdot \mathbf{p_t}]$ will consist of two parts:

$$E_{i,t}[\mathbf{p_{t+1}} + \mathbf{d_{t+1}} - R_f \cdot \mathbf{p_t}] = E_t[\mathbf{p_{t+1}} + \mathbf{d_{t+1}} - R_f \cdot \mathbf{p_t}] + \mathbf{b_{i,t+1}} \tag{4}$$

The first one, $E_t[\mathbf{p_{t+1}} + \mathbf{d_{t+1}} - R_f \cdot \mathbf{p_t}]$, is the true unbiased expected value. The second one is the biased belief about future prices and future dividends. We call these biases $\mathbf{b_{i,t+1}}$. This vector $\mathbf{b_{i,t+1}} = (b_{i,t+1,1}, b_{i,t+1,2}, \ldots, b_{i,t+1,s}, \ldots, b_{i,t+1,S})^T$ contains the current biased view of the investor i at time t regarding the future price and dividend at time $t+1$ of all stocks on the market. This biased belief can arise because of (i) the investor's irrational expectation (or sentiment) or (ii) the investor's rational expectation, but based on insufficient information. These biases ($\mathbf{b_{i,t+1}}$) of course can be zero if the investor knows the true data-generating process. We do not pay attention to $R_f \cdot \mathbf{p_t}$ because it is known at time t. Then, we can rewrite the optimal portfolios for individual i at time t as follows:

$$\mathbf{x^*_{i,t}} = \frac{1}{\gamma_i} \mathbf{\Omega_t^{-1}} E_t\left[\mathbf{p_{t+1}} + \mathbf{d_{t+1}} - R_f \cdot \mathbf{p_t}\right] + \frac{1}{\gamma_i} \mathbf{\Omega_t^{-1}} \cdot \mathbf{b_{i,t+1}} \tag{5}$$

This is the demand function for individual i. This function holds not only for biased investors (we call $i = bi$) but also for arbitrageurs ($i = ar$) who do not have a biased view. The demand function shares similar settings with the one of Kozak et al. (2018). Indeed, if the investor i has an unbiased expectation about future payoff, then $\mathbf{b_{i=ub,t+1}} = 0$. Then his demand function is just $\mathbf{x^*_{i=ub,t}} = \frac{1}{\gamma_{i=ub}} \mathbf{\Omega_t^{-1}} E_t\left[\mathbf{p_{t+1}} + \mathbf{d_{t+1}} - R_f \cdot \mathbf{p_t}\right]$.

The major difference in stocks holding between the biased investor ($i = bi$) and the unbiased one ($i = ub$) is $\frac{1}{\gamma_{i=bi}} \mathbf{\Omega_t^{-1}} \cdot \mathbf{b_{i=bi,t+1}}$. This additional component reflects his private sentimental view.

In brief, the individual demand function in Eq. (5) holds with all types of investors. In the competitive equilibrium market where supply $\mathbf{s_t}$ is equal to demand, we have

$$\mathbf{s_t} = \sum_i \mathbf{x^*_{i,t}} = \frac{1}{\gamma} \mathbf{\Omega_t^{-1}} E_t[\mathbf{p_{t+1}} + \mathbf{d_{t+1}} - R_f \cdot \mathbf{p_t}] + \mathbf{\Omega_t^{-1}} \sum_i \frac{1}{\gamma_i} \cdot \mathbf{b_{i,t+1}} \tag{6}$$

where the aggregate risk aversion is γ as $1/\gamma = \sum_i 1/\gamma^i$. We define the gross return of securities s as $R_{s,t+1} = (p_{s,t+1} + d_{s,t+1})/p_{s,t}$. From Eq. (6), we can derive the expression for the rational expectation of $R_{s,t+1}$ as follows:

Proposition 1. *The unbiased expected return of security s, which is $E[R_{s,t+1}]$, can be presented by a β model as*

$$E_t[R_{t+1,s}] = R_f + \beta_{t,s} \cdot \left[E_t[R_{t+1,M} - R_f] + \sum_{k=1}^{S} w_{t,k} \left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,k}}{p_{t,k}} \right) \right] - \sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}} \tag{7}$$

where M is the market portfolio and $\beta_{t,s} = \frac{Cov_t\left(R_{t+1,s}, R_{t+1,M}\right)}{Var_t(R_{t+1,M})}$ and $w_{t,k} = \frac{p_{t,k} \cdot x_{t,k}}{\mathbf{p_t^T} \cdot \mathbf{s_t}}$.

The term $\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}}$ is the investors' aggregate-bias-belief about stock "s" return. The term $\sum_{k=1}^{S} w_{t,k} \left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,k}}{p_{t,k}} \right)$ is the investors' aggregate-bias-belief about market return.

Proof. Appendix A.

The first proposition proposes the relationship of the true unbiased expected return of stock s with market return and the investors' aggregate-bias-belief about stocks' return. In detail, the next period's return is expected to be (i) positively correlated with its covariance risk, which relates to the market excess return $\beta_{t,s} \cdot E\left[R_{t+1,M} - R_f\right]$; (ii) positively correlated with its covariance risk with the investors' aggregate-bias-belief about the *total market return*; and (iii) negatively correlated with the investors' aggregate-bias-belief about *this stock's return*. We summarize Eq. (7) with a graphical presentation as follows in Fig. 1.

For convenience, from now on, we just call the unbiased expected return "expected return". In our model, the investors' aggregate-bias-belief about the whole market's return or a specific stock's return is made in period t for the next period $t+1$.

Note that Eq. (7) will become a standard CAPM model (Lintner, 1965; Mossin, 1966; Sharpe, 1964; Treynor, 1961, 1962) when $\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}} = 0$. This only happens in two cases. First, all investors' personal beliefs about stock s pay-off are in fact unbiased beliefs. Hence, $\frac{b_{i,t+1,s}}{p_{t,s}} = 0$, $\forall i$. This case is unlikely to happen.

Or, in the second case, the arbitrage force is efficient enough to eliminate all the effects of sentiment. This is because the arbitrageurs in the market will efficiently take the opposite trade direction from the biased investors. In this case, $\sum_i \frac{1}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}} = 0$ for each security *"s"*. When this case happens, we also have the CAPM model. The security demand function in Eq. (6) will become $\mathbf{s_t} = \sum_i \mathbf{x^*_{i,t}} = \frac{1}{\gamma} \mathbf{\Omega_t^{-1}} E_t[\mathbf{p_{t+1}} + \mathbf{d_{t+1}} - R_f \cdot \mathbf{p_t}]$. This demand function is similar to the demand function when all investors are unbiased. Previous studies show that a complete arbitrage is less likely due to arbitrage constraints such as short-sale, liquidity, funding constraints, etc. For example, Stambaugh et al. (2012,

[9] See Cochrane (2009) for a deep discussion about this optimization procedure using the Constant Absolute Risk Aversion utility function.

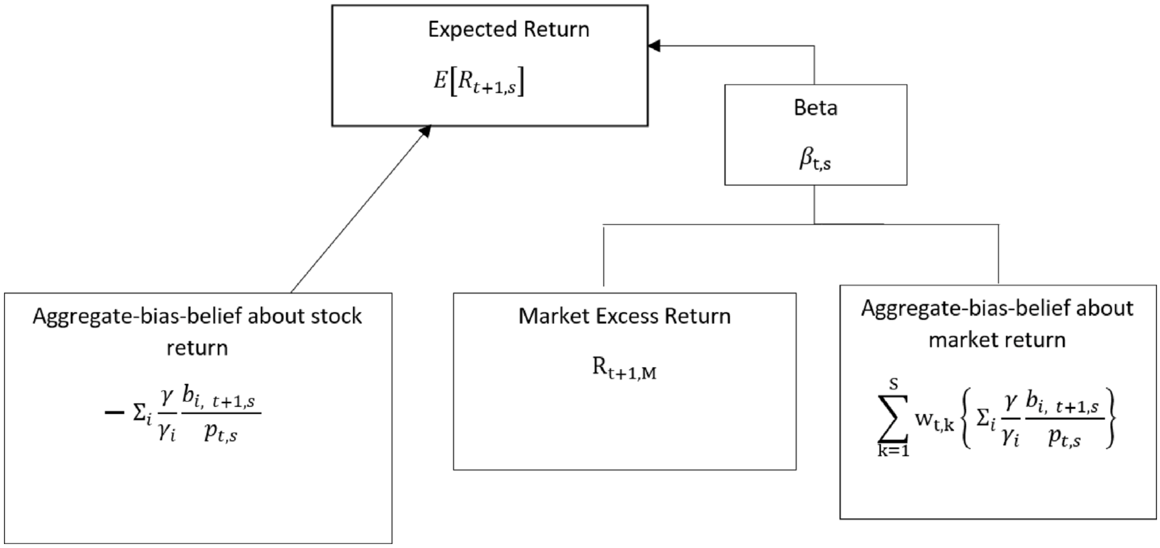

Fig. 1. Unbiased expected return under the sentiment model.

2015), Stambaugh and Yuan (2017) show how short-sale impediments are a major force restricting arbitrage on overpriced stocks.

In the empirical part, we show that these two cases are not likely to hold. The first proposition also offers us two interesting corollaries about the future return.

Corollary 1. *From Proposition 1 of our sentiment model, when keeping everything else constant, if the investors' aggregate-bias-belief (at time t) about the future market's return (at time $t+1$) increases, then the future return (at time $t+1$) of a high-β stock will increase more than the future return (at time $t+1$) of a low-β stock and vice versa.*

In our model, stocks not only co-move with the market's return but also co-move with the investors' aggregate-bias-belief about the market's future return. For each stock, these covariance-risks are specified by its β. These covariance-risks should offer a premium return.

Corollary 2. *From Proposition 1 of our sentiment model, when keeping everything else constant, if the investors' aggregate-bias-belief (at time t) about one specific future stock's future return (at time $t+1$) increases, then this stock's future return (at time $t+1$) will decrease and vice versa.*

If we have two stocks with the same β, the one that the market has a positive bias will likely have a lower future return than the other.

We already investigated the effect of $\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}}$ and $\sum_{k=1}^{S} w_{t,k}$ $\left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,k}}{p_{k,t}}\right)$ on the unbiased expected return. However, these biases are made in the present; therefore, they also have an effect on the current return and on the current price.

Proposition 2. *The current return $R_{t,s}$ and the current price $p_{t,s}$ of security s are*

(i) Positively correlated with the investors' aggregate-bias-belief (at time t) about the "stock s return" at time "t+1" $\left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}}\right)$.

(ii) Negatively correlated with the investors' aggregate-bias-belief (at time t) about the "market return" at time "t+1" $\left[\sum_{k=1}^{S} w_{t,k} \cdot \left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,k}}{p_{t,k}}\right)\right]$ when its $\beta_{t,s} > 0$.

(iii) Positively correlated with the investors' aggregate-bias-belief (at time t) about the "market return" at time "t+1" $\left[\sum_{k=1}^{S} w_{k,t} \cdot \left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,k}}{p_{t,k}}\right)\right]$ when its $\beta_{t,s} < 0$

Proof. Appendix B

Proposition 2 is consistent with Corollaries 1 and 2. In that sense, keeping everything else constant, the *current return* (and *current price*) of a high-β stock will decrease more than the *current return* (and *current price*) of low-β stock, when the overall investors' aggregate-bias-belief (in time t) about the total market's *future* return increases. In other words, a stock with high co-movement risk with the bias of future market return should offer a lower price now and thus have a higher return in the future.

Keeping everything else constant, if the investors' aggregate-bias-belief about one specific stock's *future* return increases, then this stock's *current return* (and *current price*) will increase. Now, this stock is overvalued. This *current* overvaluation will lead to a decrease in *future* return for this stock.

The Proposition 2, Corollaries 1 and 2 entail the possibility of hypotheses testing. In the empirical part, we thus test whether this proposition and two corollaries hold or not.

3. Alternative risk-based explanation

The above model shows that the stock's expected returns can be explained through the sentiment channel because of the biased part in return. The way that we decompose the individual expectation of stock's pay-off to a rational and an irrational part is different than the overconfidence approach of Daniel et al. (2001). This approach will help us to show that there is also a risk-based explanation for the biased part in the return.

Indeed, we consider the risk-based explanation or the sentiment-based explanation as two sides of the same coin in our model settings. To clarify that point, we will show that the biased part in the return, $\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}}$, in Eq. (7) can also be represented as a covariance risk between the return and a latent factor related to the aggregate-bias-belief on the market. Hence, the above exercise of mapping sentiment to the biased part in return is equivalent to mapping the covariance risk (or the stock return's exposure) to this factor to the expected returns.

In a sample space S, considering that the unbiased expectation of future payoff is constructed with a probability measure $\mathbb{P}$, and the personal biased expectation is from the probability measure $\mathbb{Q}^i \ll \mathbb{P}$ for every investor i, these probability measures are defined on sigma algebra $\mathcal{F}_t$ at time t and sigma-algebra $\mathcal{F}_{t+1}$ at time t+1, where $\mathcal{F}_t \subseteq \mathcal{F}_{t+1}$.

So, we can write the unbiased expectation as $E_t[\mathbf{p_{t+1}} + \mathbf{d_{t+1}}] = E^{\mathbb{P}}[\mathbf{p_{t+1}} + \mathbf{d_{t+1}} | \mathcal{F}_t]$ and the biased expectation of investor i as $E_{i,t}[\mathbf{p_{t+1}} + \mathbf{d_{t+1}}] = E^{\mathbb{Q}^i}[\mathbf{p_{t+1}} + \mathbf{d_{t+1}} | \mathcal{F}_t]$. Then, following the Radon–Nikodym theorem,[10] there exist

$$\begin{aligned} Random\ variable\ L_{i,t+1} &= \frac{d\mathbb{Q}^i}{d\mathbb{P}} \geq 0\ measurable\ on\ \mathcal{F}_{t+1} \\ Random\ variable\ L_{i,t} &= \frac{d\mathbb{Q}^i}{d\mathbb{P}} \geq 0\ measurable\ on\ \mathcal{F}_t \\ And\ L_{i,t} &= E^{\mathbb{P}}[L_{i,t+1}|\mathcal{F}_t] \end{aligned}$$

Combined with the abstract Bayes' formula, these above random variables satisfy the changing probability measure between $\mathbb{P}$ and $\mathbb{Q}^i$ as $E^{\mathbb{Q}^i}[\mathbf{p_{t+1}} + \mathbf{d_{t+1}}|\mathcal{F}_t] = \frac{E^{\mathbb{P}}[L_{i,t+1} \cdot (\mathbf{p_{t+1}} + \mathbf{d_{t+1}})|\mathcal{F}_t]}{E^{\mathbb{P}}[L_{i,t+1}|\mathcal{F}_t]}$. With $L_{i,t} = E^{\mathbb{P}}[L_{i,t+1}|\mathcal{F}_t]$ and $L_{i,t}$ being completely determined by $\mathcal{F}_t$, we can rewrite the biased expectation of investor i as

$$\begin{aligned} E^{\mathbb{Q}^i}[\mathbf{p_{t+1}} + \mathbf{d_{t+1}}|\mathcal{F}_t] &= E^{\mathbb{P}}[\mathbf{p_{t+1}} + \mathbf{d_{t+1}}|\mathcal{F}_t] + Cov^{\mathbb{P}}\left[\frac{L_{i,t+1}}{L_{i,t}}, \mathbf{p_{t+1}} + \mathbf{d_{t+1}}|\mathcal{F}_t\right] \\ Or\ E_{t,i}[\mathbf{p_{t+1}} + \mathbf{d_{t+1}}] &= E_t[\mathbf{p_{t+1}} + \mathbf{d_{t+1}}] + Cov\left[\frac{L_{i,t+1}}{L_{i,t}}, \mathbf{p_{t+1}} + \mathbf{d_{t+1}}\right] \end{aligned}$$

Linking with Section 2, we can see that $\mathbf{b_{i,t+1}} = Cov\left[\frac{L_{i,t+1}}{L_{i,t}}, \mathbf{p_{t+1}} + \mathbf{d_{t+1}}\right]$. Hence, we can represent the biased part in the return of stock s in Eqs. (7) as

$$\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}} = Cov\left[\sum_i \frac{L_{i,t+1} \cdot \gamma}{L_{i,t} \cdot \gamma_i}; R_{t+1,s}\right] \tag{8}$$

$Cov\left[\sum_i \frac{L_{i,t+1} \cdot \gamma}{L_{i,t} \cdot \gamma_i}; R_{t+1,s}\right]$ represents the covariance risk of stock s's return with the factor $\sum_i \frac{L_{i,t+1} \cdot \gamma}{L_{i,t} \cdot \gamma_i}$. That factor contains information about the aggregate-bias-belief of all investors. With $L_{i,t+1}$ on top and $L_{i,t}$ at the bottom, the factor carries the innovation in aggregate-bias-belief about the stock's payoff at time $t+1$ in comparison with time t. This would mean that the aggregate-bias-belief is also a risk in the market. In brief, a risk-based explanation and a sentiment explanation can be seen as two different representations of the same problem in our model settings.

By showing that risk and sentiment are two sides of the same coin, we extend the discussion of Kozak et al. (2018). Kozak et al. (2018) do not talk about the origin of mispricing. They argue that the mispricing return can have reduced factors form but do not show what are these factors. This paper enlarges these findings by showing that we can represent the mispricing return with just one factor. This risk factor actually represents the aggregate bias belief of investors on the market.

Furthermore, in the next section, we propose methods to estimate mispricing at the stock level and market level using characteristics. This is also a new empirical contribution to sentiment model literature such as Daniel et al. (2001), Kozak et al. (2018).

4. Estimation methods and data

As shown above, we know that we can either view $\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}}$ as a biased return or as a covariance risk, $Cov\left[\sum_i \frac{L_{i,t+1} \cdot \gamma}{L_{i,t} \cdot \gamma_i}; R_{t+1,s}\right]$. It is then natural to discuss how we can estimate this term.

To simplify the notification, we call the term $\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}}$ or $Cov\left[\sum_i \frac{L_{i,t+1} \cdot \gamma}{L_{i,t} \cdot \gamma_i}; R_{t+1,s}\right]$ as $\theta_{t+1,s}$. So, we will define a vector that contains the investors' aggregate-bias-belief about stock returns as $\boldsymbol{\theta}_{t+1} = (\theta_{t+1,1} \dots \theta_{t+1,s} \dots \theta_{t+1,S})^T$. We denote that the net return of stock "s" is $r_{s,t+1}$, so $\mathbf{r_{t+1}} = (r_{1,t+1} \dots r_{s,t+1} \dots r_{S,t+1})^T$. Then, the model in Eq. (7) can be written as follows:

$$E_t[\mathbf{r_{t+1}}] = r_f \cdot \mathbf{1} + \boldsymbol{\beta}_t \cdot E_t\left[r_{t+1,M} - r_f\right] + \boldsymbol{\beta}_t \cdot \left[\mathbf{w_t}^T \cdot \boldsymbol{\theta}_{t+1}\right] - \boldsymbol{\theta}_{t+1} \tag{9}$$

where $\boldsymbol{\beta}_t = (\beta_{1,t+1} \dots \beta_{s,t+1} \dots \beta_{S,t+1})^T$ and $\mathbf{1}$ is a unit vector.

