



Contents lists available at ScienceDirect

## Journal of Financial Economics

journal homepage: [www.elsevier.com/locate/jfec](http://www.elsevier.com/locate/jfec)

# Extrapolative beliefs in the cross-section: What can we learn from the crowds?<sup>☆</sup>

Zhi Da<sup>a,\*</sup>, Xing Huang<sup>b</sup>, Lawrence J. Jin<sup>c</sup><sup>a</sup> 258 Mendoza College of Business, University of Notre Dame, Notre Dame, IN 46556, USA<sup>b</sup> Olin Business School, Washington University in St. Louis, Simon Hall, One Brookings Drive, St. Louis, MO 63130, USA<sup>c</sup> Division of the Humanities and Social Sciences, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA

## ARTICLE INFO

## Article history:

Received 28 October 2019

Revised 25 March 2020

Accepted 26 March 2020

Available online xxx

## JEL classification:

G02

G11

G12

## Keywords:

Return extrapolation

Beliefs in the cross-section

Expectation formation

## ABSTRACT

Using novel data from a crowdsourcing platform for ranking stocks, we investigate how investors form expectations about stock returns over the next week. We find that investors extrapolate from stocks' recent past returns, with more weight on more recent returns, especially when recent returns are negative, salient, or from a dispersed cross-section. Such extrapolative beliefs are stronger among nonprofessionals and large stocks. Moreover, consensus rankings negatively predict returns over the next week, more so among stocks with low institutional ownership and a high degree of extrapolation. A trading strategy that sorts stocks on investor beliefs generates an economically significant profit.

© 2020 Elsevier B.V. All rights reserved.

## 1. Introduction

A central question in finance is how investors form expectations about future asset returns. Recent work (Vissing-Jorgensen, 2004; Bacchetta et al., 2009; Amromin and Sharpe, 2013; Greenwood and Shleifer, 2014; Kuchler and Zafar, 2019) provides convincing evidence of return extrapolation, the notion that investors' expectations about an asset's future return are a positive function of the asset's recent past returns. Recent models of Barberis et al. (2015) and Jin and Sui (2019) show that return extrapolation helps explain facts about the aggregate stock market such as excess volatility and predictability of stock market returns.

Despite their intuitive theoretical appeal, extrapolation models have thus far been tested primarily with data on the aggregate stock market. There has been very little direct evidence on how investors form expectations about individual stock returns, whether these expectations are

<sup>☆</sup> We thank Bill Schwert (the editor), an anonymous referee, Nicholas Barberis, Colin Camerer, Julien Cujean, Peter Cziraki, Michael Ewens, Samuel Hartzmark, Burton Hollifield, Stephen Karolyi, Lisa Kramer, Juhani Linnainmaa, Yueran Ma, Abhiroop Mukherjee, Marina Niessner, Terrance Odean, Cameron Peng, Jesse Shapiro, David Solomon, Noah Stoffman, and seminar and conference participants at Aalto University, Baylor University, Boston College, Caltech, CKGSB, CUHK Shenzhen, CUNY Baruch, Florida International University, Georgetown University, SWUFE, UC Berkeley, UC Irvine, UCLA Anderson, UC Riverside, University of Iowa, University of Washington, WUSTL, AFA 2019, CICF 2018, EFA 2018, FIRS 2018, MFA 2019, the 15th Annual Conference in Financial Economics Research, the 2018 LA Finance Day Conference, the 2019 Mitsui Finance Symposium, the 2019 PKU-CCER Summer Institute, the 2019 SFS Cavalcade, the 2018 TAU Finance Conference, the 2019 Utah Winter Finance Conference, the WAPFIN conference at NYU Stern, and the Yale Junior Finance Conference for helpful comments and suggestions. We are grateful to Leigh Dorgon, Josh Dulberger, and Aram Balian for providing data from Forcerank.

\* Corresponding author.

E-mail addresses: [zda@nd.edu](mailto:zda@nd.edu) (Z. Da), [xing.huang@wustl.edu](mailto:xing.huang@wustl.edu) (X. Huang), [jin@caltech.edu](mailto:jin@caltech.edu) (L.J. Jin).

rational, and how they relate to subsequent returns.<sup>1</sup> In this paper, we provide some of the first direct evidence of investor expectations about individual stock returns.<sup>2</sup> We find that these expectations are positively related to recent past returns but are negatively related to subsequent returns, indicating that they are not fully rational. We show that our results are consistent with a theoretical framework in which investors with extrapolative beliefs interact with more rational investors. The cross-section dimension of our analysis generates new empirical facts about extrapolative beliefs and allows us to link such beliefs to cross-sectional asset pricing patterns.

We analyze a novel data set from Forcerank, a crowdsourcing platform for ranking stocks. In each contest on this platform, participants rank ten stocks based on their perceived future performance of these stocks over the course of the contest, which is usually one week. Compared to alternative data sources, Forcerank data have a number of unique advantages for studying investor beliefs. They contain precise rankings information with a clearly specified forecasting horizon for a predetermined set of stocks. Moreover, these rankings are solicited from a highly diverse and geographically distributed group of individuals in a blind setting that rules out herding or cross-learning.<sup>3</sup>

Taking advantage of the Forcerank data, we investigate how individuals form their expectations about future returns on individual stocks and how these expectations affect asset prices. We first estimate, across stocks, a linear regression of investor expectations on past stock returns; here we use the consensus Forcerank score—each individual stock's average ranking across all participants in that contest—as a proxy for the investor expectation of the stock return. We find that individuals extrapolate from a stock's recent past returns when forming expectations about its future return. Specifically, the regression coefficients on recent past returns are all positive and mostly significant. More important, the coefficients on distant past returns are in general lower than those on recent past

returns: quantitatively, returns four weeks earlier are only about 9% as important as returns in the most recent week. Not surprisingly, individuals seem to extrapolate only from the idiosyncratic—rather than the systematic—component of past returns.

Furthermore, this extrapolative pattern remains almost identical after controlling for past fundamental news, news sentiment, and risk measures. It is also robust to alternative regression specifications. As an external validation, we observe a very similar extrapolative pattern when we use the brokerage account data of Barber and Odean (2000) to examine the relation between initial buys from a large group of short-term retail traders and past stock returns. In other words, our findings do not seem to be driven by the specific contest setting of Forcerank.

To parsimoniously capture the extrapolative pattern, we further apply an exponential decay function as the weighting scheme on past returns.<sup>4</sup> In doing so, we summarize the *degree of extrapolation* from investor expectations with two parameters. The first parameter,  $\lambda_1$ —a scaling factor that multiplies all past returns—captures a “level” effect (i.e., the overall extent to which investor expectations respond to past returns). The second parameter,  $\lambda_2$ —the weight investors put on distant past returns relative to recent past returns when forming beliefs about future returns—captures a “slope” effect. Investors' degree of extrapolation is jointly determined by  $\lambda_1$  and  $\lambda_2$ : when  $\lambda_2$  is much lower than one, investor expectations are determined primarily by most recent past returns; at the same time, investor expectations exhibit a high degree of extrapolation only when  $\lambda_1$  is high. We find that  $\lambda_1$  estimated from the Forcerank expectations data is significantly positive and  $\lambda_2$  estimated from the Forcerank expectations data is significantly lower than one. Together, these results confirm that Forcerank participants have a strong degree of extrapolation.

Our cross-sectional data allow us to study the determinants of investors' extrapolative expectation as captured by  $\lambda_1$  and  $\lambda_2$ . We find that extrapolation is *asymmetric* between positive and negative past returns: investors put more weight (a larger  $\lambda_1$ ) on negative past returns, and this weight decays more slowly into the past (a higher  $\lambda_2$ ) for negative returns. Similarly, we find both  $\lambda_1$  and  $\lambda_2$  to be higher for past stock returns during weeks when the market as a whole is doing poorly. Moreover, an individual stock's return relative to its peer performance also seems to affect investor beliefs. Specifically, we find  $\lambda_2$  to be higher for contest-weeks with more dispersed returns.

All these results so far can be attributed to salience. Negative news—both on individual stocks and on the overall market—and past stock returns that are significantly different from peers' returns can be rather salient to investors. Indeed, using news coverage as a direct measure of salience, we find investor expectations to respond more strongly to salient past returns (a larger  $\lambda_1$ ), and salient returns from both the recent past and the distant past affect investor expectations (a higher  $\lambda_2$ ).

<sup>1</sup> Cassella and Gulen (2019) analyze the relation between investor expectations about aggregate stock market returns and the relative pricing of stocks in the cross-section. Bordalo et al. (2019) examine analyst expectations about earnings growth in the cross-section. However, these studies do not directly analyze cross-sectional data of return expectations.

<sup>2</sup> As explained below, our data only allow us to study investor expectations about stock returns over a weekly horizon.

<sup>3</sup> Some social media platforms (e.g., StockTwits and Seeking Alpha) collect a textual form of user opinions about stock performance, but sometimes textual information cannot be easily converted to precise quantitative information. Equity analysts' target prices have also been used to compute return expectations. However, these return expectations can be affected by herding and “selection bias” that arise from analysts' career concerns and investment banking relations (Brav and Lehavy, 2003). (Notwithstanding these issues with target prices, we show suggestive evidence for extrapolative beliefs even among equity analysts, after removing the illiquid penny stocks; results are available upon request.) Finally, individual investors' trading decisions are sometimes used as a measure for investor beliefs (Barber et al., 2009), though they can be driven by other factors, such as liquidity shocks and preferences (Odean, 1998; Barberis and Xiong, 2012; Ingersoll and Jin, 2013). Moreover, short-sale constraints can prevent an investor from expressing negative return expectations through trading, and the investor's choice set is limited to stocks that recently caught her attention (Barber and Odean, 2008).

<sup>4</sup> Early work by Greenwood and Shleifer (2014), Barberis et al. (2015), and Cassella and Gulen (2018) has used this specification to study investors' return expectations about the aggregate stock market.

We further examine how investor and firm characteristics affect expectation formation. We show that, compared to nonprofessionals, financial professionals display a lower degree of extrapolation. Specifically, the  $\lambda_1$  estimate for professionals is significantly lower than that for nonprofessionals, suggesting that professionals rely less on past stock returns when forming expectations about returns over the next week. Moreover, the  $\lambda_2$  estimate for professionals is significantly higher than that for nonprofessionals, suggesting that professionals' expectations rely on past returns over a longer history. At the firm level, we find that  $\lambda_1$  is positively related to firm size but is negatively related to the firm's average volatility of weekly returns. We also find that  $\lambda_2$  is positively related to firm size and turnover but is negatively related to the firm's book-to-market ratio. For the effects of these firm characteristics on investor expectations, we offer some potential explanations that are related to salience and visibility. Overall, our cross-sectional analysis provides new empirical regularities that can inform future theoretical work on investor beliefs.

Given our observations on how Forcerank participants form expectations about future stock returns, a natural follow-up question is whether these expectations are accurate or systematically biased. We find that a higher Forcerank score significantly predicts a lower return over the next week in Fama-MacBeth cross-sectional regressions. We then decompose the Forcerank score into two components: (1) a predicted score—a weighted average of the stock's past 12-week returns with the weights constructed using the  $\lambda_1$  and  $\lambda_2$  estimates of Forcerank participants—and (2) the residual. We find that both components significantly predict future returns with a negative sign, indicating that the beliefs of Forcerank participants are systematically biased. Furthermore, these return predictability results survive the controls of past returns—raw returns, return ranks, and dummies capturing extreme returns—over the past one week, one month, and one quarter. Therefore, our results are not simply rediscovering the well-documented short-term return reversal phenomenon.

To clarify, we do not claim that Forcerank users alone move stock prices, nor that they represent all the market participants. Instead, we interpret our evidence as suggesting that the beliefs of these Forcerank users represent the thinking process of a broader group of behavioral investors in the market. To better understand the impact of extrapolative beliefs on asset prices, we present a cross-sectional model of return extrapolation. The model features two types of agent, extrapolators and fundamental traders. Consistent with the beliefs of Forcerank users, extrapolators form expectations about the future returns of individual stocks by extrapolating from the recent past returns of these stocks, and they trade stocks according to these extrapolative beliefs. Fundamental traders, on the other hand, serve as arbitrageurs who correct for mispricing. However, these traders are risk averse and hence cannot completely undo the mispricing caused by extrapolators. As a result, extrapolator beliefs negatively predict future stock returns, just as we have shown empirically.