We will then refer $\theta_{t+1,s}$ as both the biased in return and the covariance risk interchangeably. We will use stock characteristics as instrumental variables to estimate $\theta_{t+1,s}$. First, we discuss the logic of using characteristics both under the risk-based view and under the sentiment view. Then we move on to show our estimation methods and discuss the data we use.

4.1. The logic of using characteristics

4.1.1. The risk-based logic

On the risk-based explanation, $Cov\left[-\sum_i \frac{L_{i,t+1} \cdot \gamma}{L_{i,t} \cdot \gamma_i}; R_{t+1,s}\right] = -\theta_{t+1,s}$. The factor $-\sum_i \frac{L_{i,t+1} \cdot \gamma}{L_{i,t} \cdot \gamma_i}$ is an untradeable factor. Therefore, we can use the classic approach of taking the projection of $-\sum_i \frac{L_{i,t+1} \cdot \gamma}{L_{i,t} \cdot \gamma_i}$ on a set of well-managed portfolios to have a mimicking portfolio of this factor. The typical candidates for well-managed portfolios are characteristics-based portfolios. If we denote $\mathbf{Z_t}$ as the characteristics matrix of all stocks at time t, we can recover the projection of $-\sum_i \frac{L_{i,t+1} \cdot \gamma}{L_{i,t} \cdot \gamma_i}$ as:

$$-\sum_i \frac{L_{i,t+1} \cdot \gamma}{L_{i,t} \cdot \gamma_i} = \varepsilon + \mathbf{w^T_{t,factor}} \cdot \mathbf{Z_t^T r_{t+1}} + \epsilon_t \tag{10}$$

On one hand, Eq. (10) can be seen as a projection of the factor $-\sum_i \frac{L_{i,t+1} \cdot \gamma}{L_{i,t} \cdot \gamma_i}$ on a set of characteristics based portfolios $\mathbf{Z_t^T r_{t+1}}$ with a constant weight $\mathbf{w^T_{t,factor}}$. We also require that the $\mathbf{Z_t^T r_{t+1}}$ portfolios span the whole factor space. On the other hand, Eq. (10) can also be seen as a projection of the factor $-\sum_i \frac{L_{i,t+1} \cdot \gamma}{L_{i,t} \cdot \gamma_i}$ on the entire return space $\mathbf{r_{t+1}}$ with time-varying weight $\mathbf{w^T_{t,factor}} \cdot \mathbf{Z_t^T}$. The matrix $\mathbf{Z_t}$ is usually demeaned, so $\mathbf{Z_t^T r_{t+1}}$ are also a vector of cash neutral long-short portfolios.

As $-\sum_i \frac{L_{i,t+1} \cdot \gamma}{L_{i,t} \cdot \gamma_i}$ can be projected as a linear combination of characteristics based portfolio, we can represent $Cov\left[-\sum_i \frac{L_{i,t+1} \cdot \gamma}{L_{i,t} \cdot \gamma_i}; R_{t+1,s}\right] = -\theta_{t+1,s}$ as a linear combination of characteristics based portfolios, hence being also a linear combination of characteristics. Therefore, stocks' characteristics enter naturally into the estimation process of $Cov\left[-\sum_i \frac{L_{i,t+1} \cdot \gamma}{L_{i,t} \cdot \gamma_i}; R_{t+1,s}\right] = -\theta_{t+1,s}$.

However, since $-\sum_i \frac{L_{i,t+1} \cdot \gamma}{L_{i,t} \cdot \gamma_i}$ is latent so we cannot directly observe the factor, nor can we take the projection of this factor on the return space. Kelly et al. (2019) provide the IPCA approach to bypass this challenge. In our context, if we apply the Kelly et al. (2019) approach then we can set: $Cov\left[-\sum_i \frac{L_{i,t+1} \cdot \gamma}{L_{i,t} \cdot \gamma_i}; R_{t+1,s}\right] = -\theta_{t+1,s} = \mathbf{z^T_{t,s}} \cdot \Gamma$, where Γ is a one column matrix, and $\mathbf{z_{t,s}}$ is the characteristics vector of stock "s". So the exposure of one stock's return to the latent factor can be represented as a linear combination of the characteristics of this stock.[11]

The Kelly et al. (2019) approach treats characteristics as covariance (or proxy to covariance). They move one to use the alternating least square approach to estimate Γ and the latent factor. This alternating least square requires a sensible first guess of Γ and can be computationally costly.

[10] We refer to Björk (2009) which gives an excellent summary on measure theory, sigma algebra, and Radon–Nikodym theorem and their applications in finance.

[11] In Kelly et al. (2019), they use β to latent factor instead of covariance. Since β is just the scaled number of covariance by the factor variance, hence the above decomposition of covariance into the linear composition of characteristics shares the same spirit with Kelly et al. (2019).

In our paper, we embrace the Kelly et al. (2019) approach of using a linear combination of characteristics as a proxy for covariance to the latent factor. However, we employ a more simple, and less computational approach than the alternating least square but still deliver powerful results. We discuss the details of our approach in the below part.

4.1.2. The sentiment-based logic

In the risk-based channel, we discussed the logic of using the linear combination of characteristics to represent the covariance risk. In this part, we also discuss that using characteristics is also natural in the sentiment-based channel.

Recall that the mentioned bias of individual "i" over stock "s" is $\frac{b_{i,t+1,s}}{p_{s,t}}$. For simplification, we call that bias $\theta_{i,t+1,s}$. We make a sensible assumption that the bias of individual "i" about future stock return is not random but based on some information. So, we introduce $\mathbf{z_{t,s}}$ as a vector of the characteristics of stock "s". These characteristics can be financial ratios (i.e., book-to-market ratio, size, cash flow, etc.) or can be everything else relating to stock "s". These characteristics serve as *instrumental variables* for the bias in the return.

Subsequently, it is logical to assume that there is a function that maps these characteristics to the biased view of the return. This function will of course be different from individual to individual. This difference will lead to a different biased view among individuals about a stock's return. The bias of individual "i" over stock then is "s" $\theta_{i,t+1,s} = \mathbf{z}^T_{\mathbf{t,s}} \cdot \mathbf{f_{i,t+1}}$, where $\mathbf{f_{i,t+1}}$ is a vector that maps stock's characteristics (or information) to the individual biased view of its return. This vector $\mathbf{f_{i,t+1}}$ is unique for each investor. If we stack $\mathbf{Z_t} = \begin{bmatrix} z^T_{t,1} \\ \cdots\cdots \\ z^T_{t,s} \\ \cdots\cdots \\ z^T_{t,S} \end{bmatrix}$, then we can rewrite $\boldsymbol{\theta}_{t+1}$ in Eq. (9) as $\boldsymbol{\theta}_{t+1} = \mathbf{Z_t} \cdot \sum_i [\frac{\gamma}{\gamma_i} \cdot \mathbf{f_{i,t+1}}]$. The term $\sum_i [\frac{\gamma}{\gamma_i} \cdot \mathbf{f_{i,t+1}}]$ is a weighted view of all investors' views. The weight is decided by the individual risk aversion. We denote this aggregate view as $\mathbf{f_{t+1}}$. Each element in $\mathbf{f_{t+1}}$ reflects the importance of each characteristic to the biased return. Although each investor already formed their biased belief at time t, in our model settings, every individual only knows their bias from the last period (t) when they have the realization of return of this period ex-post ($t+1$). Therefore, $\mathbf{f_{t+1}}$ is ex-post information and unknown at time t. Investors can only estimate $\mathbf{f_{t+1}}$ at time $t+1$. However, if $\mathbf{f_{t+1}}$ is auto-correlated, then $\mathbf{f_t}$ will contain certain information of $\mathbf{f_{t+1}}$ and can be used to predict the next period's return.

By letting $\boldsymbol{\theta}_{t+1} = \mathbf{Z_t} \cdot \mathbf{f_{t+1}}$, we are saying that the bias in return can be represented as a linear combination of stocks characteristics. The sentiment approach thus has the same way to estimate $\boldsymbol{\theta}_{t+1}$ with the risk-based approach above.

4.2. Estimating the bias return

We have shown the logic of using a linear combination of characteristics to estimate the bias return (or the covariance risk to the latent factor). We impose a linear factor structure as $\boldsymbol{\theta}_{t+1} = \mathbf{Z_t} \cdot \mathbf{f_{t+1}}$. This imposition shares the same spirit with Kozak et al. (2018)'s work. Indeed, Kozak et al. (2018) show that the sentiment model does not differentiate from the reduced-factors model. Kozak et al. (2018) prove that the factor model is not only for "rational" asset pricing. Even in the case where the cross-sectional variation of returns is totally driven by the distorted beliefs of some investors, low-dimension factors of the principal components (PCs) of return-covariance can still explain the return variation. Instead of using PCs, we directly use characteristics in our model for two reasons. First, we like to know, in a sentiment model setting, which characteristics significantly contribute to the expected return. Second, with individual stocks as test assets, it would be inefficient to estimate the PCs due to their large dimensions and also because of our unbalanced panel-data structure. The use of characteristics is also in line with previous research. Indeed, Brennan et al. (1998), Daniel and Titman (1997) suggest that one can test the behavioral effect on stock returns by looking at the return variation associated with characteristics, which is orthogonal to other risk factors' returns. Kelly et al. (2019) also use characteristics as instruments for latent factors as we discussed above. Baker and Wurgler (2006) use stock characteristics to build sentiment measures. Avramov et al. (2019), Stambaugh et al. (2012, 2015), Stambaugh and Yuan (2017) use 11 to 12 characteristics or anomalies to construct mispricing measures and mispricing factors.

Our model builds a bridge between the risk-based factors model and the sentiment model by using characteristics. In the sentiment channel, the characteristics enter as initial information input to create the biased belief return. This biased belief return creates mispricing. Then, mispricing creates a variation in the expected return in the cross-section. When chasing back, characteristics create the cross-sectional variation of expected stock returns through the mispricing channel. In the risk-based channel, characteristics serve as a proxy for the exposure to a latent factor.

If we can estimate the term $\mathbf{f_{t+1}}$, we can derive the biased return of both the market and every single stock $\boldsymbol{\theta}_{t+1}$. To estimate $\mathbf{f_{t+1}}$, we can use a least square estimator as follows:

Proposition 3. *The estimated value of* $\mathbf{f_{t+1}}$ *using a least square estimator is*

$$\hat{\mathbf{f}}_{t+1} = (\mathbf{X_t^T X_t})^{-1} \mathbf{X_t^T y_{t+1}} \tag{11}$$

where $\mathbf{X_t} = [\boldsymbol{\beta}_t \cdot \mathbf{w_t^T} - \mathbf{I}] \cdot \mathbf{Z_t}$ *and* $\mathbf{y_{t+1}} = [\mathbf{r_{t+1}} - \mathbf{1} \cdot rf] - \boldsymbol{\beta}_t \cdot \left[r_{t+1,M} - r_f\right]$.

Proof. Appendix C.

Proposition 3 gives the value of $\mathbf{f_{t+1}}$ in every period. As we can perform this cross-sectional regression for example at every point in time, it shares some common points with the Fama and MacBeth (1973) regressions. However, our model is different, since it introduces the term $[\boldsymbol{\beta}_t \cdot \mathbf{w_t^T} - \mathbf{I}]$ before $\mathbf{Z_t}$. This term will serve as a weighting matrix. This property is welcomed because it will control the over-influence of micro-cap stocks.

Indeed, a common Fama and MacBeth (1973) regression approach applied to all individual stocks will give the same weight to all stocks no matter their size. It is well known that small stocks comprise a large part of the stock universe but have just a small total capitalization. Therefore, when an equal-weighting scheme is applied in the least square estimator, the effect of small-cap stocks dominates. Studies such as Green et al. (2017) and Hou et al. (2020) use the weight least square (WLS) instead of the OLS estimator to alleviate the small-cap effect. They use market capitalization as weight. In our setting with a weighting matrix $[\boldsymbol{\beta}_t \cdot \mathbf{w_t^T} - \mathbf{I}]$, our model accounts not only for the market capitalization but also for the covariance risk (β) of an individual stock's return to the market's return.

Because we cannot directly observe the true $\boldsymbol{\beta}_t$ we have to estimate it from a time series. Hence, $\hat{\boldsymbol{\beta}}_t = \boldsymbol{\beta}_t + \boldsymbol{\vartheta}_t$. Therefore, in Eq. (11), we can have an error in variable ($\boldsymbol{\vartheta}_t$) problem. Instead of $\mathbf{X_t} = [\boldsymbol{\beta}_t \cdot \mathbf{w_t^T} - \mathbf{I}] \cdot \mathbf{Z_t}$ and $\mathbf{y_{t+1}} = [\mathbf{r_{t+1}} - \mathbf{1} \cdot rf] - \boldsymbol{\beta}_t \cdot \left[r_{t+1,M} - r_f\right]$, the real terms that we put into the regression are $\hat{\mathbf{X}}_t = [\hat{\boldsymbol{\beta}}_t \cdot \mathbf{w_t^T} - \mathbf{I}] \cdot \mathbf{Z_t}$ and $\hat{\mathbf{y}_{t+1}} = [\mathbf{r_{t+1}} - \mathbf{1} \cdot rf] - \hat{\boldsymbol{\beta}}_t \cdot \left[r_{t+1,M} - r_f\right]$. In brief, due to this error in variable (EIV) problem, our OLS estimate can be biased (see Wooldridge, 2010, 2015).

Jegadeesh, Noh, Pukthuanthong, Roll, and Wang (2019) propose instrumental methods to alleviate this bias. Chordia, Goyal, and Shanken (2015) derive the asymptotic bias due to the EIV problem and eliminate it analytically. In our settings, EIV is not a serious problem. Indeed, we estimate $\mathbf{f_{t+1}}$ by using individual stocks. When the number of stocks increases, $\mathbf{w_t}$ is very small and tends to approach zero. Therefore, it is plausible for us to believe that $\boldsymbol{\vartheta}_t \cdot \mathbf{w_t^T}$ is small and near zero; hence, $\mathbf{w_t} \cdot \boldsymbol{\vartheta}_t^T \cdot \boldsymbol{\vartheta}_t \cdot \mathbf{w_t^T} \to 0$ when the number of stocks is growing. It is also plausible to say that the EIV ($\boldsymbol{\vartheta}_t$) in the estimation process of $\hat{\boldsymbol{\beta}}_t$ is unrelated to (i) the true stock beta $\boldsymbol{\beta}_t$, (ii) every stock characteristics, and (iii) and the future return $\mathbf{r_{t+1}}$; hence, $\mathbf{Z_t^T}\cdot\boldsymbol{\vartheta}_t = 0$, $\boldsymbol{\beta}_t^T\cdot\boldsymbol{\vartheta}_t = 0$, $\mathbf{r_{t+1}^T}\cdot\boldsymbol{\vartheta}_t = 0$. Taking into account these facts, $\hat{\mathbf{f}}_{t+1} = (\hat{\mathbf{X}}_t^T\hat{\mathbf{X}}_t)^{-1}\hat{\mathbf{X}}_t^T\hat{\mathbf{y}}_{t+1} \rightarrow (\mathbf{X_t^T X_t})^{-1}\mathbf{X_t^T y_{t+1}}$ when the number of stocks in the cross-section grows. So, in our settings, the EIV problem alleviates itself.

Doing a cross-sectional regression every month will give us a time series of each element in $\hat{\mathbf{f}}_{t+1}$. In the same spirit as Fama and MacBeth (1973), if these time series are stationary, we can take the average of $\hat{\mathbf{f}}_{t+1}$ over time and test whether on average, each element in $\hat{\mathbf{f}}_{t+1}$ is different from zero. Logically, we also expect that each element in $\hat{\mathbf{f}}_{t+1}$ is autocorrelated; therefore, we use Newey and West (1987)'s methods to calculate the standard error and *t*-stats.

Following the discussion in Section 2, if all investors are unbiased or if the arbitrage force is efficient enough, then Eq. (9) will converge to a CAPM model. This would mean each element in $\hat{\mathbf{f}}_{t+1}$ would be zero or not significantly different from zero. In the empirical result section, we show that this is not the case.

Another interesting feature of the vector $\hat{\mathbf{f}}_{t+1}$ is that every element is a return from a long-short self-financing (zero investment) portfolio.[12] Therefore, we can interpret each element in $\hat{\mathbf{f}}_{t+1}$ as a type of return premium to the relevant stock's characteristic. This feature is also welcomed under the risk-based channel. Since each element in $\hat{\mathbf{f}}_{t+1}$ is a well-managed portfolio hence we can represent the covariance risk to the latent factor of every stock, $\boldsymbol{\theta}_{t+1} = \mathbf{Z_t^T f_{t+1}}$, as a linear combination of well-diversified cash neutral portfolios $\mathbf{f_{t+1}}$ with their characteristics, $\mathbf{Z_t^T}$, serving as weights.

4.3. Data

We use CRSP, COMPUSTAT, and IBES databases to derive the characteristics and return of every common stock (share code: 10, 11, 12) on the NYSE, AMEX, and NASDAQ exchanges. The data frequency is monthly. We also use the risk-free rate and Fama and French (2015) five factors and momentum factor return from the Kenneth French data library.[13] We use the (Stambaugh & Yuan, 2017) mispricing factors from Stambaugh's website,[14] and the Baker and Wurgler (2006) sentiment measure from Wurgler's website.[15] We follow the procedure of Green et al. (2017) to construct 101 stocks' characteristics, and their β, then winsorize them at the 1% level. We also adjust the delisting return.[16] We estimate $\mathbf{f_t}$ for every month and compute the biased return for every stock and for the market. The list of characteristics with detailed information on calculation and publication is in Appendix D.

Different variables will have different units and hence different variances, which can impair the estimates. Therefore, every month we cross-sectionally standardized 101 stocks' characteristics to have zero mean and a unit standard deviation.[17] To deal with missing observations, we use a traditional approach, allocating missing values to be equal to their cross-sectional mean.

Our sample is from 1980 to 2018 with 2 093 322 firm-month observations. There are total 18 706 firms that enter our sample. On average, we have around 4473 firms in our sample per month. The minimum number of firms in a month is 2999, while the maximum is 6594.

By using a large number of characteristics, we share the same spirit of employing a massive number of characteristics in asset pricing models as is present in such going literature as Chen, Pelger, and Zhu (2019), Feng, He, and Polson (2018), Green et al. (2017), Gu, Kelly, and Xiu (2020), Kelly et al. (2019), Kozak (2019); etc. However, one could still raise concerns about the large number of characteristics in our studies. There can be two major concerns: one is about econometric techniques, and the other is about economic meaning.

For the econometric concern, using a large number of characteristics can cause multi-collinearity. This can make the estimates from a single cross-sectional regression volatile, which in turn can impair the model's prediction power. Although we cannot avoid multi-collinearity in our estimation methods, multicollinearity does not cause biased estimates (see Wooldridge, 2015). Therefore, in the second step, we collect estimates from month to month and take the average of these estimates. We draw inferences based on these average estimates. By doing so, we get rid of the noise in the estimates and end up with a quite precise estimation of the coefficients (see Fama & MacBeth, 1973). In the below part, we show that the moving average of $\hat{\mathbf{f}}_t$ can be used to build a strong signal to predict the expected return. Therefore, multi-collinearity is not a major concern for the model's prediction power in our paper.

One can also worry about the use of 101 characteristics as a "fishing exercise" to ensure a positive result about mispricing. First, we argue that our theoretical model and empirical estimation methods do not impose any mispricing conclusion in advance. In our theoretical model, stock can be underpriced, overpriced, or correctly priced. The data will speak for themselves about the mispricing level. Second, the "fishing exercise" concern usually arises when one randomly finds a result without any strong theoretical basis. We argue that this is not the case here, since our empirical settings are based on our theoretical model. The theory is based on two streams of literature: sentiment, and risk-based models. Therefore, the empirical exercises do not randomly come from the blue.

The characteristics in our study come from accounting measures, analyst information, price, volume, etc. One cannot deny that investors use these characteristics as information to form their judgments about a stock's price. On the sentiment-based channel, while these stock characteristics are public, inferring a stock's future payoff from this information is private. Investors can thus make private mistakes in the process. On the risk-based channel, using a large number of characteristics will create a large number of characteristics-based portfolios. These large number of portfolios will ensure that they span all the factor space, which contains the set of factors that can explain return, and the latent factor in our model. Hence, using these characteristics in our model is not a random choice but has a strong reasonable ground.

After getting the estimates $\hat{\mathbf{f}}_{t+1}$ in every month, we can calculate exactly the aggregate-bias-belief about each stock return, and the aggregate-bias-belief about the market return. We consider the market return in each month as the weighted average of the returns of all stocks that enter our sample that month. The next section will discuss empirical results from our model.

5. Empirical results

In this part, we will apply our model in different ways. First, we estimate $\mathbf{f_{t+1}}$ to clarify the effect of each characteristic on the expected return. Second, we investigate the properties of the aggregate-bias-belief return and its predicting power on the expected return. Third, we confirm Corollaries 1 and 2 and Proposition 2 by data. Fourth, using our model, we investigate the predicting power of some well-known characteristics (or anomalies) in overpriced and underpriced stocks. Fifth, we test whether the aggregate-bias-belief return can be explained by other mispricing factors.

[12] To see that, recall: $\hat{\mathbf{f}}_{t+1} = (\mathbf{X_t^T X_t})^{-1}\mathbf{X_t^T y_{t+1}}$. We have $\mathbf{X_t^T y_{t+1}} = \mathbf{Z_t^T}\cdot\mathbf{w_t}\cdot\boldsymbol{\beta}_t^T\cdot\mathbf{y_{t+1}} - \mathbf{Z_t^T}\cdot\mathbf{y_{t+1}}$. Because we cross-sectionally standardized each characteristic, for any vector of return, $(\mathbf{a})$, $\mathbf{Z_t^T}\cdot\mathbf{a}$ will give us a vector whose each element is a return from a long-short self-financing portfolio. $\mathbf{w_t}\cdot\boldsymbol{\beta}_t^T\cdot\mathbf{y_{t+1}}$ and $\mathbf{y_{t+1}}$ are both return vectors. Hence $\mathbf{X_t^T}\cdot\mathbf{y_{t+1}} = \mathbf{Z_t^T}\cdot(\mathbf{w_t}\cdot\boldsymbol{\beta}_t^T\cdot\mathbf{y_{t+1}} - \mathbf{y_{t+1}})$ is a vector whose each element is a return from a long-short zero investment portfolio. We know that any linear combination of different long-short zero investment portfolios will give a long-short zero investment portfolio. Hence, $\hat{\mathbf{f}}_{t+1} = (\mathbf{X_t^T X_t})^{-1}\mathbf{X_t^T y_{t+1}}$ is a vector whose each element is a return from a long-short zero investment portfolio.

[13] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[14] http://finance.wharton.upenn.edu/~stambaug/

[15] http://people.stern.nyu.edu/jwurgler/

[16] For more information, please consult: https://sites.google.com/site/jeremiahrgreenacctg/home

[17] Note that β is not standardized.

Table 1
Mean of $\hat{\mathbf{f}}_{t+1}$ with t-stats for each elements in $\hat{\mathbf{f}}_{t+1}$. We estimate $\hat{\mathbf{f}}_{t+1}$ for every month following Eq. (11). Our sample is from 1980 to 2018. $\hat{\mathbf{f}}_{t+1}$ is a 101-element vector corresponding to 101 characteristics. We take the average of each element in $\hat{\mathbf{f}}_{t+1}$ over time. We use the Newey and West (1987) standard error with up to seven lags. We only report the characteristics that have an absolute t-stats greater than 2. Each element in $\hat{\mathbf{f}}_{t+1}$ is a return from a net long-short portfolio and can be seen as a monthly return premium for a characteristic.

Variable	Mean f %	t-stats	Variable	Mean f %	t-stats
mom1 m	0.76**	9.13	ear	−0.11**	−7.52
mve	0.73**	4.45	std_dolvol	0.11**	3.09
turn	0.57**	11.13	bm	−0.11**	−3.61
retvol	0.53**	6.20	nincr	−0.11**	−6.51
ill	−0.41**	−8.71	IPO	0.09**	3.95
std_turn	−0.38**	−7.60	roaq	−0.09**	−2.16
rd_mve	−0.37**	−7.67	gma	−0.09*	−2.08
indmom	−0.32**	−6.60	chtx	−0.08**	−5.29
nanalyst	−0.32**	−6.15	disp	0.08**	3.84
betasq	0.30*	2.20	cfp	−0.08**	−2.70
zerotrade	0.27**	7.16	roeq	−0.08*	−2.26
dolvol	−0.23**	−2.72	convind	0.07**	4.58
stdacc	0.22*	2.23	ms	−0.07**	−2.61
mom12 m	−0.21**	−3.24	acc	0.06*	2.11
cash	−0.20**	−4.89	chnanalyst	0.06**	3.78
baspread	−0.18*	−2.37	herf	0.05**	3.02
idiovol	0.16**	3.02	grcapx	0.05**	3.27
chfeps	−0.16**	−6.70	sin	−0.05**	−3.32
ep	−0.16**	−4.32	rsup	−0.05*	−2.06
agr	0.14**	3.61	chcsho	0.05*	2.38
currat	0.13*	2.12	divi	0.05**	3.36
sue	−0.13**	−5.26	tb	−0.04**	−2.66
sfe	−0.12*	−2.00	chatoia	−0.04*	−2.22

*$p \leq 0.05$.
**$p \leq 0.01$.

5.1. The estimations of $\mathbf{f}_{t+1}$

Denote $\hat{\mathbf{f}}_{t+1}$ as the estimation of $\mathbf{f}_{t+1}$. Each element in $\hat{\mathbf{f}}_{t+1}$ represents the impact of this characteristic on the stock's return. There are 101 elements in $\hat{\mathbf{f}}_{t+1}$. Therefore, following the estimation methods in Section 4, we estimate $\hat{\mathbf{f}}_{t+1}$ for every month and come up with 101 time series. First, we use augmented Dickey–Fuller (ADF) to test if these time series are stationary.[18] We then confirm that these time series are all stationary. Therefore, it is likely that the mean of each element in $\hat{\mathbf{f}}_{t+1}$ does not shift over time. The table below reports the average value of each element in $\hat{\mathbf{f}}_{t+1}$. We also compute the t-stats. Taking into account the auto-correlation problem, we use the Newey and West (1987) standard error with up to 7 lags. We only report the ones that have an absolute t-stats greater than 2.

Note that according to Eq. (9), each element in $\hat{\mathbf{f}}_{t+1}$ has a *positive* impact on the current aggregate-bias-belief return but a *negative* impact on the real expected return. For example, the *mve* (Size) characteristic has a loading of 0.76%. This means, ceteris paribus, that an increase of one cross-sectional standard deviation of *mve* will decrease the monthly expected return by 0.76%. As another example, the *ill* characteristics ((Amihud, 2002) illiquidity measure) has a loading of −0.41%. This means that an increase of one cross-sectional standard deviation of *ill* will increase the monthly expected return by 0.41%.

Table 1 arranges the characteristics' loading in decreasing order by absolute value. Then, we can see that the size characteristic *mve* (size), liquidity characteristics (*ill, turn, basspread*, etc.), volatility characteristics (*retvol, idiovol*), and momentum (*mom1, indmom, mom12m*) are the ones that have the most important impact on expected return. After that are the characteristics related to cash flow, earnings, accrual, and investment (*rd_mve, stdacc, currat, ear, roaq, cfp, grcapx*, etc.). Only 47/101 have an absolute t-stats greater than 2. If we follow Harvey et al. (2016)'s criteria, 30/101 characteristics have an absolute t-stats greater than 3, while 18/101 have an absolute t-stats greater than 4.

Note that each loading in $\hat{\mathbf{f}}_{t+1}$ can be seen as a return from a long-short zero investment portfolio offering a monthly return premium for a characteristic. Therefore, we can also ask for the economic significance apart from the statistical significance of these loadings. Most of the characteristics offer a very small return premium. Only 11/101 have an annual return premium greater than 3%, of which 4/101 (*mve, mom1, turn, retvol*) have an annual return premium greater than 5%. In brief, this means that only a few characteristics have both statistically and economically significant impacts on the expected return. In our context, this would mean that only a few important characteristics impact the investor's biased belief about future return (or the exposure to the latent factor). These characteristics are related to liquidity, momentum, size, and volatility.

Even though these characteristics are put in a special context in the sentiment model, the results corroborate other research such as Kozak et al. (2018) which theoretically show that the stochastic discount factor (SDF) can only be presented by a few dominant factors if the absence of near-arbitrage condition holds. Our results also corroborate the empirical research of Green et al. (2017), Harvey et al. (2016), Hou et al. (2015, 2020), Kelly et al. (2019), who argue that only a handful of factors, characteristics, or anomalies are statistically and economically important. For the risk-based channel, this also means that only a handful of characteristics can present the information in the latent factor of our model.[19]

5.2. The aggregate-bias-belief of Stock's return and Market return

Knowing $\hat{\mathbf{f}}_{t+1}$, we estimate the aggregate-bias-belief of the stocks' returns ($\boldsymbol{\theta}_{t+1}$) for every month by letting $\boldsymbol{\theta}_{t+1} = \mathbf{Z_t} \cdot \hat{\mathbf{f}}_{t+1}$. We also estimate the aggregate-bias-belief of market return as $\mathbf{w^T} \cdot \boldsymbol{\theta}_{t+1}$. For convenience, we denote $RET.bias.M_{t+1} = \mathbf{w^T} \cdot \boldsymbol{\theta}_{t+1}$. The estimates are done every month. We then take a summary statistic of the cross-sectional values of $\boldsymbol{\theta}_{t+1}$, $\boldsymbol{\beta}_t$, $\mathbf{r_{t+1}}$ every month. Then, we take the time-series average of these statistics.[20] Table 2 gives the results. The results in column β are decimal, while other columns have percentage values, except for the skew and the kurtosis rows. For the $RET.bias.M$ column, "n" is the number of months; in other columns, "n" is the average number of stocks every month.

We can see, as usual, a huge variation (sd = 16.94%) in the cross-sectional distribution of returns and a high level of fat-tails (kurtosis = 94.86). However, when we move to the cross-section distribution of aggregate-bias-belief return, we can see a smaller variation (sd = 7.23%), fewer fat-tails (kurtosis = 4.17), a fair degree of symmetric (skew = 0.19), and a mean that is very small and close to zero. This is an interesting finding. It means that on a market-wide level, around *half* of the stocks have a positive aggregate-bias-belief return, and the other *half*, a negative one. These two biases cancel each other out. Therefore, when we take a cross-sectional average of $\theta_{t,s}$, it is statistically indifferent from zero.[21]

[18] There are 101 tests of stationarity. We use the ADF test for up to 7 lags. The details are available upon request.

[19] One can ask the question why liquidity, momentum, size, and volatility work well than the other characteristics. On the sentiment approach, we can see that these characteristics are easier to get at high frequency than other accounting variables. Therefore, there is a chance that investors put a lot of weight on these characteristics in their trading decision and act quickly. These quick decisions are not always optimal and can create systematic mispricing on the market level.

[20] We also have summary statistics for 101 characteristics. Because of the standardizing process, the characteristics have a mean close to zero and a standard deviation close to 1. Detailed statistics are available upon request.

[21] Even if we take a simple approach by taking the mean divided by the "se", we will come up with a very small t-stats. This t-stats does not account for the cross-sectional dependency, nor the autocorrelation. If we take into account these two problems, we will likely have an even smaller t-stats.

Table 2
Summary statistics of cross-sectional return (r), β, aggregate-bias-belief return (θ), and aggregate-bias-belief market return $RET.bias.M$ from 1980 to 2018. We estimate the aggregate-bias-belief of stocks' returns (θ_{t+1}) every month by letting: $\theta_{t+1} = \mathbf{Z_t} \cdot \hat{\mathbf{f}}_{t+1}$. We also estimate the aggregate-bias-belief of market return as $RET.bias.M_{t+1} = \mathbf{w^T} \cdot \theta_{t+1}$. First, we take summary statistics of the cross-sectional values every month. Then, we take the time-series average of these statistics. The results in column β are decimal, while the other columns have percentage values, except for the skew and kurtosis rows. For the $RET.bias.M$ column, "n" is the number of months; in other columns, "n" is the average number of stocks every month.

	β	θ(%)	r(%)	RET.bias.M (%)
n	4472.70	4472.70	4472.70	468
Mean	1.09	0.00	1.12	−0.14
Sd	0.65	7.23	16.94	3.75
Median	1.02	−0.11	0.15	−0.53
Min	−0.16	−37.19	−89.63	−11.62
Max	2.98	40.17	321.56	27.46
Skew	0.54	0.19	3.93	2.01
Kurtosis	0.18	4.17	94.86	11.62
Standard error (se)	0.01	0.11	0.25	0.17
Q0.25	0.63	−4.10	−6.34	−2.22
Q0.5	1.02	−0.11	0.15	−0.53
Q0.75	1.48	4.00	6.97	1.40

Unlike the cross-sectional mean of $\theta_{t,s}$, "$RET.bias.M$" is a cross-sectional weighted mean based on market capitalization. We also have a very small aggregate-bias-belief return on the market (−0.14% on a monthly average). The distribution of "$RET.bias.M$" is slightly right-skewed (skew = 2.01) and has quite extreme values (kurtosis = 11.62). However, with a small mean and large standard error (se), we cannot statistically differentiate "$RET.bias.M$" from zero. This means that sometimes markets are upward or downward-biased. However, it is likely that these biases were corrected over time.

In brief, these above analyses show that the aggregate-bias-belief about an individual stock's return can be huge. However, on the market-wide level, the aggregate-bias-belief about the market's return is not so severe. This can be for two reasons: (i) there is a symmetric cross-sectional distribution of aggregate-bias-belief about stocks' returns around zero, and (i) the aggregate-bias-belief about the market's return has a mean that reverts to zero over time.[22]

To get a better view of "$RET.bias.M$", we plot below its time series. We also plot the development of the market value, whose initial value is set to be 1 at the beginning of 1980.[23]

From Fig. 2, we can see that "$RET.bias.M$" is a mean reverting process. An interesting thing to notice is that during times of turmoil such as when the dot-com bubble burst (2001–03) or during the financial crisis (2008), the aggregate-bias-belief about the market return shoots up and is very volatile. This means that at this particular time, market participants are overly optimistic and do not anticipate crises.

5.3. Return prediction power of the aggregate-bias-belief return

In Eq. (9), we can see that $\boldsymbol{\theta}_{t+1} = \mathbf{Z} \cdot \mathbf{f_{t+1}}$ has a negative impact on the expected return. Unfortunately, we can only know $\mathbf{f_{t+1}}$ and $\boldsymbol{\theta}_{t+1}$ ex-post at time $t+1$ but not ex-ante at time t. However, if $\mathbf{f_t}$ contains some information about $\mathbf{f_{t+1}}$, we can use $\mathbf{f_t}$ as a return predictor. Or, as we know that each element in $\mathbf{f_{t+1}}$ is stationary, we can use past information to predict $\mathbf{f_{t+1}}$.

Motivated by this idea, we investigate three predictors: (i) the lag value of $\theta_{t+1,s}$, or $\theta_{t,s}$; (ii) a predictor called $\theta^P_{t+1,s}$, where $\theta^P_{t+1,s} = \mathbf{z^T_{t,s}} \cdot \mathbf{f_t}$; (iii) $\theta^{MA}_{t+1,s} = \mathbf{z^T_{t,s}} \cdot \mathbf{f^{MA}_t}$, where $\mathbf{f^{MA}_t}$ is the moving average of $\mathbf{f}$ from month $t-24$ to month t. Because we do not know $\mathbf{f_{t+1}}$ at time t, we use $\mathbf{f_t}$ and $\mathbf{f^{MA}_t}$ as proxies for $\mathbf{f_{t+1}}$ to construct the predictors $\theta^P_{t+1,s}$ and $\theta^{MA}_{t+1,s}$. Table 3 represents the results of the portfolio sorting procedure. In details, in every month t, we rank stocks from low to high values of either $\theta^P_{t+1,s}$, $\theta^{MA}_{t+1,s}$ or $\theta_{t,s}$. Then, we allocate stocks into 10 equal quantiles in every month t. Every quantile is a portfolio. We track each portfolio's return in the next month $t+1$. To allocate the stock's weight within a portfolio, we use both equal-weighted (EW) and value-weighted (VW) methods. We compute the high minus low portfolio (H–L). After that, we take the time-series average of each portfolio. For the H–L portfolio, we compute the Sharpe ratio, CAPM α, and Fama and French (2015) 5 factors model plus momentum factor (FF5.MOM) α. We use (Newey & West, 1987) with up to 7 lags to calculate the standard error for t-stats.[24]

We also apply the same procedure with $\theta_{t+1,s}$ as a sorting variable for the return at time $t+1$. Of course, $\theta_{t+1,s}$ is only available at time $t+1$; hence, it is not a predictor for $\mathbf{r_{t+1}}$. We use $\theta_{t+1,s}$ to see how large the difference is between the realized return of top optimistic stocks and the realized return of top pessimistic stocks. The return in Table 3 is the excess return from the risk-free rate. All portfolios' returns and α are in percentage terms, the rest are in decimal values.