Importantly, the stylized model described above makes predictions regarding the heterogeneity of return predictability in the cross-section. Specifically, return

predictability should be stronger among stocks whose clienteles are dominated by behavioral extrapolators, and among stocks with a higher degree of extrapolation—this is measured by  $\lambda_{i,1}(1 - \lambda_{i,2})$  for stock  $i$  in the model. Both predictions are strongly borne out in our sample.

Finally, we evaluate the economic significance and generalizability of our return predictability results. A trading strategy that buys stocks with low Forcerank scores and sells stocks with high Forcerank scores generates a significant profit of seven basis points (bps) per day (or, equivalently, about 18% per year) in our sample, after controlling for the Fama-French five factors, the momentum factor, and the short-term reversal factor. The returns accrue gradually over time after portfolio formation, so they go beyond bid-ask bounce and other market microstructure effects. As an external validation of our findings, we extend our analysis to stocks that are not covered by the Forcerank platform over a longer sample period. To do this, we compute predicted Forcerank scores for non-Forcerank stocks as the weighted average of their past 12-week returns where the weights are calibrated to the beliefs of Forcerank participants.

We find that these predicted scores negatively forecast next week's returns in the full sample of non-Forcerank stocks. The associated trading strategy delivers a highly significant risk-adjusted return, outperforming the standard short-term return reversal strategies that sort on either past one-week or past one-month returns. Among the largest non-Forcerank stocks, the trading strategy based on predicted scores continues to generate a significant risk-adjusted return. Moreover, predicted scores still outperform simple past returns even among this subset of stocks that are least affected by illiquidity.

Our paper adds to a literature that uses survey data to study investor beliefs (Piazzesi and Schneider, 2009; Amromin and Sharpe, 2013; Greenwood and Shleifer, 2014; Kojien et al., 2015; Kuchler and Zafar, 2019; Giglio et al., 2020; Liu et al., 2020). More broadly, our paper adds to a recent literature that analyzes the role of investor beliefs in explaining asset prices in aggregate markets and in the cross-section (Hirshleifer et al., 2015; Barberis et al., 2015; Barberis et al., 2018; Cassella and Gulen, 2018; Cassella and Gulen, 2019; Bordalo et al., 2018; Gennaioli and Shleifer, 2018; Bordalo et al., 2019; Jin and Sui, 2019; Nagel and Xu, 2019; Greenwood et al., 2019).<sup>5</sup>

Furthermore, our empirical findings provide direct support for return extrapolation, as they differ from the predictions of alternative theories that are based on fundamental extrapolation (Barberis et al., 1998; Daniel et al., 1998).<sup>6</sup> Finally, our paper contributes to the voluminous literature on the short-term return reversal starting from Fama (1965), Jegadeesh (1990), and Lehmann (1990). Our finding of significant return predictability on the largest and most liquid stocks suggests that extrapolative beliefs, in addition to liquidity shocks, can also be an important contributor to short-term return reversals.

<sup>5</sup> See Barberis (2018) for a review.

<sup>6</sup> We discuss this point further in Section 3.

While the Forcerank expectation data are weekly forecasts and hence map nicely to the horizon over which short-term return reversals operate, they do not speak directly to return anomalies over longer horizons—e.g., the medium-term momentum and long-term reversals—in the absence of additional modeling assumptions. Nevertheless, as discussed in the Appendix, some field and lab evidence suggests that the insights we obtain from our weekly data regarding investor beliefs can still be generalizable to forecasts over longer horizons.

In what follows, Section 2 provides a detailed description of the Forcerank platform and our data set. Section 3 presents our empirical analysis of the formation of investor expectations. Section 4 shows the predictive power of Forcerank scores for future stock returns as well as the performance of trading strategies. Section 5 concludes. The Appendix contains a stylized asset pricing model of extrapolative beliefs and some additional discussions.

## 2. Data and summary statistics

In this section, we describe the data from Forcerank. Forcerank is a crowdsourcing platform that organizes weekly competitions in which participants enter thematic games, and in each game, rank a list of ten stocks according to their *perceived* performance (percentage gain) of these stocks over the next week.

There are two main types of game. Most games focus on a particular industry group. For example, in one game, contestants are asked to rank ten stocks from the same e-commerce industry based on their expectations of these stocks' returns over the next week. Occasionally, the industry group is further partitioned by the market capitalizations of the stocks. For example, one game can contain only large stocks from the biotech industry. The other type of game is based on special themes, such as exchange-traded funds (ETFs) or the most heavily shorted stocks. We focus on individual firms in our study and therefore exclude games that involve ETFs. Table 1 lists the types of game in our final sample, which covers a period from February 2016 to December 2017.

In addition, most games are repeatedly conducted over time on the platform, resulting in multiple weekly contests for a given game. The goal of the participants is to precisely forecast rankings of future stock returns: they want to match their perceived rankings of the stocks with the actual rankings of these stocks based on their realized returns over the next week. Fig. 1 illustrates an example of one such contest.

Upon completion of each contest, Forcerank assigns points to participants based on the accuracy of their own rankings exclusively; it does not benchmark against the performance of other participants. For most of our sample period, points do not result in monetary compensation due to the legal risks involved.<sup>7</sup> Instead, Forcerank main-

**Table 1**

Games and contests in the sample.

The table presents the number of contests for different types of game in our sample. Each game consists of ten stocks, all of which share one or multiple specific characteristics. Most games are repeatedly conducted over time on the platform, resulting in multiple weekly contests for a given game. There are two main types of game in our sample: (1) industry games, which include stocks in a specific sector or industry; and (2) games with the most heavily shorted stocks, which include stocks with high short interest in the past month. In the final row under industry games, other industries include chemicals, pharmaceuticals, and oil services.

Types of game	Number of contests
Industry	1,318
Enterprise software (large: 69; mid/small: 67)	136
Biotech (large: 95; mid: 20)	115
Social media	111
E-commerce	108
Apparel	101
E&P (large)	96
Hardware	88
Fast food	69
Media	69
Airlines	68
Investment banks	68
Semiconductors (large)	65
Restaurants	64
Others	160
Most heavily shorted (March 2016 to December 2017)	78
Total	1,396

tains weekly leader boards where participants are ranked based on the cumulative points they have earned from the past 13 weeks. Ranking participants based on cumulative points helps alleviate strategic behavior that can arise in a “winners-take-all” tournament, since it is difficult to be both strategic and consistently accurate.

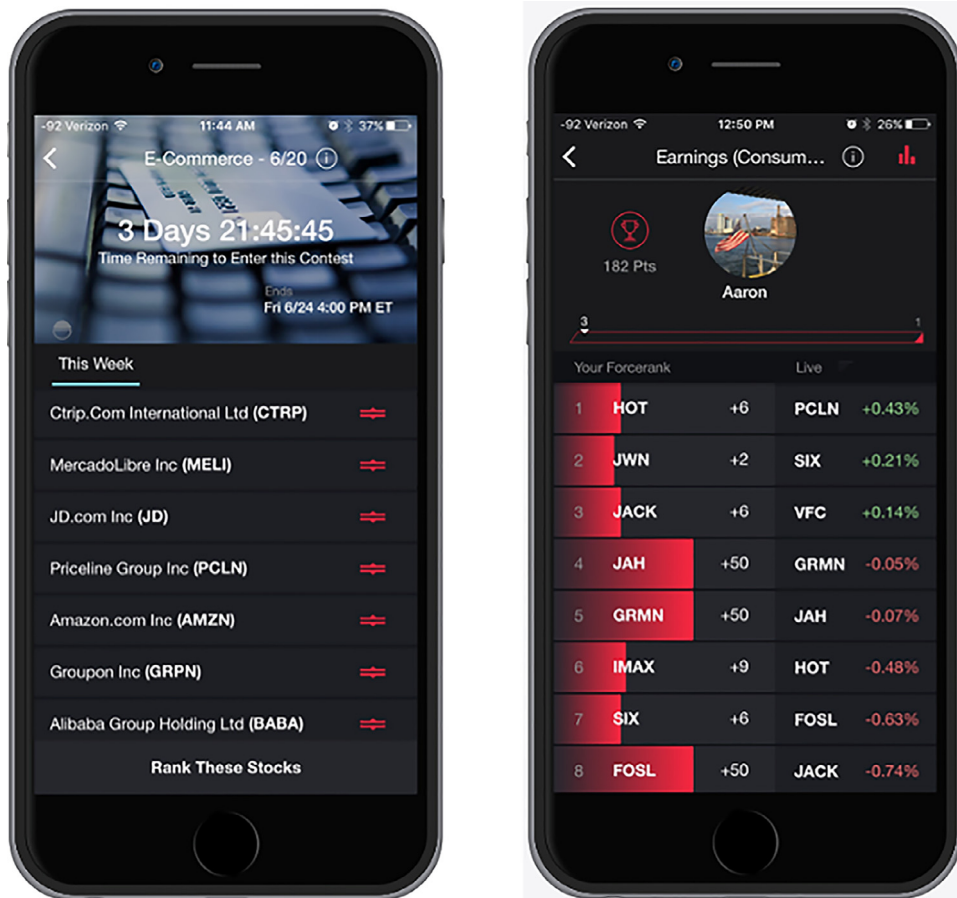
Furthermore, strategic behavior is difficult on this platform for two other reasons. First, during each contest, participants do not observe the current consensus rankings or individual rankings made by other participants. Second, the default initial rankings are randomized for every participant, so there is no common default rankings across participants to benchmark against. It is therefore our view that users who choose to participate on Forcerank are likely to truthfully reveal their return expectations across stocks in the game.

Our sample contains mostly industry contests (1,318 out of a total of 1,396 contests). Popular industries covered in our sample include enterprise software (136 weekly contests), biotech (115 weekly contests), social media (111 weekly contests), e-commerce (108 weekly contests), and apparel (101 weekly contests). Stocks covered in these contests tend to be household names that attract attention

<sup>7</sup> Monetary compensation could turn the Forcerank game into an illegal security-based swap in the eyes of the SEC (see <http://dodd-frank.com/sec-says-mobile-phone-game-is-an-illegal-security-based-swap>). Forcerank canceled its cash payments to participants in June 2016. Given this,

we split our sample into two parts, one before June 2016 and one after. We find that the degree of extrapolation from participant expectations is strong for both parts of the sample. The lack of monetary compensation, however, can explain the slow growth in user participation as well as Forcerank's decision to temporarily shut down the platform since April 2018 to focus its limited resource on developing another crowdsourcing platform called Estimote.





**Fig. 1.** Illustration of the Forcerank interface. The figure on the left presents a screenshot of the interface for a contest for ranking ten stocks from the e-commerce industry, taken on Thursday, June 16, 2016. The participant is asked to rank ten stocks based on her expectations of these stocks' returns over the subsequent "evaluation" week (from 9:30 am June 20, 2016 to 4:00 pm June 24, 2016). The current time is 11:44 am, and the time remaining to enter the rankings is three days 21 hours 45 minutes and 45 seconds. The participant could vertically drag the bars next to the company names to rank these stocks. The figure on the right presents a screenshot of the scoring page for a different contest during its "evaluation" week. The right column under "Live" displays the actual rankings of the stocks based on their realized returns thus far (since the beginning of the evaluation week). The left column under "Your Forcerank" shows the rankings submitted by the user "Aaron" with the corresponding live points earned from this contest. The point next to each stock is based on the difference between the user's ranking and the actual ranking; a larger difference leads to a lower point. In addition, a higher point is given to a more accurate ranking for the top-ranked stock or the bottom-ranked stock. The "live points" become "final points" for this contest when the evaluation week ends at 4:00 pm on Friday.

from individual investors. Over time, Forcerank expanded its game coverage to also include industries such as fast food, investment banking, airlines, and semiconductors. The only non-industry game we study involves the heavily shorted stocks (78 weekly contests that span a period from March 2016 to December 2017).

Our final sample contains 293 unique stock tickers, and it contains 12,798 contributions submitted by 1,045 distinct users. Table 2 presents a breakdown of stocks and users.

Stocks in our sample tend to be large stocks. The average stock has a market capitalization of \$56.6 billion (the median is \$15.4 billion). Using the NYSE size cutoffs, the average stock in our sample has a size quintile rank of 4.20. This fact is important for interpreting our subsequent return predictability results: given their sizes, stocks in our sample are less likely to be subject to the short-term

return reversal induced by liquidity shocks. Our sample also tilts toward growth stocks: the average stock has a book-to-market ratio  $B/M$  of 0.37 (the median is 0.26), and the average  $B/M$  quintile rank is 2.20.

The user participation in our sample is highly skewed. While about half of the users each played only three contests, the most active 1% played 355 contests covering 31 different games. In a subsequent analysis, we focus on some of these regular participants and examine whether their expectation formation process changes over time.

We observe the self-reported professional background among a fraction of users who registered before March 2017. Specifically, among the 606 users who registered before March 2017, 244 of them chose to report their professional background. Panel B of Table 2 breaks down these 244 users. Among them, 72 are financial profession-

**Table 2**

Summary statistics of stocks and users in the sample.

The table presents descriptive statistics for stocks and users in our sample. Panel A reports firm-week-level financial characteristics and user-level participation characteristics. Financial characteristics include size, book-to-market ratio ( $B/M$ ), and institutional ownership ( $IO$ ). The size and  $B/M$  quintile groups are obtained by matching each firm-week observation from July of year  $t$  to June of year  $t + 1$  with one of the 25 Fama-French size and  $B/M$  portfolios based on (1) market capitalization at the end of June of year  $t$  and (2) the book equity of the fiscal year  $t - 1$  divided by the market value of the equity at the end of December of year  $t - 1$ . Panel B reports the distribution of users in our sample by their professions; we only observe the self-reported professional background among a fraction of users who registered before March 2017.

Panel A: Stock and user characteristics							
Firm-week-level financial characteristics (Number of observations $N = 11,140$ )							
	mean	sd	p1	p25	p50	p75	p99
Size (in million)	56,602.77	102,785.18	600.73	3,949.91	15,413.54	53,054.52	515,586.56
$B/M$	0.37	0.37	0.01	0.15	0.26	0.47	1.55
Size quintile	4.20	1.08	2.00	3.00	5.00	5.00	5.00
$B/M$ quintile	2.20	1.31	1.00	1.00	2.00	3.00	5.00
$IO$	0.48	0.34	0.00	0.00	0.63	0.75	1.00
User-level participation characteristics (Number of observations $N = 1,045$ )							
	mean	sd	p1	p25	p50	p75	p99
Number of games	4.39	6.61	1.00	1.00	2.00	4.00	31.00
Number of contests	18.85	88.01	1.00	1.00	3.00	8.00	355.00
Number of weeks	3.71	6.59	1.00	1.00	2.00	4.00	38.00
Panel B: User background							
	Frequency	Percent					
Financial professional ( $N = 72$ )							
Sell side	47	7.76					
Buy side	14	2.31					
Independent	11	1.82					
Nonprofessional ( $N = 172$ )							
Financials	6	0.99					
Academia	1	0.17					
Consumer discretionary	5	0.83					
Consumer staples	2	0.33					
Energy	1	0.17					
Healthcare	6	0.99					
Industrials	1	0.17					
Information technology	22	3.63					
Materials	4	0.66					
Student	124	20.46					
Missing	362	59.74					
Total	606	100.00					

als. We conjecture that, compared to nonprofessionals, the degree of extrapolation is less pronounced among financial professionals. In our empirical analysis, we confirm this conjecture.

### 3. Expectation formation

In this section, we study the formation of investor expectations using the Forcerank data. First, we study the relation between the average beliefs of Forcerank users and past variables such as past stock returns. We then examine heterogeneity in beliefs by looking at how user and firm characteristics affect expectation formation.

To start, we analyze how past stock returns affect Forcerank users' average expectation of future stock returns. In each week  $t$ , individuals are asked to submit rankings of ten stocks according to their *perceived* performance of these stocks over week  $t + 1$ . For each stock in each contest, we measure investors' average expectation

using the consensus Forcerank score, which is the average ranking across all individuals in that contest. For each individual, the stock she ranked the highest receives a score of ten, and the stock she ranked the second highest receives a score of nine. Similarly, the stock she ranked the lowest receives a score of one, and the stock she ranked the second lowest receives a score of two.

#### 3.1. Linear model

We start with a simple linear regression model using the consensus Forcerank score as the dependent variable and past stock returns as the explanatory variables:

$$\text{Forcerank}_{i,t} = \gamma_0 + \sum_{s=0}^n \beta_s \cdot R_{i,t-s} + \varepsilon_{i,t}, \quad (1)$$

where  $\text{Forcerank}_{i,t}$  is the end-of-week- $t$  consensus ranking based on investors' average expectation about the perfor-

**Table 3**

Extrapolative beliefs: linear regression model.

The table presents the results of a contest-level linear regression specified in Eq. (1) of the main text. The dependent variable is the consensus ranking (one to ten)—a stock's average ranking across all individuals: the highest ranked stock receives a score of ten, and the lowest ranked stock receives a score of one. The explanatory variables include lagged returns from week  $t - 11$  to week  $t$ . Column (1) uses the raw level of past stock returns. Columns (2) and onwards focus on contest-adjusted returns (i.e., the stock return in excess of the average return of the ten stocks in the contest). Columns (3) and (4) separately examine the idiosyncratic and systematic components of past stock returns (according to the CAPM). Column (5) uses an ordered logit regression. Column (6) converts the explanatory variables from past weekly returns to the stocks' actual past rankings. Columns (7) and (8) repeat the analysis in Column (2) separately for the first half (before March 1, 2017) and the second half (after March 1, 2017) of our sample period. Finally, Column (9) includes controls such as tones of news coverage and the stock's CAPM expected returns from the previous 12 weeks as well as past fundamentals measured by monthly revisions in consensus earnings forecasts (or, during the earnings announcement months, by the actual earnings surprises) from the previous 12 months. The standard errors are in parentheses. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% level, respectively.

Dep. var:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	level	adj	idio	sys	ord. logit	ranking	1 <sup>st</sup> half	2 <sup>nd</sup> half	w controls
Ret( $t$ )	11.21*** (0.559)	16.98*** (0.684)	13.02*** (0.697)	-4.352* (2.223)	11.95*** (0.476)	0.276*** (0.009)	17.83*** (0.963)	19.10*** (1.060)	18.92*** (0.820)
Ret( $t - 1$ )	3.298*** (0.555)	5.139*** (0.679)	3.105*** (0.696)	-5.144** (2.173)	3.479*** (0.448)	0.0723*** (0.009)	6.601*** (0.953)	3.897*** (1.058)	4.891*** (0.817)
Ret( $t - 2$ )	3.150*** (0.560)	4.327*** (0.688)	3.346*** (0.698)	-5.083** (2.131)	2.845*** (0.445)	0.0558*** (0.009)	3.993*** (0.960)	4.883*** (1.060)	4.559*** (0.817)
Ret( $t - 3$ )	2.025*** (0.565)	2.821*** (0.694)	1.625** (0.705)	-6.398*** (2.103)	2.013*** (0.453)	0.0486*** (0.009)	3.677*** (0.968)	2.363** (1.071)	2.406*** (0.830)
Ret( $t - 4$ )	2.590*** (0.564)	3.703*** (0.695)	2.906*** (0.707)	-4.580** (2.049)	2.589*** (0.455)	0.0556*** (0.009)	4.150*** (0.973)	4.174*** (1.070)	3.754*** (0.829)
Ret( $t - 5$ )	1.911*** (0.559)	2.259*** (0.689)	1.920*** (0.710)	-4.114** (2.038)	1.485*** (0.454)	0.0457*** (0.009)	1.616* (0.958)	3.265*** (1.072)	2.080** (0.830)
Ret( $t - 6$ )	0.785 (0.542)	1.188* (0.668)	1.557** (0.698)	-3.848* (1.991)	1.110** (0.443)	0.0363*** (0.009)	-0.197 (0.946)	3.952*** (1.054)	2.030** (0.814)
Ret( $t - 7$ )	2.146*** (0.541)	3.669*** (0.668)	3.170*** (0.698)	-2.896 (1.986)	2.489*** (0.445)	0.0541*** (0.009)	2.418** (0.946)	5.894*** (1.054)	4.687*** (0.814)
Ret( $t - 8$ )	0.651 (0.542)	1.503** (0.679)	1.087 (0.692)	-3.862** (1.933)	1.171*** (0.444)	0.0251*** (0.009)	0.219 (0.949)	3.446*** (1.062)	1.294 (0.812)
Ret( $t - 9$ )	1.575*** (0.535)	2.466*** (0.669)	2.196*** (0.688)	-2.617 (1.909)	1.732*** (0.442)	0.0431*** (0.009)	1.579* (0.922)	3.903*** (1.079)	2.588*** (0.806)
Ret( $t - 10$ )	1.339** (0.520)	1.529** (0.659)	1.507** (0.688)	-4.435** (1.837)	1.125** (0.443)	0.0246*** (0.009)	0.206 (0.898)	3.706*** (1.113)	1.640** (0.804)
Ret( $t - 11$ )	0.838 (0.516)	0.812 (0.652)	0.633 (0.680)	-3.695** (1.796)	0.498 (0.436)	0.0140 (0.009)	-0.126 (0.885)	2.258** (1.098)	0.319 (0.794)
Observations	12,010	12,010	10,362	10,362	12,010	12,050	5,542	6,468	10,170
R-squared	0.042	0.064	0.042	0.004	0.0157	0.099	0.074	0.069	0.114

mance of stock  $i$  over week  $t + 1$ ;  $R_{i,t-s}$  represents the lagged return (or the lagged contest-adjusted return we define below) of stock  $i$  over week  $t - s$ , and  $s$  goes from 0 to 11.

The results are reported in Table 3. Column (1) uses the raw level of past returns. The results show clear evidence that past returns drive Forcerank scores. The coefficients on the past 12 weekly returns are all positive and mostly significant. More important, the coefficients on recent past returns are in general higher than those on distant past returns.<sup>8</sup>

Given that individuals submit relative rankings on Forcerank, it is possible that the relative levels of past

returns within a contest are more relevant to expectation formation. In Columns (2) and onwards, we adjust past returns by demeaning these return levels within each contest: we compute contest-adjusted returns by subtracting from raw returns the average return of the ten stocks in the contest. The extrapolative pattern in Column (2) remains similar to that in Column (1). At the same time, compared to Column (1), the coefficients on contest-adjusted past returns and the R-squared all increase, indicating a better fit to the data. We further decompose past weekly returns into their systematic and idiosyncratic components using the CAPM. Columns (3) and (4) show that investors extrapolate only from past idiosyncratic returns. This result is intuitive: Forcerank contests contain similar stocks from the same industry, and as a result, within-contest return variations are likely to be idiosyncratic in nature.

The positive relation between the current return expectation and recent past returns is robust to different estimation methods and sample periods. In Column (5), we estimate an ordered logit model that accounts for the ranking nature of our dependent variable. In Col-

<sup>8</sup> This regression result is supported by survey evidence from 20 Caltech undergraduate students who participated in ranking stocks on Forcerank from February 2018 to March 2018. When asked about how they came up with their stock rankings, the students typically responded by saying that the rankings are based on "the last week and last month's performance," "a quick look of past month returns," or "the last week's ranks." The survey evidence, while limited in its scope, does suggest that individuals extrapolate from past returns directly rather than from other variables that are simply correlated with past returns.

umn (6), we convert the explanatory variables from past weekly returns to the stocks' actual past rankings. Across these alternative regression specifications, the pattern of return extrapolation remains similar. The coefficients on the past 12 weekly returns (or the rankings of these past returns) are all positive and mostly significant. Moreover, the coefficients on recent past returns remain significantly higher than those on distant past returns.

Columns (7) and (8) repeat the analysis in Column (2) separately for the first half (before March 1, 2017) and the second half (after March 1, 2017) of our sample period. While we observe return extrapolation in both subperiods, the pattern seems to be stronger in the second half. A potential reason is that, with increased user participation on Forcerank, the consensus ranking became less noisy over time.