Recall that we sort portfolios from low to high values of $\theta_{t,s}$, $\theta^{MA}_{t+1,s}$, $\theta^P_{t+1,s}$, and $\theta_{t+1,s}$. So, we can say that portfolio 1 (L) has the most biased view on the pessimistic side, and portfolio 10 (H) has the most biased view on the optimistic side. Or portfolio 1 (L) is the most underpriced portfolio, while portfolio 10 (H) is the most overpriced portfolio. Following Eq. (9), we can expect that the H–L spread portfolio will offer a negative return. The empirical results show that the above conclusion is completely true with all these sorting variables.

Clearly, we can see that $\theta^P_{t+1,s}$ offers a return premium. The H–L spread portfolio monthly return is −0.91% for EW and −0.88% for VW. These returns and the H–L CAPM α are all statistically different from zero, but not the H–L FF5.MOM α. In detail, the t-stats of FF5.MOM α is just −1.03 for EW methods and −0.44 for VW methods. These 10 portfolios' returns sorting on $\theta^P_{t+1,s}$ decreases monotonically from L to H.

$\theta_{t,s}$ has some predicting power over return at time $t+1$. With EW methods, the next month's return decreases monotonically from portfolio L to H. The H–L portfolio sorting on $\theta_{t,s}$ offers a significant monthly return of −0.42% with a t-stats of −1.94. The CAPM α (−0.46%) is also significantly different from zero (t-stats: −2.20). However, the FF5.MOM α reduces almost by almost four times (−0.09%) and is not significantly different from zero anymore (t-stats: −0.24). With VW methods, in an absolute sense, the H–L portfolio offers a decent return (−0.52%). The CAPM α is almost the same (−0.56%) and significantly different from zero (t-stats: −2.65). But FF5.MOM α (−0.01%) is small and not significantly different from zero, with the small t-stats (−0.04).

Although these two predictors $\theta_{t,s}$ and $\theta^P_{t,s}$ offer a decent premium through the H–L returns, it is likely that these H–L returns are in line with factors in the FF5.MOM model.

With $\theta_{t+1,s}$, the next month's return decreases monotonically from portfolio 1 to 10. Hence, we can see a huge gap in realized return between the most optimistic stocks (overvalued) and the most pessimistic stocks (undervalued), reflected through the H–L returns (−8.51% for EW, −6.64% for VW) and through its FF5.MOM α (−8.74% for EW, −6.50% for VW). These returns and α are significantly different from zero. This would mean the return gap between the most over-valued and the most under-valued stocks is huge and cannot be explained by the FF5 and momentum factors. This is an empirical confirmation of our model in Eq. (9) and the conclusion of Corollary 2.

[22] We also do an ADF stationary test with up to 7 lags on "$RET.bias.M$". The test rejects the null hypothesis of a unit root. So, it is likely that "$RET.bias.M$" has a stable mean over-time.

[23] As discussed above, we take a value-weighted average of all the stock returns every month as the market return. We only consider stocks that show up in our sample.

[24] Because we need 24 observations to create the first $\theta^{MA}_{t+1,s}$, the sample period when we use $\theta^{MA}_{t+1,s}$ as a sorting variable is from 1982 to 2018.

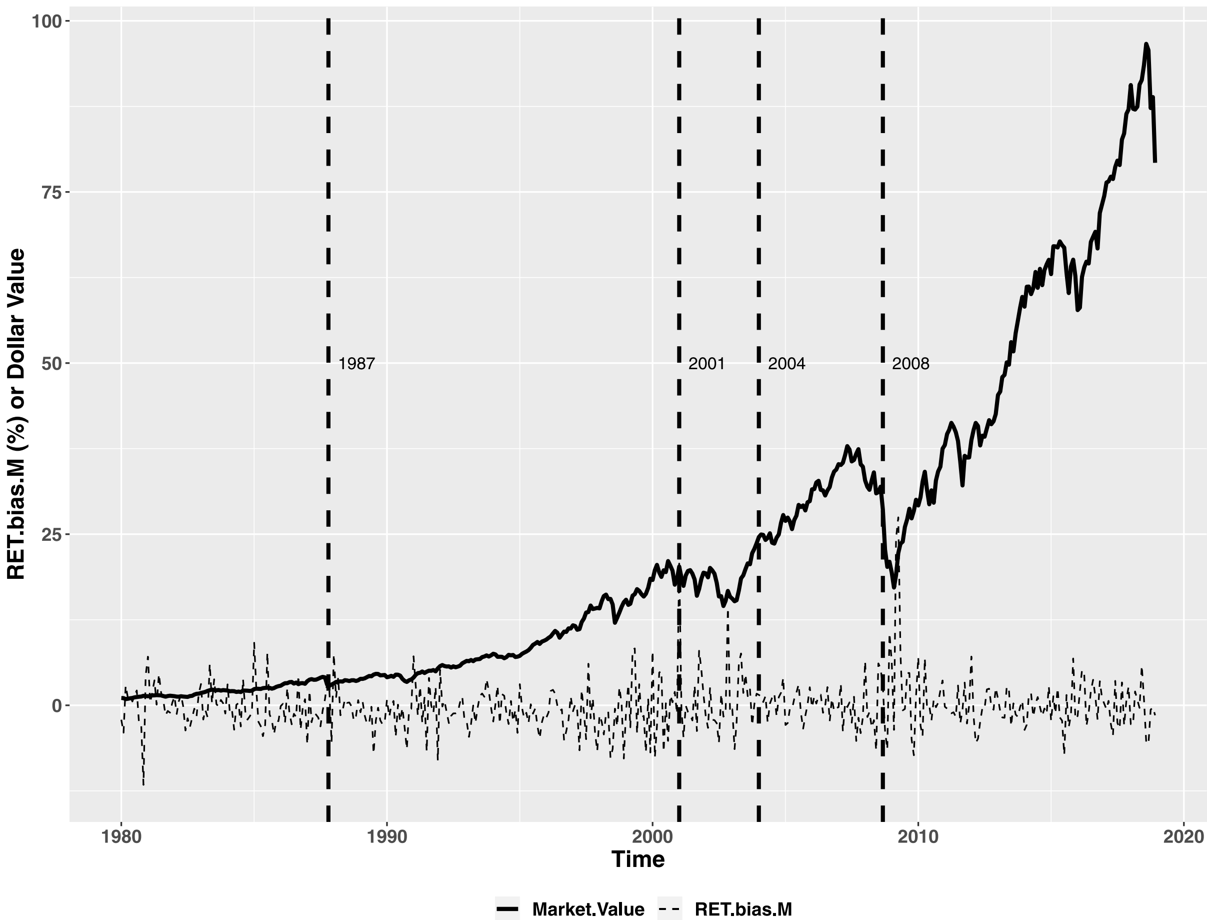

Fig. 2. This figure plots the time series of "$RET.bias.M$" (as a percentage) and the development of market value (in dollar terms). We set the initial market value to be 1 at the beginning of 1980. Our sample is from 1980 to 2018.

Table 3
Univariate portfolio sorting 1980–2018. In every month t, we rank stocks having from low to high values of either $\theta^P_{t+1,s}$, $\theta^{MA}_{t+1,s}$, or $\theta_{t,s}$. Then, we allocate stocks into 10 equal quantiles in every month t. Every quantile is a portfolio. We track each portfolio's return in the next month $t+1$. To allocate the stock's weight within a portfolio, we use both equal-weighted (EW) and value-weighted (VW) methods. We compute the high minus low portfolio (H–L). After that, we take the time-series average of each portfolio. For the H–L portfolio, we compute the Sharpe ratio, CAPM α, and Fama and French (2015) 5 factors model plus momentum factor (FF5.MOM) α. We use Newey and West (1987) with up to 7 lags to calculate the standard error for t-stats. We also apply the same procedure with $\theta_{t+1,s}$ as a sorting variable for the return in time $t+1$. Our sample is from 1980 to 2018 when we use $\theta_{t+1,s}$, $\theta_{t,s}$, $\theta^P_{t+1,s}$ as sorting variables. When we use $\theta^{MA}_{t+1,s}$ as the sorting variable, the sample is from 1982 to 2018. All the portfolios' returns and α are in percentage terms, the rest are in decimal values.

EW Portfolios	$\theta_{t,s}$	$\theta^P_{t+1,s}$	$\theta^{MA}_{t+1,s}$	$\theta_{t+1,s}$	VW Portfolios	$\theta_{t,s}$	$\theta^P_{t+1,s}$	$\theta^{MA}_{t+1,s}$	$\theta_{t+1,s}$
L	0.61	0.86	1.06	4.96	L	0.54	0.71	0.70	3.70
2	0.70	0.87	0.83	2.50	2	0.68	0.81	0.63	2.20
3	0.70	0.68	0.74	1.81	3	0.69	0.65	0.58	1.56
4	0.67	0.71	0.63	1.23	4	0.65	0.68	0.52	1.11
5	0.64	0.63	0.62	0.74	5	0.64	0.66	0.45	0.70
6	0.60	0.55	0.49	0.26	6	0.61	0.57	0.38	0.25
7	0.54	0.49	0.50	−0.18	7	0.56	0.52	0.35	−0.15
8	0.43	0.35	0.32	−0.75	8	0.44	0.37	0.18	−0.69
9	0.34	0.28	0.12	−1.62	9	0.30	0.26	0.04	−1.47
H	0.19	−0.05	−0.50	−3.55	H	0.02	−0.17	−0.52	−2.94
H–L	−0.42	−0.91	−1.56	−8.51	H–L	−0.52	−0.88	−1.22	−6.64
H–L t-stats	−1.94	−4.26	−6.88	−17.34	H–L t-stats	−2.37	−3.93	−4.64	−14.72
H–L Sharpe ratio	−0.07	−0.16	−0.33	−1.37	H–L Sharpe ratio	−0.09	−0.16	−0.24	−1.18
CAPM α	−0.46	−0.94	−1.67	−8.48	CAPM α	−0.56	−0.92	−1.41	−6.64
CAPM α t-stats	−2.20	−4.46	−7.65	−16.96	CAPM α t-stats	−2.65	−4.16	−5.66	−14.03
FF5.MOM α	−0.09	−0.36	−1.12	−8.74	FF5.MOM α	−0.01	−0.13	−0.70	−6.50
t-stats	−0.24	−1.03	−4.05	−13.15	t-stats	−0.04	−0.44	−2.36	−10.71

Unfortunately, recall that one cannot fully exploit this gap by arbitrage because $\theta_{t+1,s}$ is ex-post but not ex-ante information. $\theta_{t,s}$ and $\theta^{P}_{t+1,s}$ can be used as decent proxies for $\theta_{t+1,s}$. But they are not super-strong proxies, since they offer a premium on H–L portfolios but not a significant FF5.MOM α.

Finding a better ex-ante proxy for $\theta_{t+1,s}$ will help us a lot in constructing a better mispricing measure, or factors that can better explain returns. In that sense, $\theta^{MA}_{t+1,s}$ is a good candidate. Indeed, the H–L monthly premium return generated by sorting on $\theta^{MA}_{t+1,s}$ is big (−1.56% in the *EW* scheme and −1.22% in the *VW* scheme). The FF5.MOM α is also economically significant: −1.12% in the *EW* scheme and −0.70% in the *VW* scheme. All of these returns and α are significantly different from zero. Therefore, we can say that the spread return predicted by $\theta^{MA}_{t+1,s}$ cannot be fully explained by other common factors in the FF5 factors, nor by the momentum factor. In Section 5.6, we also show that Baker and Wurgler (2006) sentiment measure and Stambaugh and Yuan (2017) mispricing factors cannot fully explain this spread return. In the next section, we investigate whether the model's theoretical conclusions are confirmed by the data.

5.4. Testing Corollaries 1, 2, and Proposition 2

Recall that in Section 2, we draw some conclusions from Corollaries 1, 2 and Proposition 2. In brief, Corollary 1 states that when we control for the aggregate-bias-belief about future return, then we should expect that high-β stocks predict higher return than low-β stocks. Corollary 2 states that when we control for stock β, the expected return gap between the most optimistically biased and the most pessimistically biased stock is negative. Proposition 2 states that a positive aggregate-bias-belief about future return will shoot up the recent return and vice versa with a negative aggregate-bias-belief about the future return.

In the previous part, we have some empirical evidence (Table 3) that suggests Corollary 2 holds. However, we do not control for β in the previous part. In this section, we show that even when we control for β, Corollary 2 still holds. We also show that Proposition 2 holds. For Corollary 1, we find a mixed result.

5.4.1. Testing Corollary 1

To test this corollary, we perform a double sorting procedure (5 × 5). In detail, to control for the aggregate-bias-belief about future return $\theta_{t+1,s}$, in every month t, we sort stocks into 5 sections based on $\theta_{t+1,s}$. Then within each section, we continue to sort stocks into 5 portfolios based on $\beta_{t,s}$. We will have 25 portfolios each month; we then track these portfolios' returns in the next month $t+1$. To allocate the stocks' weight in each portfolio, we use both EW and VW schemes. This is the dependent-double-sorting procedure, with $\theta_{t+1,s}$ as the control variable and $\beta_{t,s}$ as the sorting variable. We also carry out the independent double-sorting procedure.[25] The result is qualitatively similar to that of the dependent double-sorting procedure, so we will not report it here.

These 25 portfolios will have 25 time series of returns. We then take the average of each time series. In each $\theta_{t+1,s}$ section, we compute the H–L (high β - low β) portfolio's return, Sharpe ratio, CAPM α, and Fama and French (2015) 5 factors model plus momentum factor (FF5.MOM) α, and t-stats. As before, we use Newey and West (1987) methods with up to 7 lags when computing the standard error.

Table 4 presents the double-sorting results. The returns of these 25 portfolios are excess returns from the risk-free rate. All the portfolios' returns and α are in percentages, the rest are in decimal values. If Corollary 1 holds, we should expect that in every section of $\theta_{t+1,s}$, the H–L return and H–L α are statistically positive. For the *EW* scheme, the H–L return and H–L FF5.MOM α are only statistically positive in the first section of $\theta_{t+1,s}$. In the second section of $\theta_{t+1,s}$, H–L return and H–L FF5.MOM α are positive but small and insignificant. In the third and the fourth sections of $\theta_{t+1,s}$, the H–L return and the H–L FF5.MOM α are not significantly different from zero. In the fifth section of $\theta_{t+1,s}$, the H–L return and H–L FF5.MOM α are significantly negative. For the *VW* scheme, we have the same qualitative results.

Recall that the first section of $\theta_{t+1,s}$ is the underpriced section, and the fifth section of $\theta_{t+1,s}$ is the overpriced section. Hence, the empirical results give a mixed conclusion on Corollary 1. As 5 sections of $\theta_{t+1,s}$ are sorted from most underpriced to most overpriced, Corollary 1 seems only hold in the most underpriced section but not in the overpriced sections.

5.4.2. Testing Corollary 2

Recall that Corollary 2 states that the expected return gap between the most optimistically biased (overvalued) and the most pessimistically biased (under-valued) stocks is negative when we control for stock β. To test that, we perform a double-sorting procedure (5 × 5) as described in Section 5.4.1. Now, we use $\beta_{t,s}$ as the control variable and $\theta_{t+1,s}$ as the sorting variable. We perform both independent and dependent double-sorts. We also use both *EW* and *VW* schemes.

If Corollary 2 holds, we should expect that in every section of β, the H–L portfolio return and H–L FF5.MOM α are significantly negative. In Table 5, we report the results coming from the dependent double-sort with the *EW* scheme. In all 5 sections of β, the H–L portfolio return and H–L FF5.MOM α are always significantly smaller than zero. The results are qualitatively similar across the sorting methods (dependent and independent double-sorts) and weighting schemes (VW, and EW). Therefore, we do not report them here.[26] In brief, Corollary 2 holds.

5.4.3. Testing Proposition 2

Proposition 2 says that a positively biased belief about the future return will hike up the current return (price), and vice versa with a negatively biased belief. Therefore, if we sort stocks into portfolios based on their aggregate-bias-belief return ($\theta_{t+1,s}$) from very negative (L) to very positive (H), we should see the *current return* at time t, $r_{t,H--L}$ of the H–L portfolio being significantly positive.

We test this implication by applying the same sorting procedure as in Section 5.3. In detail, in every month t, we sort stocks into 10 portfolios based on the aggregate-bias-belief return of the next period ($\theta_{t+1,s}$). Then we track the current return of these portfolios at time t. After that, we take the average of each portfolio's return and compute the H–L portfolio's return and statistics as in Section 5.3.

Table 6 reports the empirical results. In both *EW* and *VW* schemes, the current return increases monotonically from portfolio 1 to 10. The H–L portfolio's return and FF5.MOM α are both significantly greater than zero. In brief, we can conclude that Proposition 2 holds.

5.5. Anomalies in underpriced and overpriced stocks

Our model suggests that investors gather stock information to make their own beliefs about future returns. The information comes from the stocks' characteristics. Hence, these characteristics contribute to the aggregate-bias belief returns which makes stocks underpriced or overpriced. This mispricing will be corrected in the next period. By that channel, characteristics have a predicting power over future stocks' returns.

In the literature, researchers tend to study the predicting power of characteristics on the expected return using all stocks in the cross-section. A natural question arises: is this predicting power the same among over and underpriced stocks? Some recent studies by Baker

[25] In each month t, we sort stocks into 5 sections based on $\theta_{t+1,s}$, and into another 5 sections based on $\beta_{t,s}$ independently. Then, we combine these two sets of 5 portfolios to have 25 portfolios. The results are available upon request.

[26] Additional results are available upon request.

Table 4

Double sorting (5 × 5) with $\theta_{t+1,s}$ as the control variable and $\beta_{t,s}$ as the sorting variable. In every month t, we sort stocks into 5 sections based on $\theta_{t+1,s}$. Then, within each section, we continue to sort stocks into 5 portfolios based on $\beta_{t,s}$. We will have 25 portfolios each month; we then track these portfolios' returns in the next month $t+1$. To allocate the stock's weight in each portfolio, we use both EW and VW schemes. These 25 portfolios will have 25 time series of returns. We then take the average of each time series. In each $\theta_{t+1,s}$ section, we compute the H–L (high β - low β) portfolio's return, Sharpe ratio, CAPM α, Fama and French (2015) 5 factors model plus momentum factor (FF5.MOM) α, and t-stats. We use Newey and West (1987) methods with up to 7 lags when computing the standard error. The returns of these 25 portfolios are excess returns from the risk-free rate. Sample size: 1980–2018. All the portfolios' returns and α are in percentages, the rest are decimal values.