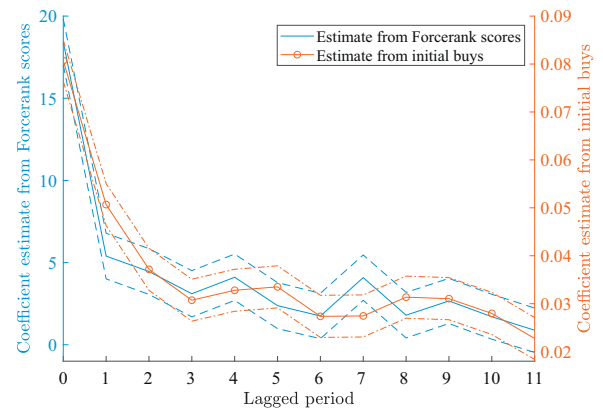
Finally, Column (9) includes some other potential determinants of investors' return expectations. These additional controls include tones of news coverage, the stock's CAPM expected returns from the previous 12 weeks, and past fundamental news measured by monthly revisions in consensus earnings forecasts (or, during the earnings announcement months, by the actual earnings surprises) from the previous 12 months. These additional controls do not alter the basic pattern of return extrapolation: the general decay pattern among coefficients for past returns remains strong and quantitatively similar with and without controls.<sup>9</sup> However, when we add controls, the additional data requirement reduces our sample size by more than 15%. For this reason, we do not include these additional controls in most of our remaining analysis.

It is important to note that the results reported in Table 3 provide direct support for return extrapolation, as they do not arise from alternative models of biased beliefs such as Barber et al. (1998) and Daniel et al. (1998). In these models, agents' expectation of future cash flows becomes overly optimistic after a sequence of positive cash flow shocks—and hence positive returns—from the recent past. However, the agents' expectation of future returns stays constant and hence does not vary with past returns. In contrast with this prediction, Table 3 shows that investors' expectations of future returns are not constant: they are a positive function of recent past returns. As such, our results help to distinguish predictions from different theories.

### 3.2. External validation

An immediate concern is that our findings so far are driven by features unique to Forcerank: the game specification, the interface, and the characteristics and incentives of users who are self-selected to participate, among others. As an external validation, we now examine the trading

<sup>9</sup> While not reported in Column (9) due to space constraints, the coefficients on past fundamental news are positive and sometimes significant, suggesting that individuals' return expectations are in part affected by past fundamentals. Nevertheless, in contrast to those on past returns, the coefficients on past fundamental news do not seem to decay into the past.



**Fig. 2.** Extrapolative beliefs: Forcerank scores versus initial buys. The figure plots the coefficient estimates from regressing the consensus Forcerank score on past 12 weekly returns (left y-axis) and the coefficient estimates from regressing the initial buys of retail investors on past 12 weekly returns (right y-axis). Initial buys are measured using individual-level transaction records from a large discount brokerage firm over the period from January 1991 to December 1996 (as in Barber and Odean, 2000). In the initial buy regression, the dependent variable is an indicator that equals one for week  $t$  if, during that week, there is at least one retail investor who purchased the stock for the first time in the sample. The initial buys are based on the trades of frequent traders whose median time of a round-trip trade is less than ten days; the round-trip time is computed as the time gap between each sale and its most recent purchase. Investors who have less than ten round-trip trades are removed from the sample. The solid line and the solid-circle line correspond to the coefficient estimates. The dashed lines and the dash-dot lines correspond to the 95% confidence level.

behavior of a large group of retail investors. Here we focus on the initial buys of these retail investors. Compared to other types of trade—e.g., additional buys of the same stock or sales, which could be driven by factors other than investor beliefs such as investor preferences or liquidity needs—initial buys are more likely to reflect investors' return expectations.

We measure initial buys using individual-level transaction records from a large discount brokerage firm over the period from January 1991 to December 1996 (as in Barber and Odean, 2000). We remove investors who have less than ten round-trip trades from the sample. Moreover, to match Forcerank's one-week forecasting horizon, we focus on frequent traders whose median time of a round-trip trade is less than ten days. These frequent traders account for 26.5% of all the investors who have more than ten round-trip trades.

We run a linear regression of initial buys on past 12 weekly returns; this is similar to the regression in Eq. (1). The dependent variable is an indicator that equals one for week  $t$  if, during that week, there is at least one retail investor who purchased the stock for the first time in the sample. Fig. 2 plots the coefficient estimates on past 12 weekly returns using initial buys as the dependent variable (right y-axis) against those from the Forcerank sample using the consensus Forcerank score as the dependent variable (left y-axis). The solid line and the solid-circle line correspond to the coefficient estimates. The dashed lines and the dash-dot lines correspond to the 95% confidence interval. We find that the extrapolative patterns are very



similar across these two settings: the changes in the two sets of coefficients are almost proportional to each other.<sup>10</sup>

### 3.3. Exponential decay model

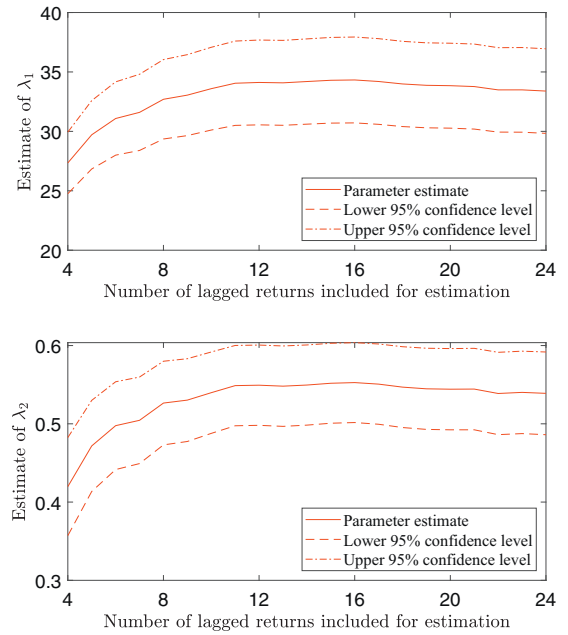
The linear regression in Eq. (1) allows for independent weights on different past returns. From this simple specification, we have observed a clear and robust decay pattern in the relation between investors' current return expectation and recent past returns. To capture this pattern parsimoniously, we now estimate a parametric regression model that assumes an exponential decay of weights on past returns:

$$\text{Forcerank}_{i,t} = \lambda_0 + \lambda_1 \cdot \sum_{s=0}^n w_s R_{i,t-s} + \varepsilon_{i,t},$$

$$\text{where } w_s = \frac{\lambda_2^s}{\sum_{j=0}^n \lambda_2^j}. \quad (2)$$

This exponential decay specification has been previously estimated by Greenwood and Shleifer (2014), Barberis et al. (2015), and Cassella and Gulen (2018), using aggregate stock market data. It allows us to characterize extrapolative expectations by two parameters. The first parameter,  $\lambda_1$ —a scaling factor that multiplies all past returns of stock  $i$ —captures a “level” effect (i.e., the overall extent to which investor expectations respond to these past returns). The second parameter,  $\lambda_2$ —which governs how past returns are relatively weighted in forming expectations—captures a “slope” effect: a  $\lambda_2$  closer to zero means that investors put much higher weight on recent past returns as opposed to distant past returns. When an investor puts more weight on all past returns of stock  $i$  and, furthermore, assigns more weight to more recent returns on a relative basis, her beliefs become more extrapolative. That is, a higher  $\lambda_1$  and a lower  $\lambda_2$  jointly lead to a higher degree of extrapolation; indeed, the extrapolation model in Appendix A shows that the appropriate measure for the degree of extrapolation is  $\lambda_1(1 - \lambda_2)$ . We first estimate the two parameters,  $\lambda_1$  and  $\lambda_2$ , by assuming them to be constant in the full sample across all stocks and individuals. Later in this section, we study the heterogeneity of these parameters.

Eq. (2) suggests that the estimates of  $\lambda_1$  and  $\lambda_2$  depend on  $n + 1$ , the total number of lagged weekly returns included in this parametric regression. Fig. 3 shows that the estimates of  $\lambda_1$  and  $\lambda_2$  both become stable when



**Fig. 3.** Estimates of  $\lambda_1$  and  $\lambda_2$  and the number of lagged returns included. The figure plots the estimates of  $\lambda_1$  and  $\lambda_2$ , each as a function of  $n + 1$ , the total number of lagged weekly returns included in the nonlinear regression specified in Eq. (2) of the main text:

$$\text{Forcerank}_{i,t} = \lambda_0 + \lambda_1 \cdot \sum_{s=0}^n w_s R_{i,t-s} + \varepsilon_{i,t}, \quad \text{where } w_s = \frac{\lambda_2^s}{\sum_{j=0}^n \lambda_2^j}, \quad 0 \leq \lambda_2 < 1.$$

we include 12 or more of the past weekly returns in the estimation ( $n \geq 11$ ). Thus, we use  $n = 11$  for the rest of our analysis.

Table 4 confirms the presence of return extrapolation using the nonlinear regression in Eq. (2). Specifically, Column (1) uses the raw level of past returns. Columns (2) and onwards focus on contest-adjusted returns. Column (2) reports an estimate of 34.12 for  $\lambda_1$  and an estimate of 0.55 for  $\lambda_2$ . These joint estimates suggest that Forcerank participants exhibit a strong degree of extrapolation. Column (3) finds similar patterns using the idiosyncratic component of past returns. To check the robustness of these estimates, Column (4) replaces stocks' past returns by their actual past rankings. Columns (5) and (6) further break up the regression results for the first half (before March 1, 2017) and the second half (after March 1, 2017) of our sample period. Finally, Column (7) includes controls of past fundamental news, tones of news coverage, and the CAPM expected returns. Across all columns, we find  $\lambda_1$  to be significantly positive and  $\lambda_2$  to be positive and significantly smaller than one.<sup>11</sup> For the rest of the paper, we focus primarily on the nonlinear specification in Eq. (2) when analyzing investor expectations, as it succinctly summarizes return extrapolation using two parameters.

<sup>10</sup> The Forcerank data only measure return expectations over one week, a short forecasting horizon. As a result, a concern is that our data are not helpful for understanding investor beliefs over longer horizons (such as six months or one year). One observation can be useful for addressing this concern: investors tend to look at returns over the past few weeks when forecasting the next week's return, and they tend to look at returns over the past few years when forecasting the next year's return. That is, when forecasting the future return over a time horizon of  $t$ —for instance,  $t$  can be a week, a quarter, or a year—investors tend to look at the past returns over a time horizon of  $N \cdot t$ , where the parameter  $N$  tends to be stable and independent of the forecasting horizon  $t$ , capturing a fundamental psychological factor such as the degree of recency bias or the speed of memory decay. Such stability is called “time-scale invariance,” an important finding from the psychology literature. We provide additional discussion about time-scale invariance in Appendix B.

<sup>11</sup> Moreover, if we replace weekly past returns on the right-hand side of Eq. (2) by daily past returns, we can estimate  $\lambda_1$  and  $\lambda_2$  using these daily returns. We find that the daily estimations of  $\lambda_1$  and  $\lambda_2$  are consistent with the weekly estimations. That is, the estimate of  $\lambda_1$  using daily past returns is about five times as big as the estimate of  $\lambda_1$  using weekly past returns. The estimate of  $\lambda_2$  using daily past returns is about the one-fifth power of the estimate of  $\lambda_2$  using weekly past returns.

**Table 4**

Extrapolative beliefs: exponential decay model.