$\beta_t \downarrow$, $\theta_{t+1,s} \rightarrow$	1	2	3	4	5	$\beta_t \downarrow$, $\theta_{t+1,s} \rightarrow$	1	2	3	4	5
Panel A: EW Scheme						*Panel B: VW Scheme*					
L	2.91	1.41	0.51	−0.07	−1.68	L	2.14	1.11	0.43	−0.07	−1.39
2	3.22	1.53	0.58	−0.20	−2.01	2	2.57	1.33	0.51	−0.15	−1.64
3	3.42	1.51	0.56	−0.39	−2.10	3	2.87	1.33	0.48	−0.39	−1.82
4	3.78	1.65	0.54	−0.49	−2.65	4	3.13	1.48	0.48	−0.48	−2.29
H	4.02	1.51	0.38	−0.87	−3.57	H	3.35	1.47	0.33	−0.91	−3.11
H–L	1.11	0.10	−0.13	−0.80	−1.89	H–L	1.22	0.35	−0.10	−0.85	−1.72
H–L.t-stats	3.06	0.34	−0.43	−2.59	−5.72	H–L.t-stats	3.17	1.08	−0.33	−2.60	−4.62
H–L.SR	0.15	0.02	−0.02	−0.14	−0.32	H–L.SR	0.16	0.06	−0.02	−0.14	−0.27
CAPM α	0.52	−0.49	−0.68	−1.36	−2.41	CAPM α	0.53	−0.27	−0.72	−1.47	−2.32
CAPM α t-stats	1.61	−1.89	−2.90	−5.55	−8.54	CAPM α t-stats	1.60	−0.97	−2.89	−5.79	−7.62
FF5.MOM α	1.17	0.06	0.03	−0.42	−1.48	FF5.MOM α	1.39	0.39	0.05	−0.53	−1.41
FF5.MOM α t-stats	2.46	0.15	0.10	−1.44	−4.00	FF5.MOM α t-stats	2.99	1.09	0.17	−2.11	−4.19

Table 5

Double sorting (5 × 5) with $\beta_{t,s}$ as the control variable and $\theta_{t+1,s}$ as the sorting variable. In every month t, we sort stocks into 5 sections based on $\beta_{t,s}$. Then, within each section, we continue to sort stocks into 5 portfolios based on $\theta_{t+1,s}$. We will have 25 portfolios each month; we then track these portfolios' returns in the next month $t+1$. To allocate the stock's weight in each portfolio, we use the EW scheme. These 25 portfolios will have 25 time series of returns. We then take the average of each time series. In each β section, we compute the H–L (high $\theta_{t+1,s}$–low $\theta_{t+1,s}$) portfolio's return, Sharpe ratio, CAPM α, Fama and French (2015) 5 factors model plus momentum factor (FF5.MOM) α, and t-stats. We use Newey and West (1987) methods with up to 7 lags when computing the standard error. The returns of these 25 portfolios are excess returns from the risk-free rate. Sample size: 1980–2018. All the portfolios' returns and α are in percentages, the rest are decimal values.

$\theta_{t+1,s} \downarrow$, $\beta_t \rightarrow$	1	2	3	4	5
L	2.60	2.66	2.77	3.29	4.16
2	1.28	1.32	1.26	1.53	1.46
3	0.59	0.61	0.56	0.52	0.17
4	−0.01	0.03	−0.12	−0.38	−0.98
H	−1.31	−1.23	−1.55	−2.23	−3.54
H–L	−3.91	−3.88	−4.32	−5.52	−7.70
H–L.t-stats	−14.47	−16.23	−18.13	−17.23	−16.73
H–L.SR	−1.01	−1.15	−1.18	−1.20	−1.16
CAPM α	−3.91	−3.87	−4.32	−5.48	−7.58
CAPM α t-stats	−14.83	−16.35	−17.79	−16.56	−16.81
FF5.MOM α	−4.51	−3.84	−4.09	−5.31	−8.10
FF5.MOM α t-stats	−9.71	−8.37	−10.52	−10.83	−12.70

and Wurgler (2006), Liu et al. (2018), Stambaugh et al. (2012, 2015), Stambaugh and Yuan (2017) address the above question. They find that the predicting power of certain characteristics is different between under and overpriced stocks. Motivated by these studies, using our model, we try to investigate the predicting power on stock returns of some well-known characteristics conditioned on over and underpriced stocks. On one hand, we confirm major points from previous studies. On the other hand, we find that more characteristics only offer a return premium in overpriced stocks. When it comes to underpriced stocks, the return premium from characteristics either disappears or turns to the opposite sign.

To perform such analysis, we apply a double sort (2 × 5) as described in previous parts. First, in each month *"t"*, we sort stocks into two equal sections of under and overpriced based on the aggregate-bias-belief about future returns ($\theta_{t+1,s}$). Then, within each section, we sort the stocks into 5 sub-sections from low to high values of each characteristic. For each portfolio, we use both EW and VW weighting schemes. This is the dependent double-sort procedure. We also perform the independent double-sort. The results are qualitatively similar, so we do not report them here. As before, with each characteristic, we compute the H–L portfolio's return for each underpriced or overpriced section and the FF5.MOM α of the H–L portfolio.

Table 7 reports these analyses. All the portfolio returns and α are percentages. The first 3 rows report the results for β, size (*mve*), and idiosyncratic volatility (*idiovol*). For the size premium, Baker and Wurgler (2006) find that after a high-sentiment period when stocks are usually overpriced, the expected returns of small stocks tend to be lower than those of big stocks. This is because the overpricing in small stocks is hard to arbitrage in the high-sentiment period. Hence, a correction after that period will make the small stocks' returns lower than those of the big stocks. Thus, the H–L portfolio sorted on size should be positive. The logic is inverted when we consider the return after a low-sentiment period. Now, small stocks tend to deliver higher returns than big stocks. Because the underpricing in small stocks is rigid to arbitrage in a low-sentiment period, the correction after that period makes small stocks' returns higher than big stocks. Therefore, the H–L portfolio's return sorted on size should be negative.

With our model, our results corroborate those of Baker and Wurgler (2006). In detail, small stocks only have a higher expected return than big stocks in the under-priced section. This makes the H–L portfolio and its α deliver a negative return. For the overpriced section, the relationship is inverted. Small stocks are expected to have a lower return than big stocks; hence, the H–L portfolio's return and its α are positive. This is true for both EW and VW schemes.

Having the same logic as Baker and Wurgler (2006), Stambaugh et al. (2015) argue that stocks with high *idiovol* are rigid to arbitrage. Hence, in the underpriced section, they find that high-*idiovol* stocks have a higher expected return than low-*idiovol* stocks. The relation is inverted in the overpriced section. With our model, we find a similar result with both EW and VW schemes in Table 7. The idiosyncratic volatility anomaly happens only in the overpriced section. In the underpriced section, we have the inverted result when high-*idiovol* stocks have higher expected returns (and α) than low-*idiovol* stocks.

Liu et al. (2018) show that β is highly correlated with *idiovol*. Hence, they show that the β anomaly only happens in overpriced stocks, while in underpriced stocks, high-β stocks have higher expected returns (and α) than low-β stocks. With our model, we also confirm these conclusions when testing Corollary 1 in Section 5.4.1. Here, we have a similar result. Indeed, H–L portfolio return and its α are positive in the underpriced section and negative in the overpriced section.

Using our model, we also study more characteristics. For the ill measure of Amihud (2002), we should expect a positive H–L return and a positive α. However, this is only true with the underpriced section. The H–L return is negative, and its α is statistically indifferent from zero, in the overpriced section. Since ill and mve (size) have a high negative correlation, this phenomenon is in fact in the opposite direction from the one in mve (size).

Table 6
Univariate portfolio sorting 1980–2018. In every month t, we rank stocks from low to high value of $\theta_{t+1,s}$. Then, we allocate stocks into 10 equal quantiles for every month t. Every quantile is a portfolio. We track each portfolio's return in this current month t. To allocate the stock's weight within a portfolio, we use both equal-weighted (EW) and value-weighted (VW) methods. We compute the high minus low portfolio (H–L). After that, we take the time-series average of each portfolio. For the H–L portfolio, we compute the Sharpe ratio, CAPM α, and Fama and French (2015) 5 factors model plus momentum factor (FF5.MOM) α. We use Newey and West (1987) with up to 7 lags to calculate the standard error for t-stats. Our sample is from 1980–2018.

EW Portfolios	Sorting based on θ_{t+1}	VW Portfolios	Sorting based on θ_{t+1}
L	−1.96	1	−0.67
2	−0.80	2	−0.01
3	−0.23	3	0.33
4	0.16	4	0.65
5	0.46	5	0.82
6	0.72	6	1.04
7	1.03	7	1.30
8	1.38	8	1.59
9	1.97	9	2.06
H	3.61	10	3.45
H–L	5.56	H–L	4.11
H–L t-stats	9.22	H–L t-stats	8.80
H–L Sharpe Ratio	0.54	H–L Sharpe Ratio	0.50
CAPM α	5.24	CAPM α	3.86
CAPM α t-stats	8.67	CAPM α t-stats	8.26
FF5.MOM α	3.33	FF5.MOM α	2.27
FF5.MOM α t-stats	3.62	FF5.MOM α t-stats	3.32

Table 7
Double sorting (2 × 5) with $\theta_{t+1,s}$ as the control variable and different sorting variables. In every month t, we sort stocks into 2 sections based on $\theta_{t+1,s}$. The bottom section we call Underpriced, while the other is called Overpriced. Then, within each section, we continue to sort stocks into 5 portfolios from low to high values of each characteristic. We will have 10 portfolios each month for each characteristic, we then track these portfolios' returns in month $t+1$. To allocate the stock's weight in each portfolio, we use both EW and VW schemes. These 10 portfolios will have 10 time series of returns. In each $\theta_{t+1,s}$ section, we compute the H–L (high–low) portfolio's return and the Fama and French (2015) 5 factors model plus momentum factor (FF5.MOM) α. We use Newey and West (1987) methods with up to 7 lags when computing the standard error. The returns of H–L portfolios are excess returns from the risk-free rate. The sample size is from 1980 to 2018. All the portfolios' returns and α are percentages.

Sorting variable	Return/ FF5.MOM α	Underpriced	Overpriced	Underpriced	Overpriced
		EW		VW	
β	H–L	0.79*	−1.47**	0.86**	−1.42**
	FF5.MOM_α	0.85*	−1.21**	0.93**	−1.2**
idiovol	H–L	1.7**	−2.11**	1.06**	−2.18**
	FF5.MOM_α	2.39**	−1.76**	1.77**	−1.38**
mve	H–L	−1.7**	0.67**	−1.26**	0.8**
	FF5.MOM_α	−2.33**	0.86**	−1.75**	0.9**
ill	H–L	0.95**	−0.51**	0.39	−0.7**
	FF5.MOM_α	1.36**	−0.4	0.69**	−0.45
bm	H–L	0.24	1.65**	0.2	1.12**
	FF5.MOM_α	−0.02	0.81**	−0.01	0.23
mom12 m	H–L	−0.37	1.63**	0.25	1.32**
	FF5.MOM_α	−1.35*	1.51**	−0.65	1.15**
mom1 m	H–L	−1.48**	−0.49**	−0.77**	−0.14
	FF5.MOM_α	−1.35**	0.18	−0.69*	0.35
maxret	H–L	1.27**	−2.22**	0.8**	−2.12**
	FF5.MOM_α	1.96**	−2.04**	1.42**	−1.45**
indmom	H–L	0.62**	0.83**	0.37	0.49*
	FF5.MOM_α	0.18	0.67	0.04	0.47
invest	H–L	−1.11**	−0.77**	−0.46**	−0.76**
	FF5.MOM_α	−1**	−0.33**	−0.36*	−0.27
roeq	H–L	−0.4	1.35**	−0.13	1.03**
	FF5.MOM_α	−1.22**	1.11**	−0.86**	0.78**
std_turn	H–L	1.26**	−0.73**	1.03**	−0.55*
	FF5.MOM_α	1.32**	−1.05**	1.19**	−0.61**

*$p \leq 0.05$.
**$p \leq 0.01$.

For the book-to-market ratio (bm), we can see that the H–L portfolio and α are only significantly positive in the overpriced section. In the underpriced section, the H–L portfolio and α are still positive but very small and insignificantly different from zero.

For momentum in the last 12 months, *mom12m*, the H–L return and its α are only significantly positive in the overpriced section. The momentum effect is almost insignificant in the underpriced section.

For momentum in the last month, $mom1m$, the short-term reversal effect happens in both the underpriced and the overpriced section. However, in the overpriced section, this reversal effect reflected in the H–L return is smaller, and its α is statistically indifferent from zero.

For industry momentum, $indmom$, the H–L portfolio's return is positive in both sections, although its α is not different from zero.

For the maximum return measure, $maxret$, of Bali, Cakici, and Whitelaw (2011), we find an interesting result. According to Bali et al. (2011), people like winning the lottery by over-buying stocks with the maximum past return. Hence, we should expect a negative H–L return and negative α when sorting on $maxret$. However, with our model, this is true only with the overpriced section. In the underpriced section, the H–L return and its α in fact have the opposite sign (positive).

For the investment measure, *invest*, of Chen and Zhang (2010), we find a consistently negative H–L return and α in both the overpriced and the underpriced section. This result corroborates previous studies on the investment asset pricing models, such as (Fama & French, 2015; Hou et al., 2015).

With return on equity, *roeq*, a measure of profitability, according to Hou et al. (2015), we should expect a positive H–L return and a positive α. However, this is true only with the overpriced section. In the under-priced section, the H–L return and its α are both negative. Although in the underpriced section, the H–L return is not statistically different from zero, the α is significantly negative.

Last but not least, we also find an interesting result with the standard deviation of turnover, *std_turn*, a measure of liquidity volatility. In a normal sense, liquidity volatility can be seen as a risk and should be a reward for the return. Hence, the relationship between liquidity–volatility and expected returns should be positive. However, Chordia, Subrahmanyam, and Anshuman (2001) document a robust, strange, negative relation between liquidity–volatility and expected returns. With our model, we find that this negative relation is only significant in the overpriced section with the negative H–L return and α. In the underpriced section, the relation is in fact positive between liquidity volatility and expected returns with a significantly positive H–L return and a significantly positive α.

The explanation of this result with *std_turn* and the result with *maxret* may be similar to the explanation of the *idiovol* result. In detail, stocks with a high value in *maxret* or *std_turn* are harder to trade, and hence rigid to current arbitrage activities. Therefore, their expected return will be higher in the future when they are underpriced, and vice versa.

In brief, some well-documented anomaly premia and risk premia from β, *idiovol, bm, mom12m, maxret*, and *invest, std_turn* are only significant in overpriced stocks. The premium either disappears (*bm, mom12m*) or changes sign (β, *idiovol, maxret, invest, std_turn*) in underpriced stocks.

The size premium and illiquidity premium are only significant in underpriced stocks. These premia change signs in overpriced stocks. The premium from *mom1 m, indmom, invest* is consistent in both underpriced and overpriced stocks.

These results add another layer to the new research orientation in our field surrounding anomalies conditioned on mispricing. This empirical evidence suggests that we should also control for mispricing when investigating an anomaly. Most of the documented anomalies happen only in overpriced stocks. One plausible reason for this is the existence of the arbitrage risk when taking a short position on overpriced stocks. Other explanations for this phenomenon would be an interesting topic for further research.

5.6. Control for other mispricing factors and sentiment measures

The gap in expected return between underpriced and overpriced stocks is reflected in the H–L return when we do a univariate sort on $\theta_{t+1,s}$. However, as $\theta_{t+1,s}$ is ex-post information, we cannot use $\theta_{t+1,s}$ to predict a stock's return. In Section 5.3, we show that $\theta^{MA}_{t+1,s}$, ex-ante information, is a good predictor of future stock returns. The mispricing H–L return generated by this predictor cannot be explained by FF5 factors nor by the momentum factor. Hence, here we ask a further question: can other existing mispricing factors or sentiment measures explain these returns generated by $\theta_{t+1,s}$, or predicted by $\theta^{MA}_{t+1,s}$? The short answer is "No".

Stambaugh and Yuan (2017) use 11 anomalies to construct a 4 factors model, whose 2 factors are related to mispricing. The model explains the cross-sectional stock return very well. We test to see whether the spread return generated $\theta_{t+1,s}$ and the spread return predicted by $\theta^{MA}_{t+1,s}$ can be explained by the (Stambaugh & Yuan, 2017) model.

Table 8
The α from Stambaugh and Yuan (2017) mispricing factors model. Univariate portfolio sorting. In every month t, we rank stocks from low to high values of either $\theta^{MA}_{t+1,s}$ or $\theta_{t+1,s}$. Then, we allocate stocks into 10 equal quantiles in every month t. Every quantile is a portfolio. We track each portfolio's return in the next month $t+1$. To allocate the stock's weight within a portfolio, we use both equal-weighted (EW) and value-weighted (VW) methods. We compute α from the missing factors model of Stambaugh and Yuan (2017) for each portfolio. We use Newey and West (1987) with up to 7 lags to calculate the standard error for t-stats. Our sample is from 1980 to 2016. All the α are percentages.

Sort on $\theta_{t+1,s}$	EW α	VW α	*Sort on* $\theta^{MA}_{t+1,s}$	EW α	VW α
L	5.7**	4.34**	L	1.46**	0.84**
2	2.88**	2.56**	2	1.11**	0.66**
3	2.02**	1.78**	3	0.96**	0.63**
4	1.42**	1.29**	4	0.85**	0.55*
5	0.89**	0.84**	5	0.94**	0.55*
6	0.42	0.41	6	0.82**	0.52
7	0	0.03	7	0.88**	0.58
8	−0.49*	−0.42	8	0.85*	0.5
9	−1.25**	−1.03**	9	0.78*	0.47
H	−2.99**	−2.31**	H	0.31	0.06
H–L	−8.69**	−6.65**	H–L	−1.15**	−0.77**

*$p \le 0.05$.
**$p \le 0.01$.

To do that, we take all the portfolios sorted on $\theta_{t+1,s}$ and sorted on $\theta^{MA}_{t+1,s}$ in Table 3. Then, we calculate α from the (Stambaugh & Yuan, 2017) 4 factors model for each portfolio. We get the data from the 4 factors model from Stambaugh's website.[27] The 4 factors time series end in 2016. The results are in Table 8.

Clearly, we can see that most of the Stambaugh and Yuan (2017) α of 10 portfolios sorting on θ_{t+1} cannot be explained by the mispricing 4 factors model. The α is not different from zero in portfolios 6, 7, and 8 in the VW scheme, and portfolios 6, and 7 in the EW scheme. The H–L portfolio's α is huge and cannot be explained by the 4 factors mispricing model in both weighting schemes. Hence, the return premium generated by our $\theta_{t+1,s}$ cannot be explained by Stambaugh and Yuan (2017) 4 factors model.

However, recall that $\theta_{t+1,s}$ is ex-post information, which cannot fully be known at time t. Hence, we also perform the previous exercise with $\theta^{MA}_{t+1,s}$, an ex-ante variable fully known at time t, as the sorting variable. The result is also in Table 8. Once again, the mispricing 4 factors model cannot explain the α in almost all portfolios. Only the "H" portfolio α in both weighting schemes and the 6, 7, 8, and 9 portfolios' α in the VW scheme are not different from zero. The H–L portfolio α is still big: −1.15% monthly in the EW scheme, and −0.77% monthly in the VW scheme. These α values are all significantly different from zero. That would mean the Stambaugh and Yuan (2017) 4 factors model also cannot fully explain the return generated by $\theta^{MA}_{t+1,s}$. That would suggest the H–L portfolio return sorting on $\theta^{MA}_{t+1,s}$ can be used as a new mispricing factor that can complement the Stambaugh and Yuan (2017) model.

As the Stambaugh and Yuan (2017) mispricing factors model cannot explain the return premium generated by $\theta_{t+1,s}$ or by $\theta^{MA}_{t+1,s}$, we ask whether another measure of sentiment can explain it. Baker and Wurgler (2006) propose a market-wide sentiment measure, which is orthogonal to other macro factors. The measure is named $SENT^{\perp}$. The evidence from Avramov et al. (2019), Baker and Wurgler (2006), Stambaugh et al. (2012) shows that the measure has predicting power on the future return. Therefore, we try to test whether $SENT^{\perp}$ can explain the return premium generated by $\theta_{t+1,s}$, and the predicted return premium by $\theta^{MA}_{t+1,s}$.

To do that, we repeat the same univariate sorting exercise as in Table 8. After sorting, for each portfolio p, we run two time series regressions:

$$r_{p,t+1} - r_{f,t+1} = \alpha + b \cdot SENT^{\perp}_{t} + c \cdot MKTRF_{t+1} + d \cdot SMB_{t+1} + e \cdot MGMT_{t+1} + f \cdot PERF_{t+1} + \eta_{t+1} \quad (12)$$

[27] http://finance.wharton.upenn.edu/~stambaug/

Table 9
The effect of Baker and Wurgler (2006) $SENT_t^{\perp}$ measure on mispricing returns. In every month t, we rank stocks from low to high values of either $\theta^{MA}_{t+1,s}$ or $\theta_{t+1,s}$. Then, we allocate stocks into 10 equal quantiles in every month t. Every quantile is a portfolio. We track each portfolio's return in the next month $t+1$. To allocate the stock's weight within a portfolio, we use both equal-weighted (EW) and value-weighted (VW) methods. In panel A, we run the time series regression $r_{p,t+1} - r_f = \alpha + b \cdot SENT_t^{\perp} + c \cdot MKTRF_{t+1} + d \cdot SMB_{t+1} + e \cdot MGMT_{t+1} + f \cdot PERF_{t+1} + \eta_{t+1}$ for each portfolio p. In panel B, we run the time series regression $r_{p,t+1} - r_f = \alpha + b \cdot SENT_t^{\perp} + \eta_{t+1}$, where $MKTRF$, SMB, $MGMT$, and $PERF$ are the 4 factors of Stambaugh and Yuan (2017) models. We report α and b for each portfolio's regression. We use Newey and West (1987) with up to 7 lags to calculate the standard error for t-stats. Our sample is from 1980 to 2016 for Panel A and 1980 to 2018 for Panel B. All the α are percentages.

Sort on $\theta_{t+1,s}$	EW α	EW b	VW α	VW b	*Sort on* $\theta^{MA}_{t+1,s}$	EW α	EW b	VW α	VW b
Panel A: $r_{p,t+1} - r_f = \alpha + b \cdot SENT_t^{\perp} + c \cdot MKTRF + d \cdot SMB + e \cdot MGMT + f \cdot PERF + \eta_{t+1}$									
1	5.5**	0.01	4.07**	0.01	1	1.37**	0.01	0.78*	0
2	2.64**	0.01	2.33**	0.01	2	1.03**	0	0.56	0.01
3	1.88**	0.01	1.62**	0.01	3	0.86**	0.01	0.55	0.01
4	1.3**	0	1.15**	0.01	4	0.77**	0	0.49	0
5	0.8**	0	0.73*	0	5	0.86**	0	0.5	0
6	0.37	0	0.33	0	6	0.75*	0	0.45	0
7	−0.03	0	−0.01	0	7	0.82*	0	0.5	0
8	−0.46	0	−0.38	0	8	0.81*	0	0.47	0
9	−1.18**	0	−0.94**	0	9	0.71	0	0.42	0
10	−2.91**	0	−2.17**	−0.01	10	0.3	0	0.02	0
H–L	−8.4**	−0.01	−6.24**	−0.02	H–L	−1.07**	0	−0.76**	0
Sort on $\theta_{t+1,s}$	EW α	EW b	VW α	VW b	*Sort on* $\theta^{MA}_{t+1,s}$	EW α	EW b	VW α	VW b
Panel B: $r_{p,t+1} - r_f = \alpha + b \cdot SENT_t^{\perp} + \eta_{t+1}$									
1	5.32**	−0.01	3.97**	−0.01	1	1.35*	−0.01	0.98	−0.01
2	2.71**	−0.01	2.41**	−0.01	2	1.09*	−0.01	0.84	−0.01
3	2.08**	−0.01	1.81**	−0.01	3	0.97	−0.01	0.81	−0.01
4	1.5**	−0.01	1.37**	−0.01	4	0.88	−0.01	0.77	−0.01
5	1.05*	−0.01	0.99*	−0.01	5	0.88	−0.01	0.72	−0.01
6	0.6	−0.01	0.58	−0.01	6	0.77	−0.01	0.65	−0.01
7	0.21	−0.01*	0.24	−0.01*	7	0.81	−0.02	0.64	−0.01
8	−0.26	−0.02**	−0.17	−0.02**	8	0.67	−0.02	0.53	−0.02
9	−1.05*	−0.02**	−0.85	−0.02**	9	0.48	−0.02	0.41	−0.02
10	−2.89**	−0.02**	−2.18**	−0.03**	10	−0.03	−0.02	−0.07	−0.02*
H–L	−8.21**	−0.01	−6.15**	−0.02	H–L	−1.38**	−0.01**	−1.05**	−0.01*

*$p \le 0.05$.
**$p \le 0.01$.

$$r_{p,t+1} - r_{f,t+1} = \alpha + b \cdot SENT_t^{\perp} + \eta_{t+1} \quad (13)$$

where $MKTRF$, SMB, $MGMT$, and $PERF$ are the 4 factors in Stambaugh and Yuan (2017)'s model and $r_{f,t+1}$ is the risk-free rate. We report α and b from the first regression in Panel A of Table 9, while α and b from the second regression are in Panel B.

If $SENT_t^{\perp}$ has some explanatory power over $r_{p,t+1}$, we should expect b to be significantly negative. This is because a rise in market-wide sentiment tends to make assets relatively overpriced, thus decreasing the expected return. When a decrease in market-wide sentiment happens, stocks tend to be relatively underpriced, hence creating an increase in expected return. If the explaining power of $SENT_t^{\perp}$ is big enough, the α in these two above regressions should be statistically non-different from zero.

However, in Panel A of Table 9, all b values are not significantly different from zero. This is true for both EW and VW schemes, and also true when we use both ex-post information $\theta_{t+1,s}$ and ex-ante information $\theta^{MA}_{t+1,s}$ as sorting variables. All the α in Panel A of Table 9 are qualitatively the same as in Table 8. The α from the combining model between the 4 factors model and the $SENT_t^{\perp}$ measure is still the same and significantly different from zero. This is true for both weighting schemes and for both ex-post and ex-ante sorting variables.

When we drop the 4 factors of Stambaugh and Yuan (2017) in Panel B of Table 9, we can see b being significantly negative in portfolios 7, 8, 9, and H sorting on $\theta_{t+1,s}$. The α is quasi-similar to the α in Panel A.

When we sort on $\theta^{MA}_{t+1,s}$, b is significantly negative in portfolios H, and H–L in the VW scheme, and in portfolios H–L in the EW scheme.

Note that in both panels A and B, the α of the H–L return, regardless of weighting schemes and sorting variables, is always huge and significantly different from zero.

In brief, neither the Stambaugh and Yuan (2017) 4 factors model nor the Baker and Wurgler (2006) sentiment measure has strong explanatory power over the mispricing return premium generated by an ex-post or ex-ante variable from our model. This would suggest that our model captures some information that these cited factors and measures cannot capture. The spread portfolio return thus can be used as a new mispricing factor that complements the previous factors and measures.

6. Conclusion

This paper proposes a bridge between the sentiment model and the risk-based factor model in asset pricing. According to our model, the aggregate-bias-belief about future stock returns is a major force creating variation in future stock returns. This is because the aggregate-bias-belief about future stock returns ($\boldsymbol{\theta}_{t+1}$) will immediately shoot up or shoot down the current stock price. This makes stocks relatively overpriced or underpriced. Hence, in the next period, a price correction happens. Then, the unbiased expected next-period return is negatively correlated with the aggregate-bias-belief about the return. We also show that this mispricing return can be represented as a return premium because of the exposure of stock returns to a latent factor. We perform empirical analyses to show the above implications of the model.

To do that, we use characteristics both as proxies for the covariance risk to the latent factor on the risk-based explanation and as information that investors use to form their beliefs about the future returns on the sentiment-based explanation. Therefore, we impose a factor structure based on characteristics for $\boldsymbol{\theta}_{t+1} = \mathbf{Z}_{t+1} \cdot \mathbf{f}_{t+1}$. We estimate $\mathbf{f}_{t+1}$ and $\boldsymbol{\theta}_{t+1}$ in every month and take the average of $\mathbf{f}_{t+1}$. The results show that only a handful of characteristics are both statistically and economically significant relating to size, volatility, momentum, and liquidity, for example, market size, return volatility, price momentum of the previous month and of the previous twelve months, industry momentum, Amihud (2002)'s illiquidity and trading turnover. This result corroborates previous research on these characteristics and factors (see, for example, Green et al., 2017; Harvey, 2017; Harvey et al., 2016; Hou et al., 2015, 2020; Kelly et al., 2019).

Although a reduced number of characteristics can explain well the cross-sectional of stocks return, our study does not ensure trading profitability when using characteristics. Since the return generated by these characteristics relates to a latent risk factor in our model, the return is risky in its nature. Therefore, proper risk management and trading cost optimization are essential to any investors who would like to capture the return from firms' characteristics and can be an interesting topic for further research. In addition, studying the wealth effect, and the reputation effect on the sentiment model can also be an interesting venue for further research. For example, which investors (rich or poor, famous or not) will dominate in the economy and impact the stock's return?

Our model entails two Corollaries 1 and 2 and the Proposition 2 that can be tested. Corollary 1 states that after controlling for the aggregate-bias-belief about future return, high-β stocks predict higher returns than low-β stocks. Corollary 2 states that after controlling for stock β, the difference in expected return between the most optimistically biased and the most pessimistically biased stocks is negative. Proposition 2 states that a positive bias view about future return will increase the recent return and vice versa with a negative bias view about the future return.

We can confirm that the Proposition 2 and Corollary 2 hold. These results align with the empirical and theoretical evidence from the literature (see i.e. Avramov et al., 2019; Baker & Wurgler, 2006; Daniel et al., 2001; Kozak et al., 2018; Liu et al., 2018; Stambaugh et al., 2012, 2015; Stambaugh & Yuan, 2017).

However, we find a mixed result on the Corollary 1. In our test, the high β stocks do not always have a higher expected return than the low β stocks. The result adds additional evidence to the β anomaly in the literature where higher β stocks can have lower expected returns than low β stocks. Frazzini and Pedersen (2014), Liu et al. (2018) and the references therein give a deep analysis of the anomaly and provide possible explanations. For the scope of this paper, we are silent about the cause of this anomaly.

At the individual stock level, using an ex-ante sorting variable, we also show that the expected return gap between top overpriced and top underpriced stocks can be from 0.72% to 1.38% a month after controlling for FF5 factors plus the momentum factor, other mispricing factors, such as Stambaugh and Yuan (2017) 4 factors model, and the sentiment measure of Baker and Wurgler (2006). This return gap, or premium, is a potential candidate for a mispricing factor that can complement other factors from previous studies. On the market-wide level, the aggregate-bias-belief about the market return (i) has a mean reverting to zero, (ii) shoots up and is volatile in financial crises. Using our model, we show that some well-known anomalies and risk premia hold only in either under or overpriced stocks, but not in the entire cross-section.

Data availability

The authors do not have permission to share data.

Appendix A. Proof of Proposition 1

We can rearrange Eq. (6) to get $\mathbf{p_t}$

$$\mathbf{p_t} = \frac{E_t[\mathbf{p_{t+1}} + \mathbf{d_{t+1}}] - \gamma \cdot \boldsymbol{\Omega}_t \cdot \mathbf{s_t} + \sum_i \frac{\gamma}{\gamma_i} \cdot \mathbf{b_{i,t+1}}}{R_f} \tag{A.1}$$

Define $\mathbf{e_s}$ as a vector $(S \times 1)$ with 1 in the sth row and zero elsewhere. For a security s, we have

$$E_t[R_{t+1,s}] = R_f + \frac{\mathbf{e_s^T} \cdot \gamma \cdot \boldsymbol{\Omega}_t \cdot \mathbf{s_t} - \sum_i \frac{\gamma}{\gamma_i} \cdot b_{i,t+1,s}}{p_{t,s}}$$

$$= R_f + \gamma \cdot \frac{Cov_t\left(p_{t+1,s} + d_{t+1,s}, [\mathbf{p_{t+1}} + \mathbf{d_{t+1}}]^{\mathbf{T}} \cdot \mathbf{s_t}\right)}{p_{t,s}} - \sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}}$$

$$E_t[R_{t+1,s}] = R_f + \gamma \cdot Cov_t\left(R_{t+1,s}, R_{t+1,M}\right) \cdot \mathbf{p_t}^T \cdot s_t - \sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}} \tag{A.2}$$

Note that $\mathbf{s_t}$ is the market portfolio (M), which has a value of $[\mathbf{p_t} + \mathbf{d_t}]^{\mathbf{T}} \cdot \mathbf{s_t}$ at time t and $[\mathbf{p_{t+1}} + \mathbf{d_{t+1}}]^{\mathbf{T}} \cdot \mathbf{s_t}$ at time $t+1$. From Eq. (A.2), we get

$$E_t[R_{t+1,s}] = R_f + \gamma \cdot \beta_{t,s} \cdot \mathbf{p_t}^T \mathbf{s_t} \cdot Var_t(R_{t+1,M}) - \sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}} \tag{A.3}$$

where $\beta_{t,s} = \frac{Cov_t\left(R_{t+1,s}, R_{t+1,M}\right)}{Var_t(R_{t+1,M})}$. Now, call the weight of stock s in portfolio M as $w_{t,s} = \frac{p_{t,s} \cdot x_{t,s}}{\mathbf{p_t^T} \cdot \mathbf{s_t}}$. Multiply Eq. (A.3) with $w_{t,s}$ and summing over S will give

$$E_t[R_{t+1,M}] = R_f + \gamma \cdot \mathbf{p_t}^T \cdot \mathbf{s_t} \cdot Var_t(R_{t+1,M}) - \sum_{k=1}^{S} w_{t,k}\left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,k}}{p_{t,k}}\right) \tag{A.4}$$

Obviously, $E_t[R_{t+1,M}] - R_f + \sum_{k=1}^{S} w_{t,k}\left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,k}}{p_{t,k}}\right) = \gamma \cdot \mathbf{p_t}^T \cdot \mathbf{s_t} \cdot Var_t(R_{t+1,M})$. Plug this result into Eq. (A.3), we come up with the result in Eq. (7) $E_t[R_{t+1,M}] = R_f + \beta_{t,s}\left[\cdot E_t\left[R_{t+1,M} - R_f\right] + \sum_{k=1}^{S} w_{t,k}\left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,k}}{p_{t,k}}\right)\right] - \sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}}$. □

Appendix B. Proof of Proposition 2

At the time t, the price $t-1$ is known and treated as constant. We also have $R_{t,s} = \frac{p_{t,s} + d_{t,s}}{p_{t-1,s}}$. **First**, we prove that $\frac{\partial R_{t,s}}{\partial\left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}}\right)} > 0$ and $\frac{\partial p_{t,s}}{\partial\left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}}\right)} > 0$. We have

$$\frac{\partial R_{t,s}}{\partial\left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}}\right)} = \frac{\partial R_{t,s}}{\partial\left(\sum_i \frac{\gamma}{\gamma_i} \cdot b_{i,t+1,s}\right)} \cdot \frac{\partial\left(\sum_i \frac{\gamma}{\gamma_i} \cdot b_{i,t+1,s}\right)}{\partial\left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}}\right)}$$

$$= \frac{1}{p_{t-1,s}} \cdot \frac{\partial[p_{t,s} + d_{t,s}]}{\partial\left(\sum_i \frac{\gamma}{\gamma_i} \cdot b_{i,t+1,s}\right)} \cdot p_{t,s}$$

Taking $d_{s,t}$ as exogenous, constant, and positive, then we will have

$$sgn\left(\frac{\partial R_{t,s}}{\partial\left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}}\right)}\right) = sgn\left(\frac{\partial p_{t,s}}{\partial\left(\sum_i \frac{\gamma}{\gamma_i} \cdot b_{i,t+1,s}\right)}\right) \tag{B.1}$$

We know that $\frac{\partial p_{t,s}}{\partial\left(\sum_i \frac{\gamma}{\gamma_i} \cdot b_{i,t+1,s}\right)} > 0$, following Eq. (A.1). Now, applying Eq. (B.1), we will have $\frac{\partial R_{s,t}}{\partial\left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}}\right)} > 0$. □

With $R_{t,s} = \frac{p_{t,s} + d_{t,s}}{p_{t-1,s}}$, then $sgn\left[\frac{\partial R_{s,t}}{\partial\left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}}\right)}\right] = sgn\left[\frac{\partial p_{t,s}}{\partial\left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}}\right)}\right]$. Therefore, $\frac{\partial p_{t,s}}{\partial\left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}}\right)} > 0$. □

Second, we prove

$$sgn\left(\frac{\partial R_{t,s}}{\partial \sum_{k=1}^{S} w_{t,k}\left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,k}}{p_{t,k}}\right)}\right) = sgn\left(\frac{\partial p_{t,s}}{\partial \sum_{k=1}^{S} w_{t,k}\left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,k}}{p_{t,k}}\right)}\right) = -sgn(\beta_{t,s})$$

In this proof, we keep the bias of the market about the stock s return $\left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,s}}{p_{t,s}}\right)$ unchanged. So the change in $\sum_{k=1}^{S} w_{t,k}\left(\sum_i \frac{\gamma}{\gamma_i} \cdot \frac{b_{i,t+1,k}}{p_{t,k}}\right)$ is

Table D.1
Name of characteristics and their descriptions.