The table presents the results of a contest-level nonlinear regression specified in Eq. (2) of the main text:

$$\text{Forcerank}_{i,t} = \lambda_0 + \lambda_1 \cdot \sum_{s=0}^n w_s R_{i,t-s} + \varepsilon_{i,t}, \quad \text{where } w_s = \frac{\lambda_2^s}{\sum_{j=0}^n \lambda_2^j}, \quad 0 \leq \lambda_2 < 1.$$

The dependent variable is the consensus ranking (one to ten)—a stock’s average ranking across all individuals: the highest ranked stock receives a score of ten, and the lowest ranked stock receives a score of one. The explanatory variables include lagged returns from week  $t - 11$  to week  $t$ . Column (1) uses the raw level of past stock returns. Columns (2) and onwards focus on contest-adjusted returns (i.e., the stock return in excess of the average return of the ten stocks in the contest). Column (3) uses the idiosyncratic component of past stock returns. Column (4) converts the explanatory variables from past weekly returns to the stocks’ actual past rankings. Columns (5) and (6) repeat the analysis in Column (2) separately for the first half (before March 1, 2017) and the second half (after March 1, 2017) of our sample period. Finally, Column (7) includes controls such as tones of news coverage and the stock’s CAPM expected returns from the previous 12 weeks as well as past fundamentals measured by monthly revisions in consensus earnings forecasts (or, during the earnings announcement months, by the actual earnings surprises) from the previous 12 months. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	Forcerank score						
	level	adj	idio	ranking	1 <sup>st</sup> half	2 <sup>nd</sup> half	w controls
$\lambda_0$	5.401*** (0.027)	5.498*** (0.025)	5.578*** (0.028)	2.789*** (0.118)	5.491*** (0.037)	5.503*** (0.035)	4.576*** (0.067)
$\lambda_1$	23.83*** (1.586)	34.12*** (1.820)	23.14*** (1.688)	0.493*** (0.021)	35.57*** (2.521)	50.74*** (3.460)	31.95*** (1.911)
$\lambda_2$	0.590*** (0.030)	0.549*** (0.026)	0.484*** (0.039)	0.471*** (0.023)	0.534*** (0.035)	0.739*** (0.025)	0.443*** (0.033)
Observations	12,010	12,010	10,362	12,010	5,542	6,468	10,170
R-squared	0.037	0.056	0.036	0.085	0.070	0.052	0.106

3.4. Past return characteristics

To develop a deeper understanding of expectation formation, we generalize the regression in Eq. (2) by separately estimating  $\lambda_1$  and  $\lambda_2$  for past returns of different characteristics. Recent empirical, experimental, and neuroscience studies suggest that expectation formation differs depending on whether past outcomes are positive or negative. In particular, negative past outcomes can have a particularly strong influence on investors’ beliefs about future outcomes (see [Kuhnen, 2015](#) for a review of this evidence). To test this hypothesis in our setting, we separate past weekly returns into positive returns and negative returns and then run a generalized nonlinear regression of the form:

$$\text{Forcerank}_{i,t} = \lambda_0 + \lambda_{1,p} \cdot \sum_{s=0}^n \mathbb{1}_{\{R_{i,t-s} \geq 0\}} \cdot w_{s,p} R_{i,t-s} + \lambda_{1,n} \cdot \sum_{s=0}^n \mathbb{1}_{\{R_{i,t-s} < 0\}} \cdot w_{s,n} R_{i,t-s} + \varepsilon_{i,t}, \quad (3)$$

where  $w_{s,p} = \frac{\lambda_{2,p}^s}{\sum_{j=0}^n \lambda_{2,p}^j}$  and  $w_{s,n} = \frac{\lambda_{2,n}^s}{\sum_{j=0}^n \lambda_{2,n}^j}$ . In other words, this regression allows  $\lambda_1$  and  $\lambda_2$  to differ across positive ( $p$ ) versus negative ( $n$ ) returns of individual stocks.

Column (1) of [Table 5](#) reports the empirical estimates of  $\lambda_{1,p}$ ,  $\lambda_{2,p}$ ,  $\lambda_{1,n}$ , and  $\lambda_{2,n}$ . The results show that return extrapolation is *asymmetric*. In particular, individuals seem to put more weight on negative past returns— $\lambda_{1,n}$  is much larger than  $\lambda_{1,p}$ —and this weight decays more slowly into the past for negative past returns— $\lambda_{2,n}$  is much higher than  $\lambda_{2,p}$  and is therefore much closer to one. While coefficients on positive contest-adjusted past returns become insignificant beyond past one week, the coefficients on

negative contest-adjusted past returns stay strongly significant for many past weeks: returns four weeks earlier are 45% as important as returns in the most recent week in determining the current expectation about future returns.

Column (2) allows  $\lambda_1$  and  $\lambda_2$  to differ across up ( $u$ ) versus down ( $d$ ) markets, where up (down) markets are defined as weeks when the market returns are above (below) the median weekly market return in our sample period (0.23%). We again find both  $\lambda_1$  and  $\lambda_2$  to be higher for past returns during down markets. Related to this finding, [Cassella and Gulen \(2018\)](#) show that  $\lambda_2$  estimated from return expectations about the aggregate stock market is significantly higher in bear markets than in bull markets. Our result complements theirs by showing that (1) a similar pattern of asymmetry in  $\lambda_2$  holds in the cross-section, and (2) the asymmetry in the degree of extrapolation also shows up in the difference between  $\lambda_{1,u}$  and  $\lambda_{1,d}$ . Put differently, negative—rather than positive—information at both the individual stock level and the market level is more important in driving expectations about future stock returns (a higher  $\lambda_1$ ), and it has a more persistent impact on these expectations (a higher  $\lambda_2$ ).<sup>12</sup>

The regression in Column (3) allows  $\lambda_1$  and  $\lambda_2$  to differ across contest-weeks with dispersed (*disp*) versus

<sup>12</sup> Neuroscience studies suggest that extrapolation of positive returns and extrapolation of negative returns take place in different parts of the human brain. [Kuhnen and Knutson \(2005\)](#) find that, when investors are faced with good investment outcomes, the nucleus accumbens of the ventral striatum, a particular region of the brain, is activated, leading to optimistic beliefs about future investment outcomes. On the other hand, when investors are faced with bad investment outcomes, anterior insula, a different region of the brain, is activated, leading to pessimistic beliefs about future investment outcomes. Understanding the differences in the process of activating these two regions of the brain can provide a micro-foundation for the asymmetry in return extrapolation we documented.

**Table 5**

Extrapolative beliefs: past return characteristics.

The table presents the results of contest-level nonlinear regressions for various past return characteristics. The dependent variable in these regressions is the consensus Forcerank score. The explanatory variables are lagged contest-adjusted returns. The regression in Column (1) is specified in Eq. (3) of the main text:

$$\text{Forcerank}_{i,t} = \lambda_0 + \lambda_{1,p} \cdot \sum_{s=0}^n \mathbb{1}_{\{R_{i,t-s} \geq 0\}} \cdot w_{s,p} R_{i,t-s} + \lambda_{1,n} \cdot \sum_{s=0}^n \mathbb{1}_{\{R_{i,t-s} < 0\}} \cdot w_{s,n} R_{i,t-s} + \varepsilon_{i,t},$$

where  $w_{s,p} = \frac{\lambda_{2,p}^s}{\sum_{j=0}^n \lambda_{2,p}^j}$  and  $w_{s,n} = \frac{\lambda_{2,n}^s}{\sum_{j=0}^n \lambda_{2,n}^j}$ . In other words, this regression allows  $\lambda_1$  and  $\lambda_2$  to differ across positive ( $p$ ) versus negative ( $n$ ) returns of individual stocks. Similarly, the regression in Column (2) allows  $\lambda_1$  and  $\lambda_2$  to differ across up ( $u$ ) versus down ( $d$ ) markets, where up (down) markets are defined as weeks when the market returns are above (below) the median weekly market return in our sample period (0.23%). The regression in Column (3) allows  $\lambda_1$  and  $\lambda_2$  to differ across contest-weeks with dispersed ( $disp$ ) versus close ( $close$ ) returns. Contest-weeks with dispersed (close) returns are those where the cross-sectional standard deviation of the ten stocks' returns is above (below) the median (2.77%). Finally, the regression in Column (4) allows  $\lambda_1$  and  $\lambda_2$  to differ across salient ( $s$ ) versus nonsalient ( $ns$ ) returns. To measure the salience level associated with the return of stock  $i$  in week  $t$ , we count the number of news articles written on that stock in that week. We obtain the news coverage data from RavenPack. To ensure that extreme returns are not driving the variation in  $\lambda_1$  and  $\lambda_2$  mechanically, we orthogonalize news coverage with respect to the absolute returns. Specifically, for each week in the sample period, we run a cross-sectional regression of the total amount of news coverage of a stock on its absolute return over the same week. We then define a stock return as salient (nonsalient) if the residual from this cross-sectional regression is above (below) median. The standard errors are in parentheses. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% level, respectively.

	(1)		(2)		(3)		(4)
Dependent variable:	Forcerank score						
$\lambda_0$	6.213*** (0.047)	$\lambda_0$	5.497*** (0.025)	$\lambda_0$	5.498*** (0.025)	$\lambda_0$	5.508*** (0.025)
$\lambda_{1,p}$	18.44*** (1.674)	$\lambda_{1,u}$	27.63*** (2.242)	$\lambda_{1,disp}$	31.10*** (2.056)	$\lambda_{1,s}$	55.84*** (3.462)
$\lambda_{2,p}$	0.0538 (0.063)	$\lambda_{2,u}$	0.387*** (0.048)	$\lambda_{2,disp}$	0.564*** (0.032)	$\lambda_{2,s}$	0.765*** (0.021)
$\lambda_{1,n}$	77.93*** (3.115)	$\lambda_{1,d}$	50.39*** (3.067)	$\lambda_{1,close}$	66.67*** (4.344)	$\lambda_{1,ns}$	28.88*** (2.038)
$\lambda_{2,n}$	0.818*** (0.016)	$\lambda_{2,d}$	0.700*** (0.024)	$\lambda_{2,close}$	0.440*** (0.037)	$\lambda_{2,ns}$	0.356*** (0.043)
Observations	12,010		12,010		12,010		12,010
R-squared	0.081		0.062		0.072		0.063

close ( $close$ ) returns. Contest-weeks with dispersed (close) returns are those for which the cross-sectional standard deviation of the ten stocks' returns is above (below) the median (2.77%). It is not surprising that dispersed contest-weeks are associated with a lower  $\lambda_1$ . The higher standard deviation of the independent variable (the past return) mechanically results in a lower regression coefficient for those contest-weeks. What is more interesting is that these dispersed contest-weeks are also associated with a higher  $\lambda_2$ , suggesting that past returns from dispersed contest-weeks have a more persistent impact on investor expectations.

All the results in Columns (1) to (3) can be explained by salience. Negative returns or returns during down markets are more salient than positive returns or returns during up markets, and contests with dispersed stock performance are more salient than those with similar performance, therefore affecting investor expectations to a larger extent. Consistent with this explanation, Garcia (2018) and Niessner and So (2018) show that financial press is more likely to cover negative stock market returns and individual stocks with deteriorating performance. Moreover, Reyes (2019) finds that investors pay more attention to negative stock market returns than comparable positive returns.

To directly test the impact of salience on expectation formation, the regression in Column (4) of Table 5 allows  $\lambda_1$  and  $\lambda_2$  to differ across salient ( $s$ ) versus nonsalient ( $ns$ ) past returns. To measure the salience level associated with the return of stock  $i$  in week  $t$ , we count the number of

news articles written on the stock in that week. We obtain the news coverage data from RavenPack. To ensure that extreme returns are not driving the variation in  $\lambda_1$  and  $\lambda_2$  mechanically, we orthogonalize news coverage with respect to the absolute returns. Specifically, for each week in the sample period, we run a cross-sectional regression of the total amount of news coverage of a stock on its absolute return over the same week. We then define a stock return as salient (nonsalient) if the residual from this cross-sectional regression is above (below) median.

Column (4) reports the empirical estimates of  $\lambda_{1,s}$ ,  $\lambda_{2,s}$ ,  $\lambda_{1,ns}$ , and  $\lambda_{2,ns}$ . We find that investor expectations respond more strongly to salient returns— $\lambda_{1,s}$  is much larger than  $\lambda_{1,ns}$ —and salient returns have a more persistent impact on the current expectation than nonsalient returns— $\lambda_{2,s}$  is significantly higher than  $\lambda_{2,ns}$ . Our result is also consistent with an attention story: news coverage draws investor attention to the stock return, making it more salient and hence better encoded into investor expectations.

### 3.5. User and stock characteristics

So far we have been studying the heterogeneity of  $\lambda_1$  and  $\lambda_2$  for past returns of different characteristics. Our cross-sectional setting allows us to study heterogeneity in expectation formation along other dimensions: we can link return expectations to different user and stock characteristics. Panel A of Table 6 estimates  $\lambda_1$  and  $\lambda_2$  separately for financial professionals and nonprofessionals.

**Table 6**

Extrapolative beliefs: user and stock characteristics.