Name	Reference	Label	Date, Journal	Definition
absacc	Bandyopadhyay, Huang, and Wirjanto (2010)	Absolute accruals	2010, WP	Absolute value of acc
acc	Sloan (1996)	Working capital accruals	1996, TAR	Annual income before extraordinary items (ib) minus operating cash flows (oancf) divided by average total assets (at); if oancf is missing then set to change in act - change in che - change in lct+ change in dlc+change in txp-dp
aeavol	Lerman, Livnat, and Mendenhall (2011)	Abnormal earnings announcement volume	2007, WP	Average daily trading volume (vol) for 3 days around earnings announcement minus average daily volume for 1-month ending 2 weeks before earnings announcement divided by 1-month average daily volume. Earnings announcement day from Compustat quarterly (rdq)
age	Jiang, Lee, and Zhang (2005)	Years since first Compustat coverage	2005, RAS	Number of years since first Compustat coverage
agr	Cooper, Gulen, and Schill (2008)	Asset growth	2008, JF	Annual percent change in total assets (at)
baspread	Amihud and Mendelson (1989)	Bid–ask spread	1989, JF	Monthly average of daily bid–ask spread divided by average of daily spread
beta	Fama and MacBeth (1973)	Beta	1973, JPE	Estimated market beta from weekly returns and equal weighted market returns for 3 years ending month t-1 with at least 52 weeks of returns
betasq	Fama and MacBeth (1973)	Beta squared	1973, JPE	Market beta squared
bm	Rosenberg, Reid, and Lanstein (1985)	Book-to-market	1985, JPM	Book value of equity (ceq) divided by end of fiscal year-end market capitalization
bm_ia	Asness, Porter, and Stevens (2000)	Industry-adjusted book to market	2000, WP	Industry adjusted book-to-market ratio
cash	Palazzo (2012)	Cash holdings	2012, JFE	Cash and cash equivalents divided by average total assets
cashdebt	Ou and Penman (1989)	Cash flow to debt	1989, JAE	Earnings before depreciation and extraordinary items (ib+dp) divided by avg. total liabilities (lt)
cashpr	Chandrashekar and Rao (2009)	Cash productivity	2009, WP	Fiscal year-end market capitalization plus long-term debt (dltt) minus total assets (at) divided by cash and equivalents (che)
cfp	Desai, Rajgopal, and Venkatachalam (2004)	Cash-flow-to-price ratio	2004, TAR	Operating cash flows divided by fiscal-year-end market capitalization
cfp_ia	Asness et al. (2000)	Industry-adjusted cash-flow-to-price ratio	2000, WP	Industry adjusted cfp
chatoia	Soliman (2008)	Industry-adjusted change in asset turnover	2008, TAR	2-digit SIC - fiscal-year mean-adjusted change in sales (sale) divided by average total assets (at)
chcsho	Pontiff and Woodgate (2008)	Change in shares outstanding	2008, JF	Annual percent change in shares outstanding (csho)
chempia	Asness et al. (2000)	Industry-adjusted change in employees	200, WP	Industry-adjusted change in number of employees
chfeps	Hawkins, Chamberlin, and Daniel (1984)	Change in forecasted EPS	1984, FAJ	Mean analyst forecast in month prior to fiscal period end date from I/B/E/S summary file minus same mean forecast for prior fiscal period using annual earnings forecasts
chinv	Thomas and Zhang (2002)	Change in inventory	2002, RAS	Change in inventory (inv) scaled by average total assets (at)
chmom	Gettleman and Marks (2006)	Change in 6-month momentum	2006, WP	Cumulative returns from months t-6 to t-1 minus months t-12 to t-7
chnanalyst	Scherbina (2007)	Change in number of analysts	2008 RF	Change in nanalyst from month t-3 to month t
chpmia	Soliman (2008)	Industry-adjusted change in profit margin	2008, TAR	2-digit SIC - fiscal-year mean adjusted change in income before extraordinary items (ib) divided by sales (sale)
chtx	Thomas and Zhang (2011)	Change in tax expense	2011, JAR	Percent change in total taxes (txtq) from quarter t-4 to t
cinvest	Titman, Wei, and Xie (2004)	Corporate investment	2004, JFQA	Change over one quarter in net PP&E (ppentq) divided by sales (saleq) - average of this variable for prior 3 quarters; if saleq== 0, then scale by 0.01

(continued on next page)

Table D.1 (*continued*).

convind	Valta (2016)	Convertible debt indicator	2016, JFQA	An indicator equal to 1 if company has convertible debt obligations
currat	Ou and Penman (1989)	Current ratio	1989, JAE	Current assets/current liabilities
depr	Holthausen and Larcker (1992)	Depreciation/PP&E	1992, JAE	Depreciation divided by PP&E
disp	Diether, Malloy, and Scherbina (2002)	Dispersion in forecasted EPS	2002, JF	Standard deviation of analyst forecasts in the month prior to fiscal period end date divided by the absolute value of the mean forecast; if meanest = 0, then scalar set to 1. Forecast data from I/B/E/S summary files
divi	Michaely, Thaler, and Womack (1995)	Dividend initiation	1995, JF	An indicator variable equal to 1 if the company pays dividends but did not in prior year
divo	Michaely et al. (1995)	Dividend omission	1995, JF	An indicator variable equal to 1 if the company does not pay dividends but did in prior year
dolvol	Chordia, Subrahmanyam, and Anshuman (2001)	Dollar trading volume	2001, JFE	Natural log of trading volume times price per share from month t-2
dy	Litzenberger and Ramaswamy (1982)	Dividend to price	1982, JF	Total dividends (dvt) divided by market capitalization at fiscal year-end
ear	Brandt, Kishore, Santa-Clara, and Venkatachalam (2008)	Earnings announcement return	2008, WP	Sum of daily returns in three days around an earnings announcement. Earnings announcement from Compustat quarterly file (rdq)
egr	Richardson, Sloan, Soliman, and Tuna (2005)	Growth in common shareholder equity	2005, JAE	Annual percent change in book value of equity (ceq)
ep	Basu (1977)	Earnings to price	1977, JF	Annual income before extraordinary items (ib) divided by end of fiscal year market cap
fgr5yr	Bauman and Dowen (1988)	Forecasted growth in 5-year EPS	1988, FAJ	Most recently available analyst forecasted 5-year growth
gma	Novy-Marx (2013)	Gross profitability	2013, JFE	Revenues (revt) minus cost of goods sold (cogs) divided by lagged total assets (at)
grCAPX	Anderson and Garcia-Feijóo (2006)	Growth in capital expenditures	2006, JF	Percent change in capital expenditures from year t-2 to year t
grltnoa	Fairfield, Whisenant, and Yohn (2003)	Growth in long-term net operating assets	2003, TAR	Growth in long-term net operating assets
herf	Hou and Robinson (2006)	Industry sales concentration	2006, JF	2-digit SIC - fiscal-year sales concentration (sum of squared percent of sales in the industry for each company).
hire	Belo, Lin, and Bazdresch (2014)	Employee growth rate	2014, JPE	Percent change in number of employees (emp)
idiovol	Ali, Hwang, and Trombley (2003)	Idiosyncratic return volatility	2003, JFE	Standard deviation of residuals of weekly returns on weekly equal weighted market returns for 3 years prior to month end
ill	Amihud (2002)	Illiquidity	2002, JFM	Average of daily (absolute return/dollar volume).
indmom	Moskowitz and Grinblatt (1999)	Industry momentum	1999, JF	Equal weighted average industry 12-month returns
invest	Chen and Zhang (2010)	Capital expenditures and inventory	2010, JF	Annual change in gross property, plant, and equipment (ppegt) + annual change in inventories (invt) all scaled by lagged total assets (at)
IPO	Loughran and Ritter (1995)	New equity issue	1995, JF	An indicator variable equal to 1 if first year available on CRSP monthly stock file
lev	Bhandari (1988)	Leverage	1988, JF	Total liabilities (lt) divided by fiscal year-end market capitalization
lgr	Richardson et al. (2005)	Growth in long-term debt	2005, JAE	Annual percent change in total liabilities (lt)
maxret	Bali et al. (2011)	Maximum daily return	2011, JFE	Maximum daily return from returns during calendar month t-1
mom12 m	Jegadeesh (1990)	12-month momentum	1990, JF	11-month cumulative returns ending one month before month end
mom1m	Jegadeesh and Titman (1993)	1-month momentum	1993, JF	1-month cumulative return
mom36m	Jegadeesh and Titman (1993)	36-month momentum	1993, JF	Cumulative returns from months t-36 to t-13
mom6m	Jegadeesh and Titman (1993)	6-month momentum	1993, JF	5-month cumulative returns ending one month before month end
ms	Mohanram (2005)	Financial statement score	2005, RAS	Sum of 8 indicator variables for fundamental performance
mve	Banz (1981)	Size	1981, JFE	Natural log of market capitalization at end of month t-1

(*continued on next page*)

Table D.1 (*continued*).

mve_ia	Asness et al. (2000)	Industry-adjusted size	2000, WP	2-digit SIC industry-adjusted fiscal year-end market capitalization
nanalyst	Elgers, Lo, and Pfeiffer (2001)	Number of analysts covering stock	2001, TAR	Number of analyst forecasts from most recently available I/B/E/S summary files in the month prior to month of portfolio formation. nanalyst set to zero if not covered in I/B/E/S summary file
nincr	Barth, Elliott, and Finn (1999)	Number of earnings increases	1999, JAR	Number of consecutive quarters (up to eight quarters) with an increase in earnings (ibq) over the same quarter in the prior year
operprof	Fama and French (2015)	Operating profitability	2015, JFE	Revenue minus the cost of goods sold - SG&A expense - interest expense divided by lagged common shareholders' equity
orgcap	Eisfeldt and Papanikolaou (2013)	Organizational capital	2013, JF	Capitalized SG&A expenses
pchcapx_ia	Abarbanell and Bushee (1998)	Industry adjusted % change in capital expenditures	1998, TAR	2-digit SIC - fiscal-year mean-adjusted percent change in capital expenditures (capx)
pchcurrat	Ou and Penman (1989)	% change in current ratio	1989, JAE	Percent change in currat.
pchdepr	Holthausen and Larcker (1992)	% change in depreciation	1992, JAE	Percent change in depr
pchgm_pchsale	Abarbanell and Bushee (1998)	% change in gross margin - % change in sales	1998, TAR	Percent change in gross margin (sale-cogs) minus percent change in sales (sale)
pchquick	Ou and Penman (1989)	% change in quick ratio	1989, JAE	Percent change in quick
pchsale_pchinvt	Abarbanell and Bushee (1998)	% change in sales - % change in inventory	1998, TAR	Annual percent change in sales (sale) minus annual percent change in inventory (invt).
pchsale_pchrect	Abarbanell and Bushee (1998)	% change in sales -% change in A/R	1998, TAR	Annual percent change in sales (sale) minus annual percent change in receivables (rect)
pchsale_pchxsga	Abarbanell and Bushee (1998)	% change in sales - % change in SG&A	1998, TAR	Annual percent change in sales (sale) minus annual percent change in SG&A (xsga)
pchsaleinv	Ou and Penman (1989)	% change sales-to-inventory	1989, JAE	Percent change in saleinv
pctacc	Hafzalla, Lundholm, and Matthew Van Winkle (2011)	Percent accruals	2011, TAR	Same as acc except that the numerator is divided by the absolute value of ib; if ib= 0 then ib set to 0.01 for denominator
pricedelay	Hou and Moskowitz (2005)	Price delay	2005, RFS	The proportion of variation in weekly returns for 36 months ending in month explained by 4 lags of weekly market returns incremental to contemporaneous market return
ps	Piotroski et al. (2000)	Financial statements score	2000, JAR	Sum of 9 indicator variables to form fundamental health score
quick	Ou and Penman (1989)	Quick ratio	1989, JAE	(current assets - inventory)/current liabilities
rd	Eberhart, Maxwell, and Siddique (2004)	R&D increase	2004, JF	An indicator variable equal to 1 if R&D expense as a percentage of total assets has an increase greater than 5%.
rd_mve	Guo, Lev, and Shi (2006)	R&D to market capitalization	2006, JBFA	R&D expense divided by end-of-fiscal-year market capitalization
rd_sale	Guo et al. (2006)	R&D to sales	2006, JBFA	R&D expense divided by sales (xrd/sale)
realestate	Tuzel (2010)	Real estate holdings	2010, RFS	Buildings and capitalized leases divided by gross PP&E
retvol	Ang, Hodrick, Xing, and Zhang (2006)	Return volatility	2006, JF	Standard deviation of daily returns from month t-1
roaq	Balakrishnan, Bartov, and Faurel (2010)	Return on assets	2010, JAE	Income before extraordinary items (ibq) divided by one quarter lagged total assets (atq)
roavol	Francis, LaFond, Olsson, and Schipper (2004)	Earnings volatility	2004, TAR	Standard deviation for 16 quarters of income before extraordinary items (ibq) divided by average total assets (atq)
roeq	Hou et al. (2015)	Return on equity	2015 RFS	Earnings before extraordinary items divided by lagged common shareholders' equity
roic	Brown and Rowe (2007)	Return on invested capital	2007, WP	Annual earnings before interest and taxes (ebit) minus nonoperating income (nopi) divided by non-cash enterprise value (ceq+lt-che)
rsup	Kama (2009)	Revenue surprise	2009, JBFA	Sales from quarter t minus sales from quarter t-4 (saleq) divided by fiscal-quarter-end market capitalization (cshoq * prccq)
salecash	Ou and Penman (1989)	Sales to cash	1989, JAE	Annual sales divided by cash and cash equivalents

(*continued on next page*)

Table D.1 (*continued*).

saleinv	Ou and Penman (1989)	Sales to inventory	1989, JAE	Annual sales divided by total inventory
salerec	Ou and Penman (1989)	Sales to receivables	1989, JAE	Annual sales divided by accounts receivable
secured	Valta (2016)	Secured debt	2016, JFQA	Total liability scaled secured debt
securedind	Valta (2016)	Secured debt indicator	2016, JFQA	An indicator equal to 1 if company has secured debt obligations
sfe	Elgers et al. (2001)	Scaled earnings forecast	2001, TAR	Analysts mean annual earnings forecast for the nearest upcoming fiscal year from most recent month available prior to the month of portfolio formation from I/B/E/S summary files scaled by price per share at fiscal quarter end
sgr	Lakonishok, Shleifer, and Vishny (1994)	Sales growth	1994, JF	Annual percent change in sales (sale)
sin	Hong and Kacperczyk (2009)	Sin stocks	2009, JFE	An indicator variable equal to 1 if a company's primary industry classification is in smoke or tobacco, beer or alcohol, or gaming
SP	Barbee, Mukherji, and Raines (1996)	Sales to price	1996, FAJ	Annual revenue (sale) divided by fiscal year-end market capitalization
std_dolvol	Chordia, Roll, and Subrahmanyam (2001)	Volatility of liquidity (dollar trading volume)	2001, JFE	Monthly standard deviation of daily dollar trading volume
std_turn	Chordia, Roll, and Subrahmanyam (2001)	Volatility of liquidity (share turnover)	2001, JFE	Monthly standard deviation of daily share turnover
stdacc	Bandyopadhyay et al. (2010)	Accrual volatility	2010, WP	Standard deviation for 16 quarters of accruals (measured with quarterly Compustat) scaled by sales; if saleq= 0, then scale by 0.01
stdcf	Huang (2009)	Cash flow volatility	2009, JEF	Standard deviation for 16 quarters of cash flows divided by sales (saleq); if saleq= 0, then scale by 0.01. Cash flows are defined as ibq minus quarterly accruals
sue	Rendleman, Jones, and Latané (1982)	Unexpected quarterly earnings	1982, JFE	Unexpected quarterly earnings divided by fiscal-quarter-end market cap. Unexpected earnings are I/B/E/S actual earnings minus median forecasted earnings if available, else it is the seasonally differenced quarterly earnings before extraordinary items from Compustat quarterly file
tang	Almeida and Campello (2007)	Debt capacity/firm tangibility	2007, RFS	Cash holdings + 0.715 ×receivables + 0.547 ×inventory + 0.535 ×PPE/ total assets
tb	Lev and Nissim (2004)	Tax income to book income	2004, TAR	Tax income, calculated from current tax expense divided by the maximum federal tax rate, divided by income before extraordinary items
turn	Datar, Naik, and Radcliffe (1998)	Share turnover	1998, JFM	Average monthly trading volume for most recent 3 months scaled by number shares outstanding in current month
zerotrade	Liu (2006)	Zero trading days	2006, JFE	Turnover weighted number of zero trading days for most

driven by other stocks than s. Subsequently,

$$\frac{\partial p_{s,t}}{\partial \sum_{k=1}^{S} w_{t,k}\left(\sum_i \frac{\gamma}{\gamma_i}\cdot\frac{b_{i,t+1,k}}{p_{t,k}}\right)} = \frac{\partial p_{t,s}}{\partial E_t[R_{t+1,s}]}\cdot\frac{\partial E_t[R_{t+1,s}]}{\partial \sum_{k=1}^{S} w_{k,t}\left(\sum_i \frac{\gamma}{\gamma_i}\cdot\frac{b_{i,t+1,k}}{p_{t,k}}\right)}.$$

$$= \frac{\partial p_{t,s}}{\partial E[R_{t+1,s}]}\cdot \beta_{t,s} \tag{B.2}$$

With $\frac{\partial p_{t,s}}{\partial E_t[R_{t+1,s}]} < 0$, then, $sgn\left[\frac{\partial p_{t,s}}{\partial \sum_{k=1}^{S} w_{t,k}\left(\sum_i \frac{\gamma}{\gamma_i}\cdot\frac{b_{i,t+1,k}}{p_{t,k}}\right)}\right] = -sgn(\beta_{t,s})$.

Since we have

$$\frac{\partial R_{t,s}}{\partial \sum_{k=1}^{S} w_{t,k}\left(\sum_i \frac{\gamma}{\gamma_i}\cdot\frac{b_{i,t+1,k}}{p_{t,k}}\right)} = \frac{\partial R_{t,s}}{\partial p_{t,s}}\cdot\frac{\partial p_{t,s}}{\partial \sum_{k=1}^{S} w_{k,t}\left(\sum_i \frac{\gamma}{\gamma_i}\cdot\frac{b_{i,t+1,k}}{p_{t,k}}\right)}$$

and $\frac{\partial R_{t,s}}{\partial p_{t,s}} > 0$, then, $sgn\left[\frac{\partial R_{t,s}}{\partial \sum_{k=1}^{S} w_{k,t}\left(\sum_i \frac{\gamma}{\gamma_i}\cdot\frac{b_{i,t+1,k}}{p_{t,k}}\right)}\right] = -sgn(\beta_{t,s})$ □

Appendix C. Proof of Proposition 3

Given that $\boldsymbol{\theta}_{t+1} = \mathbf{Z_t}\cdot\mathbf{f_{t+1}}$, we can rewrite Eq. (9) as

$$E_t[\mathbf{r_{t+1}}] = r_f\cdot\mathbf{1} + \boldsymbol{\beta}_t\cdot E_t\left[r_{t+1,M} - r_f\right] + \boldsymbol{\beta}_t\cdot\left[\mathbf{w_t}^T\cdot\mathbf{Z_t}\cdot\mathbf{f_{t+1}}\right] - \mathbf{Z_t}\cdot\mathbf{f_{t+1}}$$

$$E_t[\mathbf{r_{t+1}}] - \mathbf{1}\cdot rf - \boldsymbol{\beta}_t\cdot E_t\left[r_{t+1,M} - r_f\right] = [\boldsymbol{\beta}_t\cdot\mathbf{w_t^T} - \mathbf{I}]\cdot\mathbf{Z_t}\cdot\mathbf{f_{t+1}}$$

With $E_t[\mathbf{r_{t+1}}]$ is the unbiased expectation of $\mathbf{r_{t+1}}$, then we can set up a regression as

$$\mathbf{r_{t+1}} - 1\cdot rf - \boldsymbol{\beta}_t\cdot\left[r_{t+1,M} - r_f\right] = [\boldsymbol{\beta}_t\cdot\mathbf{w_t^T} - \mathbf{I}]\cdot\mathbf{Z_t}\cdot\mathbf{f_{t+1}} + \boldsymbol{\upsilon}_{t+1}$$

$$\mathbf{y_{t+1}} = \mathbf{X_t}\cdot\mathbf{f_{t+1}} + \boldsymbol{\upsilon}_{t+1}$$

where $\boldsymbol{\upsilon}_{t+1}$ is a residual vector. The vector $\mathbf{f_{t+1}}$ in the above equation can be estimated by a standard least square estimator. □

Appendix D. Variables list

See Table D.1.

References

Abarbanell, J. S., & Bushee, B. J. (1998). Abnormal returns to a fundamental analysis strategy. *The Accounting Review*, *73*(1), 19–45.

Al-Nasseri, A., Menla Ali, F., & Tucker, A. (2021). Investor sentiment and the dispersion of stock returns: Evidence based on the social network of investors. *International Review of Financial Analysis*, *78*, Article 101910, URL https://linkinghub.elsevier.com/retrieve/pii/S1057521921002362.

Ali, A., Hwang, L.-S., & Trombley, M. A. (2003). Arbitrage risk and the book-to-market anomaly. *Journal of Financial Economics*, *69*(2), 355–373.

Almeida, H., & Campello, M. (2007). Financial constraints, asset tangibility, and corporate investment. *The Review of Financial Studies*, *20*(5), 1429–1460.

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31–56.

Amihud, Y., & Mendelson, H. (1989). The effects of beta, bid-ask spread, residual risk, and size on stock returns. *The Journal of Finance*, *44*(2), 479–486.

Anderson, C. W., & Garcia-Feijóo, L. (2006). Empirical evidence on capital investment, growth options, and security returns. *The Journal of Finance*, *61*(1), 171–194.

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. *The Journal of Finance*, *61*(1), 259–299.

Asness, C. S., Porter, R. B., & Stevens, R. L. (2000). Predicting stock returns using industry-relative firm characteristics. Available at SSRN 213872.

Avramov, D., Chordia, T., Jostova, G., & Philipov, A. (2019). Bonds, stocks, and sources of mispricing. In *George Mason University School of Business research paper, no. 18-5*.

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance*, *61*(4), 1645–1680.

Baker, M., Wurgler, J., & Yuan, Y. (2012). Global, local, and contagious investor sentiment. *Journal of Financial Economics*, *104*(2), 272–287.

Balakrishnan, K., Bartov, E., & Faurel, L. (2010). Post loss/profit announcement drift. *Journal of Accounting and Economics*, *50*(1), 20–41.

Bali, T. G., Cakici, N., & Whitelaw, R. F. (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics*, *99*(2), 427–446.

Bandyopadhyay, S. P., Huang, A. G., & Wirjanto, T. S. (2010). The accrual volatility anomaly. University of Waterloo, Unpublished manuscript.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*(1), 3–18.

Barbee, W. C., Jr., Mukherji, S., & Raines, G. A. (1996). Do sales–price and debt–equity explain stock returns better than book–market and firm size? *Financial Analysts Journal*, *52*(2), 56–60.

Barth, M. E., Elliott, J. A., & Finn, M. W. (1999). Market rewards associated with patterns of increasing earnings. *Journal of Accounting Research*, *37*(2), 387–413.

Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. *The Journal of Finance*, *32*(3), 663–682.

Bauman, W. S., & Dowen, R. (1988). Growth projections and commin stock returns. *Financial Analysts Journal*, *44*(4), 79.

Belo, F., Lin, X., & Bazdresch, S. (2014). Labor hiring, investment, and stock return predictability in the cross section. *Journal of Political Economy*, *122*(1), 129–177.

Benhabib, J., Liu, X., & Wang, P. (2016). Sentiments, financial markets, and macroeconomic fluctuations. *Journal of Financial Economics*, *120*(2), 420–443.

Bhandari, L. C. (1988). Debt/equity ratio and expected common stock returns: Empirical evidence. *The Journal of Finance*, *43*(2), 507–528.

Björk, T. (2009). *Arbitrage Theory in Continuous Time*. Oxford University Press.

Brandt, M. W., Kishore, R., Santa-Clara, P., & Venkatachalam, M. (2008). Earnings announcements are full of surprises. *SSRN eLibrary*.

Brennan, M. J., Chordia, T., & Subrahmanyam, A. (1998). Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal of Financial Economics*, *49*(3), 345–373.

Brown, D. P., & Rowe, B. (2007). The productivity premium in equity returns. Available at SSRN 993467.

Chandrashekar, S., & Rao, R. K. (2009). The productivity of corporate cash holdings and the cross-section of expected stock returns. In *McCombs research paper series no. FIN-03-09*.

Chen, L., Pelger, M., & Zhu, J. (2019). Deep learning in asset pricing. *SSRN Electronic Journal*.

Chen, L., & Zhang, L. (2010). A better three-factor model that explains more anomalies. *The Journal of Finance*, *65*(2), 563–595.

Chordia, T., Goyal, A., & Shanken, J. A. (2015). Cross-sectional asset pricing with individual stocks: betas versus characteristics. Available at SSRN 2549578.

Chordia, T., Roll, R., & Subrahmanyam, A. (2001). Market liquidity and trading activity. *The Journal of Finance*, *56*(2), 501–530.

Chordia, T., Subrahmanyam, A., & Anshuman, V. R. (2001). Trading activity and expected stock returns. *Journal of Financial Economics*, *59*(1), 3–32.

Cochrane, J. H. (2009). *Asset Pricing: Revised Edition*. Princeton University Press.

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

Cortés, K., Duchin, R., & Sosyura, D. (2016). Clouded judgment: The role of sentiment in credit origination. *Journal of Financial Economics*, *121*(2), 392–413.

Daniel, K. D., Hirshleifer, D., & Subrahmanyam, A. (2001). Overconfidence, arbitrage, and equilibrium asset pricing. *The Journal of Finance*, *56*(3), 921–965.

Daniel, K., & Titman, S. (1997). Evidence on the characteristics of cross sectional variation in stock returns. *The Journal of Finance*, *52*(1), 1–33.

Datar, V. T., Naik, N. Y., & Radcliffe, R. (1998). Liquidity and stock returns: An alternative test. *Journal of Financial Markets*, *1*(2), 203–219.

Desai, H., Rajgopal, S., & Venkatachalam, M. (2004). Value-glamour and accruals mispricing: One anomaly or two? *The Accounting Review*, *79*(2), 355–385.

Diether, K. B., Malloy, C. J., & Scherbina, A. (2002). Differences of opinion and the cross section of stock returns. *The Journal of Finance*, *57*(5), 2113–2141.

Dong, H., & Gil-Bazo, J. (2020). Sentiment stocks. *International Review of Financial Analysis*, *72*, Article 101573.

Eberhart, A. C., Maxwell, W. F., & Siddique, A. R. (2004). An examination of long-term abnormal stock returns and operating performance following R&D increases. *The Journal of Finance*, *59*(2), 623–650.

Edmans, A., Fernandez-Perez, A., Garel, A., & Indriawan, I. (2022). Music sentiment and stock returns around the world. *Journal of Financial Economics*, *145*(2), 234–254.

Eisfeldt, A. L., & Papanikolaou, D. (2013). Organization capital and the cross-section of expected returns. *The Journal of Finance*, *68*(4), 1365–1406.

Elgers, P. T., Lo, M. H., & Pfeiffer, R. J., Jr. (2001). Delayed security price adjustments to financial analysts' forecasts of annual earnings. *The Accounting Review*, *76*(4), 613–632.

Fairfield, P. M., Whisenant, J. S., & Yohn, T. L. (2003). Accrued earnings and growth: Implications for future profitability and market mispricing. *The Accounting Review*, *78*(1), 353–371.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, *81*(3), 607–636.

Fang, H., Chung, C. P., Lu, Y. C., Lee, Y. H., & Wang, W. H. (2021). The impacts of investors' sentiments on stock returns using fintech approaches. *International Review of Financial Analysis*, *77*, Article 101858.

Feng, G., He, J., & Polson, N. G. (2018). Deep learning for predicting asset returns. arXiv preprint arXiv:1805.01104.

Francis, J., LaFond, R., Olsson, P. M., & Schipper, K. (2004). Costs of equity and earnings attributes. *The Accounting Review*, *79*(4), 967–1010.

Frazzini, A., & Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, *111*(1), 1–25.

Gettleman, E., & Marks, J. M. (2006). Acceleration strategies. *SSRN Electronic Journal*.

Gong, X., Zhang, W., Wang, J., & Wang, C. (2022). Investor sentiment and stock volatility: New evidence. *International Review of Financial Analysis*, *80*, Article 102028.

Green, J., Hand, J. R. M., & Zhang, X. F. (2017). The characteristics that provide independent information about average U.S. monthly stock returns. *The Review of Financial Studies*, *30*(12), 4389–4436.

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, *33*(5), 2223–2273.

Guo, R.-J., Lev, B., & Shi, C. (2006). Explaining the short- and long-term IPO anomalies in the US by R&D. *Journal of Business Finance & Accounting*, *33*(3–4), 550–579.

Hafzalla, N., Lundholm, R., & Matthew Van Winkle, E. (2011). Percent accruals. *The Accounting Review*, *86*(1), 209–236.

Harvey, C. R. (2017). Presidential address: The scientific outlook in financial economics. *The Journal of Finance*, *72*(4), 1399–1440.

Harvey, C. R., Liu, Y., & Zhu, H. (2016). …and the cross-section of expected returns. *The Review of Financial Studies*, *29*(1), 5–68.

Hawkins, E. H., Chamberlin, S. C., & Daniel, W. E. (1984). Earnings expectations and security prices. *Financial Analysts Journal*, *40*(5), 24–38.

Holthausen, R. W., & Larcker, D. F. (1992). The prediction of stock returns using financial statement information. *Journal of Accounting and Economics*, *15*(2–3), 373–411.

Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, *93*(1), 15–36.

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.

Hou, K., & Robinson, D. T. (2006). Industry concentration and average stock returns. *The Journal of Finance*, *61*(4), 1927–1956.

Hou, K., Xue, C., & Zhang, L. (2015). Digesting anomalies: An investment approach. *The Review of Financial Studies*, *28*(3), 650–705.

Hou, K., Xue, C., & Zhang, L. (2020). Replicating anomalies. *The Review of Financial Studies*, *33*(5), 2019–2133.

Huang, A. G. (2009). The cross section of cashflow volatility and expected stock returns. *Journal of Empirical Finance*, *16*(3), 409–429.

Islam, M. A. (2021). Investor sentiment in the equity market and investments in corporate-bond funds. *International Review of Financial Analysis*, *78*, Article 101898.

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *The Journal of Finance*, *45*(3), 881–898.

Jegadeesh, N., Noh, J., Pukthuanthong, K., Roll, R., & Wang, J. (2019). Empirical tests of asset pricing models with individual assets: Resolving the errors-in-variables bias in risk premium estimation. *Journal of Financial Economics*, *133*(2), 273–298.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), 65–91.

Jiang, F., Lee, J., Martin, X., & Zhou, G. (2019). Manager sentiment and stock returns. *Journal of Financial Economics*, *132*(1), 126–149.

Jiang, G., Lee, C. M., & Zhang, Y. (2005). Information uncertainty and expected returns. *Review of Accounting Studies*, *10*(2–3), 185–221.

Kama, I. (2009). On the market reaction to revenue and earnings surprises. *Journal of Business Finance & Accounting*, *36*(1–2), 31–50.

Kelly, B. T., Pruitt, S., & Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*.

Kim, K., Ryu, D., & Yang, H. (2021). Information uncertainty, investor sentiment, and analyst reports. *International Review of Financial Analysis*, *77*, Article 101835.

Kozak, S. (2019). Kernel trick for the cross section. *SSRN Electronic Journal*.

Kozak, S., Nagel, S., & Santosh, S. (2018). Interpreting factor models. *The Journal of Finance*, *73*(3), 1183–1223.
Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. *The Journal of Finance*, *49*(5), 1541–1578.
Lerman, A., Livnat, J., & Mendenhall, R. R. (2011). The high-volume return premium and post-earnings announcement drift. *SSRN Electronic Journal*.
Lev, B., & Nissim, D. (2004). Taxable income, future earnings, and equity values. *The Accounting Review*, *79*(4), 1039–1074.
Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, *47*(1), 13–37.
Litzenberger, R. H., & Ramaswamy, K. (1982). The effects of dividends on common stock prices tax effects or information effects? *The Journal of Finance*, *37*(2), 429–443.
Liu, W. (2006). A liquidity-augmented capital asset pricing model. *Journal of Financial Economics*, *82*(3), 631–671.
Liu, J., Stambaugh, R. F., & Yuan, Y. (2018). Absolving beta of volatility's effects. *Journal of Financial Economics*, *128*(1), 1–15.
Loughran, T., & Ritter, J. R. (1995). The new issues puzzle. *The Journal of Finance*, *50*(1), 23–51.
Michaely, R., Thaler, R. H., & Womack, K. L. (1995). Price reactions to dividend initiations and omissions: Overreaction or drift? *The Journal of Finance*, *50*(2), 573–608.
Mohanram, P. S. (2005). Separating winners from losers among LowBook-to-market stocks using financial statement analysis. *Review of Accounting Studies*, *10*(2–3), 133–170.
Moskowitz, T. J., & Grinblatt, M. (1999). Do industries explain momentum? *The Journal of Finance*, *54*(4), 1249–1290.
Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica*, *34*(4), 768.
Newey, W. K., & West, K. D. (1987). Hypothesis testing with efficient method of moments estimation. *International Economic Review*, *28*(3), 777–787.
Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, *108*(1), 1–28.
Obaid, K., & Pukthuanthong, K. (2022). A picture is worth a thousand words: Measuring investor sentiment by combining machine learning and photos from news. *Journal of Financial Economics*, *144*(1), 273–297.
Ou, J. A., & Penman, S. H. (1989). Financial statement analysis and the prediction of stock returns. *Journal of Accounting and Economics*, *11*(4), 295–329.
Palazzo, B. (2012). Cash holdings, risk, and expected returns. *Journal of Financial Economics*, *104*(1), 162–185.
Piotroski, J. D., et al. (2000). Value investing: The use of historical financial statement information to separate winners from losers. *Journal of Accounting Research*, *38*, 1–52.
Pontiff, J., & Woodgate, A. (2008). Share issuance and cross-sectional returns. *The Journal of Finance*, *63*(2), 921–945.
Rendleman, R. J., Jones, C. P., & Latané, H. A. (1982). Empirical anomalies based on unexpected earnings and the importance of risk adjustments. *Journal of Financial Economics*, *10*(3), 269–287.
Richardson, S. A., Sloan, R. G., Soliman, M. T., & Tuna, İ. (2005). Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics*, *39*(3), 437–485.
Rosenberg, B., Reid, K., & Lanstein, R. (1985). Persuasive evidence of market inefficiency. *The Journal of Portfolio Management*, *11*(3), 9–16.
Scherbina, A. (2007). Suppressed negative information and future underperformance. *Review of Finance*, *12*(3), 533–565.
Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*(3), 425–442.
Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting Review*, *71*(3), 289–315.
Soliman, M. T. (2008). The use of DuPont analysis by market participants. *The Accounting Review*, *83*(3), 823–853.
Song, Z., & Yu, C. (2022). Investor sentiment indices based on k-step PLS algorithm: A group of powerful predictors of stock market returns. *International Review of Financial Analysis*, *83*, Article 102321.
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. *The Journal of Finance*, *70*(5), 1903–1948.
Stambaugh, R. F., & Yuan, Y. (2017). Mispricing factors. *The Review of Financial Studies*, *30*(4), 1270–1315.
Thomas, J. K., & Zhang, H. (2002). Inventory changes and future returns. *Review of Accounting Studies*, *7*(2–3), 163–187.
Thomas, J., & Zhang, F. X. (2011). Tax expense momentum. *Journal of Accounting Research*, *49*(3), 791–821.
Titman, S., Wei, K. C. J., & Xie, F. (2004). Capital investments and stock returns. *Journal of Financial and Quantitative Analysis*, *39*(4), 677–700.
Treynor, J. L. (1961). Market value, time, and risk. *Time, and Risk*.
Treynor, J. L. (1962). Toward a theory of market value of risky assets. Elsevier BV, Unpublished manuscript. "Rough Draft" dated By Mr. Treynor to the Fall of 1962.
Tuzel, S. (2010). Corporate real estate holdings and the cross-section of stock returns. *The Review of Financial Studies*, *23*(6), 2268–2302.
Valta, P. (2016). Strategic default, debt structure, and stock returns. *Journal of Financial and Quantitative Analysis*, *51*(1), 197–229.
Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). The MIT Press.
Wooldridge, J. M. (2015). *Introductory Econometrics: A Modern Approach*. Cengage learning.