The table in Panel A presents the results of contest-level nonlinear regressions for financial professionals versus nonprofessionals. The regression is specified in Eq. (2) of the main text. The dependent variable is a stock's consensus ranking averaged across professional users (Column (1)) or nonprofessional users (Column (2)). The highest ranked stock receives a score of ten, and the lowest ranked stock receives a score of one. The explanatory variables include lagged contest-adjusted returns from week  $t - 11$  to week  $t$ . The table in Panel B analyzes how the extrapolation parameters vary with firm characteristics. For each stock  $i$ , we estimate  $\lambda_{i,1}$  and  $\lambda_{i,2}$  from the following specification:

$$\text{Forcerank}_{i,t} = \lambda_{i,0} + \lambda_{i,1} \cdot \sum_{s=0}^n w_{i,s} R_{i,t-s} + \varepsilon_{i,t},$$

$$\text{where } w_{i,s} = \frac{\lambda_{i,2}^s}{\sum_{j=0}^n \lambda_{i,2}^j}, \quad 0 \leq \lambda_{i,2} < 1.$$

Firms that appear for less than 30 weeks in our sample are removed from this analysis. The dependent variable is  $\lambda_{i,1}$  in Column (1) and  $\lambda_{i,2}$  in Column (2). The explanatory variables include (1)  $\ell n$  (Size): the logarithm of a firm's market capitalization at the end of year 2015; (2)  $B/M$ : the book-to-market ratio, which is a firm's book equity of the fiscal year 2015 divided by its market value of the equity at the end of year 2015; (3) volatility: a firm's average weekly return volatility during our sample period; and (4) turnover: a firm's average weekly turnover during our sample period. The standard errors are in parentheses. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A: User characteristics		
	(1)	(2)
Dependent variable:	Professionals	Nonprofessionals
$\lambda_0$	5.498*** (0.029)	5.494*** (0.028)
$\lambda_1$	26.35*** (2.784)	33.77*** (2.055)
$\lambda_2$	0.773*** (0.035)	0.552*** (0.030)
Observations	9,658	10,261
R-squared	0.015	0.050
Panel B: Stock characteristics		
	(1)	(2)
Dependent variable:	$\lambda_{i,1}$	$\lambda_{i,2}$
$\ell n$ (Size)	8.173** (4.074)	0.066*** (0.024)
$B/M$	7.034 (14.30)	-0.153* (0.083)
Volatility	-12.03** (5.221)	-0.014 (0.030)
Turnover	1.548 (7.602)	0.113** (0.044)
Observations	137	137
R-squared	0.174	0.101

Interestingly, between professional and nonprofessional users, the extrapolation parameters are quite different. We find that financial professionals have a  $\lambda_1$  of 26.35, which is lower than that of the nonprofessionals (33.77), suggesting that professionals rely less on past stock returns when forming expectations about stock returns over the next week. Moreover, professionals have a  $\lambda_2$  of 0.773, which is higher than that of the nonprofessionals (0.552). This result suggests that nonprofessionals display a stronger degree of extrapolation, as they overweight recent past returns more strongly. The weight that nonprofessionals

put on returns decays by about 90% one month into the past; in comparison, the weight applied by professionals takes more than two months to decay by 90%.

Next, we examine how the extrapolation parameters vary across different stocks. First, we estimate the belief parameters,  $\lambda_{i,1}$  and  $\lambda_{i,2}$ , for each stock  $i$ . We then regress  $\lambda_{i,1}$  and  $\lambda_{i,2}$  on firm characteristics: size, book-to-market ratio, return volatility, and turnover. Panel B of Table 6 presents these results.

We find that the market capitalization of a firm is positively related to both  $\lambda_{i,1}$  and  $\lambda_{i,2}$ . One possible explanation of this finding is that data from larger firms are more visible or accessible to investors (Begenau et al., 2018). As a result, salience implies that information about larger firms plays a bigger role in the formation of investor expectations. We also find that a firm's turnover—a measure that is positively related to the firm's size—positively affects  $\lambda_{i,2}$ . In addition, a firm's return volatility averaged across all weeks in our sample period is negatively related to  $\lambda_{i,1}$ : higher volatility of a stock's past returns makes it more difficult for investors to identify a trend in the stock price, or it reduces investors' confidence in perceiving a trend, therefore reducing their degree of extrapolation. Finally,  $\lambda_{i,2}$  is higher for growth stocks than for value stocks.

How does an investor's degree of extrapolation change over time? To address this question, we take a closer look at the 35 most active users who regularly participated in the contests. For these users, we can reliably estimate  $\lambda_1$  and  $\lambda_2$  at the individual level and examine the time-series evolution of these extrapolation parameters. We find that almost all of these active users are extrapolators; only one is a "contrarian" with a significantly negative  $\lambda_1$ . We also find that, among these users, extrapolative beliefs do not seem to diminish with time. In fact, both  $\lambda_1$  and  $\lambda_2$  generally increased over time, possibly because weekly returns became more salient as users started to follow these stocks regularly on Forcerank.

To summarize this section, we analyze the Forcerank data and find strong evidence that individuals extrapolate from recent past returns when forming expectations about future stock returns, especially when recent returns are negative, more dispersed within a contest, or salient. Such extrapolative beliefs are stronger among nonprofessionals and stocks with certain characteristics. Overall, our findings suggest that salience is an important source for return extrapolation. As an external validation, we show a quantitatively similar pattern of return extrapolation among retail investors who trade frequently, suggesting that our findings are not particularly driven by features unique to the crowdsourcing platform. With our observations on expectation formation in hand, a natural follow-up question to examine is whether the return expectations from Forcerank users are accurate or systematically biased. We address this question in the next section.

#### 4. Return predictability

In this section, we study the asset pricing implications of investor expectations. First, we examine the accuracy or plausibility of the return expectations from Forcerank users. We show that the consensus Forcerank score signifi-

icantly predicts future stock returns with a negative sign. Moreover, we decompose the Forcerank score into two components: a predicted component explained by past returns and the residual component that is orthogonal to past returns. We find that both components negatively predict future stock returns. These return predictability results suggest that the beliefs of Forcerank users are systematically biased.

Next, we make the assumption that the weekly forecasts from Forcerank users represent the thinking process of a broader group of behavioral investors in the market. This assumption allows us to study the asset pricing implications of biased beliefs, and it is consistent with the finding (presented in Fig. 2) that the extrapolative patterns of return expectations from Forcerank users and a large group of short-term retail traders are quantitatively similar. Moreover, a recent study by Giglio et al. (2020) uses surveys to elicit beliefs of a large panel of retail investors who have substantial wealth invested in financial markets and shows that these self-reported beliefs indeed affect investors' portfolio choices. Given this assumption, we present a simple model in which a fraction of investors have extrapolative beliefs about stock returns. The model makes specific predictions regarding the heterogeneity of return predictability in the cross-section. We empirically test and confirm these predictions.

Finally, we evaluate the economic magnitude of the return predictability of extrapolative beliefs using trading strategies, both in sample among stocks covered on Forcerank and out of sample among all stocks over a longer period. Across different specifications, our trading strategies generate risk-adjusted profits that are economically significant.

#### 4.1. Return predictability of consensus beliefs

We first examine whether the consensus beliefs of Forcerank users are accurate or systematically biased. We address this question using Fama-MacBeth forecasting regressions, in which the dependent variable is the daily return of an individual stock over the next week. Panel A of Table 7 reports these regression results.

As Column (1) shows, the consensus Forcerank scores significantly predict the next week's stock returns with a negative sign. This return predictability can arise from extrapolative beliefs—investors form expectations about future returns by extrapolating from recent past returns—or it can arise from general “sentiment” above and beyond return extrapolation. To understand the source of the return predictability, we further decompose the Forcerank score into two components: a predicted score and the residual. The predicted score is computed as the fitted value from the nonlinear regression in Eq. (2) using contest-adjusted past returns as the explanatory variables. In other words, it is the weighted average of past 12 weekly returns that best predicts the Forcerank score. The residual of this regression is labeled as the residual score.

Column (2) shows that the predicted score also significantly predicts future stock returns with a negative sign. The magnitude of the regression coefficient on the predicted score is slightly greater than that on the raw

Forcerank score from Column (1). Also note that, although past returns altogether explain only about 6% of the variation in the Forcerank score, they contribute significantly to the predictive power of the Forcerank score for future stock returns.

Of course, a large literature on the short-term return reversal has already shown that the past return of a stock negatively predicts its future return and this reversal can be driven by liquidity shocks unrelated to return extrapolation (see Jegadeesh and Titman, 1995 and Campbell et al., 1993, among others). Given this literature, a natural question is whether the predictive power of the Forcerank score simply reflects liquidity-shock-induced return reversals. A priori, we do not expect liquidity shocks to be the main explanation for return predictability because stocks in our sample tend to be very large stocks, as seen in Table 2.

To directly address this question, we examine the short-term return reversal explicitly in regressions. For an apple-to-apple comparison, we convert each stock's contest-adjusted past return into a decile rank.<sup>13</sup> The contest-adjusted return is either over the past one-week return ( $Ret(t)$ ), the past one-month return ( $Ret(t-3, t)$ ), or the past one-quarter return ( $Ret(t-11, t)$ ). Contest adjustment effectively controls for the industry-level return, hence making past returns more likely to predict future return reversals.<sup>14</sup> Nonetheless, Columns (4) and (6) show that neither the past one-week return nor the past one-quarter return has significant predictive power for the next week's return, even after contest adjustment. Column (5) shows that the past one-month return has significant predictive power for the next week's return. Overall, the evidence suggests that only a weak standard short-term return reversal is present in our sample. More important, Columns (7) and (8) show that the Forcerank score and the predicted score both drive out past-return measures when they are included in the same regression.

Finally, we examine the predictive power of the residual score for future stock returns. By construction, the residual score is orthogonal to past returns. Interestingly, Columns (3) and (9) show that the residual score also negatively predicts the next week's return, with or without controlling for past returns. This finding suggests that the predictive power of the Forcerank score is not completely driven by its association with past returns. The Forcerank score contains additional information about investor “sentiment” above and beyond return extrapolation.

Panel B of Table 7 contains various robustness checks where Forcerank-based score variables are horseshod against continuous variables of past returns and dummy variables that represent extreme past returns in the top or the bottom decile; Atkins and Dyl (1990) and Kumar et al. (2019), among others, demonstrate that extreme winners and losers can disproportionate

<sup>13</sup> This conversion allows us to better compare the regression results with those using the Forcerank (ranking) score, the predicted score, or the residual score as the explanatory variable.

<sup>14</sup> Da et al. (2013) show that, compared to sorting stocks across industries, sorting stocks within an industry gives rise to stronger return reversals.



**Table 7**

Return predictability: Fama-MacBeth regressions.

This table presents the results of Fama-MacBeth forecasting regressions. For each week  $t$  and each stock  $i$ , the dependent variable is the daily return of stock  $i$  over week  $t + 1$ . Panel A compares various score variables as return predictors. They include the consensus Forcerank score, the predicted score, the residual score, and a decile rank based on the stock's contest-adjusted return over the past one week, one month, and one quarter ( $\text{Ret}(t)$  score,  $\text{Ret}(t - 3, t)$  score, and  $\text{Ret}(t - 11, t)$  score, respectively). The consensus Forcerank score is the average of the Forcerank consensus rankings of the same stock across contests. The predicted score is computed as the fitted value from the nonlinear regression in Eq. (2) of the main text using the consensus Forcerank score defined above as the dependent variable and using contest-adjusted past returns as the explanatory variables. The residual of this regression is labeled as the residual score. Panel B contains robustness checks; Forcerank-based score variables are horseshoed against continuous variables of past returns and dummy variables that represent extreme past returns in the top or the bottom decile. In addition, stock characteristics such as size and book-to-market ratio are also controlled for. Returns are in daily percent, and the  $t$ -statistics are in parentheses. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A: Score variables									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable:	Daily return in week $t + 1$								
Forcerank score	-0.0246*** (-2.75)						-0.0293*** (-3.22)		
Predicted score		-0.0321*** (-4.22)						-0.0716*** (-4.41)	
Residual score			-0.0162* (-1.86)						-0.0238*** (-2.71)
$\text{Ret}(t)$ score				0.0025 (0.25)			-0.0221 (-1.02)	0.0245 (1.32)	-0.0255 (-1.18)
$\text{Ret}(t - 3, t)$ score					-0.0267** (-2.30)		0.0049 (0.24)	0.0264 (1.20)	-0.0011 (-0.06)
$\text{Ret}(t - 11, t)$ score						-0.0051 (-0.40)	0.0058 (0.44)	0.0032 (0.22)	0.0084 (0.64)
Observations	59,929	59,929	59,929	59,929	59,929	59,929	59,929	59,929	59,929
R-squared	0.019	0.013	0.018	0.024	0.027	0.035	0.096	0.094	0.096
Panel B: Robustness checks									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable:	Daily return in week $t + 1$								
Forcerank score	-0.0285*** (-3.36)	-0.0149* (-1.88)	-0.0164** (-2.01)						
Predicted score				-0.0831*** (-5.65)	-0.0358*** (-2.60)	-0.0479*** (-4.76)			
Residual score							-0.0179** (-2.17)	-0.0139* (-1.79)	-0.0089 (-1.10)
$\text{Ret}(t)$	-0.0093 (-0.81)			0.0193 (1.55)			-0.0133 (-1.20)		
Top decile of $\text{Ret}(t)$	0.200** (1.99)			0.0457 (0.63)			0.204* (1.95)		
Bottom decile of $\text{Ret}(t)$	-0.116 (-1.29)			-0.119 (-1.32)			-0.111 (-1.25)		
$\text{Ret}(t - 3, t)$		-0.0023 (-0.47)			0.0058 (0.97)			-0.0030 (-0.61)	
Top decile of $\text{Ret}(t - 3, t)$		-0.436*** (-5.58)			-0.479*** (-6.26)			-0.441*** (-5.57)	
Bottom decile of $\text{Ret}(t - 3, t)$		-0.0055 (-0.06)			-0.0184 (-0.21)			-0.0036 (-0.04)	
$\text{Ret}(t - 11, t)$			0.0027 (1.21)			0.0036 (1.47)			0.0034 (1.44)
Top decile of $\text{Ret}(t - 11, t)$			-0.427*** (-4.06)			-0.323*** (-3.51)			-0.445*** (-4.12)
Bottom decile of $\text{Ret}(t - 11, t)$			-0.0365 (-0.46)			-0.0954 (-1.17)			-0.0361 (-0.45)
$\ln(\text{Size})$	-0.0643*** (-3.19)	-0.0779*** (-3.64)	-0.0644*** (-3.29)	-0.0909*** (-3.79)	-0.0810*** (-3.59)	-0.0716*** (-3.38)	-0.0679*** (-3.34)	-0.0786*** (-3.66)	-0.0687*** (-3.49)
$\ln(B/M)$	0.0311 (0.82)	0.0351 (0.88)	0.0194 (0.46)	0.0269 (0.76)	0.0607* (1.71)	0.0265 (0.72)	0.0302 (0.79)	0.0373 (0.93)	0.0197 (0.47)
Observations	59,929	59,929	59,929	59,929	59,929	59,929	59,929	59,929	59,929
R-squared	0.138	0.139	0.149	0.135	0.135	0.146	0.137	0.139	0.148

ately drive short-term return reversals.<sup>15</sup> In addition, stock characteristics such as size and book-to-market ratio are also controlled for in the regressions. The predictive power of the Forcerank score and the predicted score remains significant with all the controls; it is not driven by extreme past returns over any specific horizons. The horserace in Column (6) is particularly interesting. The past one-quarter return can be viewed as an equal-weighted average of past 12 weekly returns, whereas the predicted score is a weighted average of the same past 12 weekly returns with the weights calibrated to extrapolative beliefs. The fact that the predicted score drives out the past one-quarter return supports the predictions of the extrapolation model we describe in [Appendix A](#). The last three columns show that the predictive power of the residual score remains significant except when the past one-quarter return and its extreme value dummies are present.

In unreported tests, we further examine the fundamental predictability of Forcerank scores. After controlling for past returns and past fundamentals (proxied by analyst earnings forecast revisions and earnings surprises), we find Forcerank scores to predict neither the standardized earnings surprise (SUE) nor the analyst-consensus-based earnings surprise in the next quarter. The lack of fundamental predictability, combined with Forcerank scores' negative return predictability, reinforces the notion that the beliefs of Forcerank users are systematically biased.

In summary, [Table 7](#) shows that the Forcerank score, its component that is related to past returns, and the residual component all significantly predict the next week's stock returns with a negative sign. In addition, these return predictability results remain strong after controlling for returns over the past one week, one month, and one quarter, among other variables. Altogether, our results show that the beliefs of Forcerank participants are systematically biased: when these participants are optimistic about future stock returns, returns tend to be low, on average; conversely, when they are pessimistic about future stock returns, returns tend to be high, on average. As such, extrapolative expectation can be another important contributor to the well-documented short-term return reversal phenomenon, consistent with [Subrahmanyam \(2005\)](#).

#### 4.2. Heterogeneity of return predictability

Given that our sample only includes 1,045 distinct Forcerank users, we do not claim that these users alone move stock prices, nor that they represent all the market participants. Instead, we interpret our evidence as suggesting that the beliefs of these Forcerank users represent the thinking process of a broader group of behavioral investors in the market.

By assuming that the beliefs of the Forcerank users are representative of the beliefs of a broader investor population, we can now study the asset pricing implications of

these biased beliefs. To do so, we develop a cross-sectional model of return extrapolation. The model features two types of agent, extrapolators and fundamental traders. Extrapolators form expectations about the future stock returns by extrapolating from the recent past returns, and they trade stocks based on these extrapolative beliefs. Fundamental traders, on the other hand, serve as arbitrageurs who correct for mispricing. We leave the detail of the model in [Appendix A](#).

Importantly, the model makes two specific predictions regarding the cross-sectional heterogeneity of return predictability. First, return predictability should be stronger among stocks whose clienteles are dominated by behavioral extrapolators. Second, return predictability should also be stronger among stocks with a higher degree of extrapolation—this is measured by  $\lambda_{i,1}(1 - \lambda_{i,2})$  for stock  $i$  in the model. In this section, we empirically test and confirm these two predictions.

First, we use institutional ownership to measure a stock's clientele of fundamental traders: stocks whose clienteles are dominated by extrapolators are assumed to have low institutional ownership. In Panel A of [Table 8](#), we run Fama-MacBeth forecasting regressions separately for stocks with below-median institutional ownership—we assume these stocks are traded more by extrapolators—and for stocks with above-median institutional ownership—we assume these stocks are traded less by extrapolators. Consistent with the model prediction, our results show that the return predictability of the Forcerank score and the predicted score is only present among stocks with low institutional ownership.

Second, for each stock  $i$ , we estimate  $\lambda_{i,1}$  and  $\lambda_{i,2}$  based on the nonlinear regression in [Eq. \(2\)](#). We then follow the model and use  $\lambda_{i,1}(1 - \lambda_{i,2})$  to measure stock  $i$ 's degree of extrapolation. Panel B of [Table 8](#) runs Fama-MacBeth regressions separately for stocks with below-median degree of extrapolation and for stocks with above-median degree of extrapolation. The results confirm the model prediction that the Forcerank score and the predicted score both have much stronger predictive power for the future returns of stocks that have a higher degree of extrapolation.

#### 4.3. Trading strategies

To evaluate the economic significance of our return predictability results, we form trading strategies. At the beginning of each week, we sort the stocks into five quintiles based on different variables. The portfolio is rebalanced every week. Stocks with prices below \$5 at the beginning of the week are removed to reduce the impact of illiquidity. Panel A of [Table 9](#) presents our results.

Row (1) sorts stocks based on the consensus Forcerank scores. It shows that Forcerank scores negatively predict future stock returns: the low-score-minus-high-score return spread is 8.11 bps per day ( $t$ -value of 2.33).<sup>16</sup> The

<sup>15</sup> Extreme winners and losers can be particularly salient to investors due to regular coverage by financial media. As a result, the attention-induced price pressure can amplify return reversals.

<sup>16</sup> Consistent with our prior analyses, the Forcerank scores are computed using all individual rankings during a contest, including those submitted after the closing time on Friday. Removing these late rankings does not alter our return predictability results but makes our trading strategy implementable in real time.

**Table 8**

Return predictability: heterogeneity.

This table presents the results of Fama-MacBeth forecasting regressions. For each week  $t$  and each stock  $i$ , the dependent variable is the daily return of stock  $i$  over week  $t + 1$ . The explanatory variables include the Forcerank score, the predicted score, and a decile rank based on the stock's contest-adjusted return over the past one week, one month, and one quarter (Ret( $t$ ) score, Ret( $t - 3, t$ ) score, and Ret( $t - 11, t$ ) score, respectively). The consensus Forcerank score is the average of the Forcerank consensus rankings of the same stock across contests. The predicted score is computed as the fitted value from the nonlinear regression in Eq. (2) of the main text using the consensus Forcerank score defined above as the dependent variable and using contest-adjusted past returns as the explanatory variables. In Panel A, the stocks covered on Forcerank are partitioned into two groups based on institutional ownership that we obtained from the Thomson-Reuters Institutional Holdings (13F) Database and measured at the end of December 2015; ownership is set to zero if no institution in the database reports its ownership of the stock. Stocks with low (high) institutional ownership have a below-median (above-median) fraction of shares owned by institutions. In Panel B, the stocks covered on Forcerank are partitioned into two groups based on the degree of extrapolation—for each stock  $i$ , this is measured by  $\lambda_{i,1}(1 - \lambda_{i,2})$ . Returns are in daily percent, and the  $t$ -statistics are in parentheses. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A: Institutional ownership				
	(1)	(2)	(3)	(4)
Dependent variable:	Daily return in week $t + 1$			
	Low IO	High IO	Low IO	High IO
Forcerank score	-0.0398*** (-3.13)	-0.0125 (-1.12)		
Predicted score			-0.0880*** (-4.28)	-0.0062 (-0.30)
Ret( $t$ ) score	-0.0132 (-0.86)	0.0230 (1.50)	0.0408** (2.45)	0.0146 (0.77)
Ret( $t - 3, t$ ) score	-0.0023 (-0.33)	-0.0057 (-0.66)	0.0052 (0.27)	-0.0455** (-1.99)
Ret( $t - 11, t$ ) score	-0.0029 (-0.30)	0.0018 (0.41)	0.0398** (2.14)	-0.0061 (-0.34)
Observations	30,014	29,915	30,014	29,915
R-squared	0.148	0.176	0.135	0.171
Panel B: Degree of extrapolation				
	(1)	(2)	(3)	(4)
Dependent variable:	Daily return in week $t + 1$			
	Low degree	High degree	Low degree	High degree
Forcerank score	-0.0196* (-1.72)	-0.235** (-2.28)		
Predicted score			-0.0558*** (-3.11)	-0.161*** (-4.89)
Ret( $t$ ) score	0.0243** (2.27)	-0.0112 (-0.51)	0.0655*** (3.95)	0.0962*** (3.35)
Ret( $t - 3, t$ ) score	-0.0424*** (-6.85)	-0.0154 (-1.38)	-0.0456** (-2.46)	0.0626** (2.24)
Ret( $t - 11, t$ ) score	0.0166*** (4.59)	0.0359** (2.23)	0.0221 (1.21)	-0.0603** (-2.28)
Observations	19,617	18,730	19,617	18,730
R-squared	0.160	0.234	0.150	0.237

return spread remains significant after risk adjustments using the CAPM, the Fama-French five-factor model, or the five-factor model augmented with the momentum factor and the short-term reversal factor. Moreover, our estimate of direct transaction cost for our trading strategies, accounting for the bid-ask spread for the round-trip trades with an average weekly portfolio turnover rate of 50%, is 1.7 bps per day, which is far below the risk-adjusted return spread of 7 bps per day. Therefore, the profits to our trading strategies are likely to survive transaction costs.

Row (2) sorts stocks based on the predicted scores. It shows that predicted scores also negatively predict future stock returns: the low-score-minus-high-score return spread is 6.51 bps per day ( $t$ -value of 2.01). The return spread remains significant even after controlling for the Fama-French five factors, the momentum factor, and the

short-term reversal factor. The seven-factor alpha is still 5.47 bps per day ( $t$ -value of 1.70).

Row (3) shows that the predictive power of the residual score for future stock returns is slightly larger than that of the predicted score. The low-score-minus-high-score return spread is 6.89 bps per day ( $t$ -value of 2.07). The return spread remains significant after controlling for the Fama-French five factors, the momentum factor, and the short-term reversal factor. The seven-factor alpha is 6.67 bps per day ( $t$ -value of 2.01). In other words, the Forcerank score contains information about investor sentiment above and beyond return extrapolation.

Rows (4) and (5) show that the standard short-term return reversals are actually not economically significant in our sample. Neither sorting on past one-week returns nor sorting on past one-month returns generates a significant return spread, even though Table 7 Panel A showed that

**Table 9**

Return predictability: trading strategies.

This table shows daily calendar-time portfolio returns. Panel A examines stocks covered by the Forcerank platform. At the beginning of every calendar week, the stocks are ranked in an ascending order into five quintiles on the basis of their consensus Forcerank scores, the predicted scores (the fitted values from the nonlinear regression in Eq. (2) of the main text using contest-adjusted past returns as the explanatory variables), the residual scores (the difference between the Forcerank scores and the predicted scores), the contest-adjusted return over the past week, and the contest-adjusted return over the past month. “L” in Column (1) corresponds to the bottom quintile, and “H” in Column (2) corresponds to the top quintile. The stocks are equally weighted within a given portfolio, and the portfolio is rebalanced every calendar week. Calendar-time alphas are estimated using raw returns, the CAPM, the Fama-French five-factor model alone, and the Fama-French five-factor model augmented with the momentum factor and the short-term reversal factor. Panels B and C examine an out-of-sample period from April 9, 2001 to December 31, 2017. Panel B includes all stocks; Panel C includes the top size quintile stocks based on the CRSP cap-based portfolio assignment. Panels B and C sort stocks into ten deciles instead of five quintiles. Also, for these two panels, the predicted scores are constructed using industry-adjusted returns over the past 12 weeks. The explanatory variables also include “predicted score PN,” the fitted value from the nonlinear regression in Eq. (3) of the main text: this fitted value allows for the asymmetry between positive and negative past returns when estimating  $\lambda_1$  and  $\lambda_2$ . Stocks with prices below \$5 per share at the beginning of each calendar week are removed from the sample. Returns are in daily percent, and the  $t$ -statistics are in parentheses. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% level, respectively.

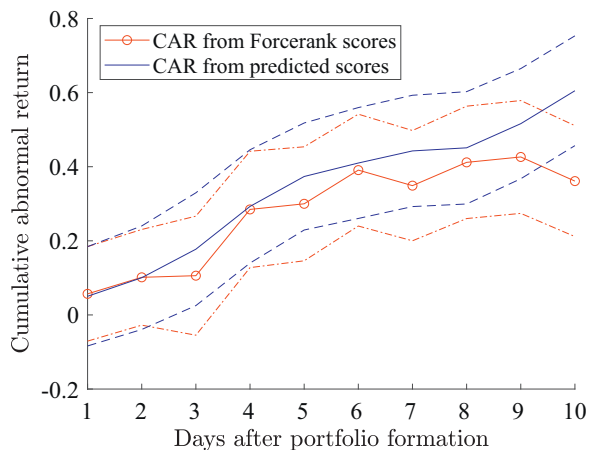
	(1)	(2)	(3)	(4)	(5)	(6)
	<i>L</i>	<i>H</i>	<i>L</i> – <i>H</i> Raw	<i>L</i> – <i>H</i> CAPM	<i>L</i> – <i>H</i> FF5	<i>L</i> – <i>H</i> FF5 + Mom + Rev
<b>Panel A: In sample</b>						
Forcerank score	0.155*** (3.07)	0.0740* (1.67)	0.0811** (2.33)	0.0680* (1.96)	0.0706** (2.16)	0.0705** (2.14)
Predicted score	0.137*** (2.88)	0.0666 (1.39)	0.0651** (2.01)	0.0675** (2.07)	0.0679** (2.08)	0.0547* (1.70)
Residual score	0.154*** (3.03)	0.0853* (1.81)	0.0689** (2.07)	0.0635* (1.89)	0.0670** (2.03)	0.0667** (2.01)
Contest-adjusted Ret( <i>t</i> )	0.128*** (2.70)	0.0803 (1.54)	0.0367 (1.09)	0.0441 (1.31)	0.0429 (1.28)	0.0351 (1.05)
Contest-adjusted Ret( <i>t</i> – 3, <i>t</i> )	0.117** (2.41)	0.0806 (1.56)	0.0445 (1.25)	0.0455 (1.26)	0.0462 (1.30)	0.0301 (0.86)
<b>Panel B: Out of sample</b>						
	(1)	(2)	(3)	(4)	(5)	(6)
	<i>L</i>	<i>H</i>	<i>L</i> – <i>H</i> Raw	<i>L</i> – <i>H</i> CAPM	<i>L</i> – <i>H</i> FF5	<i>L</i> – <i>H</i> FF5 + Mom + Rev
Predicted score	0.242*** (9.04)	–0.0025 (–0.11)	0.244*** (14.29)	0.238*** (14.34)	0.246*** (14.91)	0.215*** (16.59)
Predicted score PN	0.246*** (8.95)	–0.0001 (–0.00)	0.246*** (13.78)	0.238*** (13.94)	0.255*** (15.38)	0.231*** (16.09)
Industry-adjusted Ret( <i>t</i> )	0.224*** (8.44)	0.0027 (0.12)	0.221*** (12.13)	0.216*** (12.07)	0.223*** (12.55)	0.198*** (12.55)
Industry-adjusted Ret( <i>t</i> – 3, <i>t</i> )	0.200*** (7.13)	0.0285 (1.29)	0.172*** (9.08)	0.164*** (9.02)	0.175*** (9.73)	0.136*** (11.01)
<b>Panel C: Out-of-sample large cap</b>						
Predicted score	0.114*** (3.29)	0.0106 (0.36)	0.104*** (4.14)	0.0967*** (3.92)	0.110*** (4.50)	0.0671*** (3.36)
Predicted score PN	0.156*** (3.58)	0.0185 (0.85)	0.137*** (4.18)	0.120*** (3.91)	0.150*** (5.12)	0.121*** (4.50)
Industry-adjusted Ret( <i>t</i> )	0.113*** (3.17)	0.0208 (0.70)	0.0921*** (3.34)	0.0843*** (3.11)	0.0984*** (3.68)	0.0608** (2.56)
Industry-adjusted Ret( <i>t</i> – 3, <i>t</i> )	0.0842** (2.29)	0.0252 (0.87)	0.0590** (2.04)	0.0489* (1.74)	0.0633** (2.27)	0.0069 (0.34)

past one-month returns negatively predict future returns at the 5% significance level.

Next, we examine how the performance of the trading strategies evolves over time. Fig. 4 plots the daily cumulative abnormal returns (CARs) over the first two weeks after forming the trading strategy based on either the Forcerank score (solid-circle line) or the predicted score (solid line). The daily abnormal returns are estimated using the Fama-French five-factor model augmented with the momentum factor and the short-term reversal factor. The

CAR plot demonstrates that returns for trading strategies accrue gradually over time, instead of coming exclusively from the first day since portfolio formation. As such, our return predictability results clearly go beyond the bid-ask bounce and other market microstructure effects.

Finally, we examine the generalizability of our return predictability results. We conduct an out-of-sample validation test by studying return predictability among all stocks—not just those covered by the Forcerank platform—over a longer period from April 9, 2001 to December



**Fig. 4.** Trading strategies: daily CAR plot. The figure plots the daily cumulative abnormal returns (CARs) over the first two weeks after forming the trading strategy based on either the Forcerank score or the predicted score. At the beginning of every calendar week, the stocks are ranked in an ascending order into five quintiles on the basis of their consensus Forcerank scores or the predicted scores (the fitted values from the nonlinear regression in Eq. (2) of the main text using contest-adjusted past returns as the explanatory variables) over the past week. The solid-circle (solid) line corresponds to the trading strategy that buys stocks with low Forcerank scores (predicted scores) and sells stocks with high Forcerank scores (predicted scores). The stocks are equally weighted within a given portfolio, and the portfolio is rebalanced every calendar week. The daily abnormal returns are estimated using the Fama-French five-factor model augmented with the momentum factor and the short-term reversal factor. The dashed lines and the dash-dot lines correspond to the 95% confidence level.

31, 2017.<sup>17</sup> If the beliefs of Forcerank users represent the thinking process of a broader group of behavioral investors in the market, we would expect that the predicted scores for non-Forcerank stocks also have predictive power for the future returns of these stocks.

For each stock in each week, we first compute a predicted score as the fitted value from the nonlinear regression in Eq. (2); we use the stock's industry-adjusted returns over the past 12 weeks as the explanatory variables, and we use the estimates of  $\lambda_1$  and  $\lambda_2$  from Column (2) of Table 4 to construct the weights in Eq. (2). We also compute a second predicted score (called "PN") as the fitted value from the nonlinear regression in Eq. (3), allowing for the asymmetry between positive and negative past returns when estimating  $\lambda_1$  and  $\lambda_2$ ; we use the estimates of  $\lambda_{1,p}$ ,  $\lambda_{2,p}$ ,  $\lambda_{1,n}$ , and  $\lambda_{2,n}$  from Column (1) of Table 5 to construct the weights in Eq. (3). To evaluate the economic magnitude of the return predictability, we examine trading strategies that are similar to those in Panel A. Stocks with prices below \$5 at the beginning of the week are removed to reduce the impact of illiquidity. The results are reported in Panels B and C of Table 9.

<sup>17</sup> The starting date of the out-of-sample period is the date of full implementation of decimalization for all equities and options on exchanges. We chose this date to alleviate the concern that our analysis simply captures the short-term return reversal due to the bid-ask bounce or other liquidity issues.

Row (1) of Panel B sorts all the stocks based on their predicted scores and reports the top- and bottom-decile portfolio performance of the daily return over the next week. The low-score-minus-high-score return spread is 24.4 bps per day ( $t$ -value of 14.29). The spread remains highly significant after various risk adjustments. Row (2) shows that allowing the asymmetry between positive and negative past returns in constructing the predicted scores further increases the return spread.

As a comparison, Rows (3) and (4) of Panel B report the performance of the standard industry-neutral short-term return reversal strategies that sort on past one-week returns or past one-month returns. Although the return spreads produced by these trading strategies are also statistically significant, they are smaller in magnitude relative to those in Rows (1) and (2). In other words, extrapolative beliefs, by applying declining weights to past weekly returns and by allowing for different weights on positive and negative past returns, predict future returns better than past one-week returns and past one-month returns.

To further alleviate the concern that our return predictability is simply due to liquidity shocks, we repeat the trading strategies from Panel B but only among the largest stocks (those in the top CRSP size quintile). These stocks are least likely to be affected by illiquidity. Panel C shows that, even among this subset of large-cap stocks, the predicted scores still outperform past one-week and one-month returns.

## 5. Conclusion

Taking advantage of novel data from Forcerank, a crowdsourcing platform for ranking stocks, we provide strong empirical evidence that investors extrapolate from recent past returns of individual stocks when forming expectations about future returns. We then study how investors extrapolate. We find that extrapolation is asymmetric between positive and negative past returns: investors put more weight on negative past returns, and this weight decays more slowly into the past for these negative returns. The weight also decays more slowly for a more dispersed cross-section of past returns. Finally, investor expectations respond more strongly to salient past returns, and salient returns from both the recent past and the distant past affect investor expectations. Moreover, we examine the effect of user and firm characteristics on expectation formation. We find a stronger degree of extrapolation among users who are not financial professionals. We also find that extrapolation is affected by firm characteristics such as size, book-to-market ratio, return volatility, and turnover.

Next, we examine whether the return expectations from Forcerank users are accurate or systematically biased. We show that the consensus Forcerank score significantly predicts future stock returns with a negative sign. Furthermore, we decompose the Forcerank score into two components: a predicted component explained by past returns and the residual component orthogonal to the past returns. We find that both components negatively predict future stock returns. Altogether, our results sug-



gest that beliefs of Forcerank users are systematically biased.

Finally, we examine additional asset pricing implications of these biased beliefs. In the cross-section, we find that return predictability of the Forcerank score is stronger among stocks with lower institutional ownership and a higher degree of extrapolation. This heterogeneity is consistent with the predictions of a cross-sectional model of return extrapolation. We also form weekly rebalanced trading strategies to evaluate the economic magnitude of the return predictability of extrapolative beliefs, both in sample among stocks covered by the Forcerank platform and out of sample among all stocks over a longer period. Across different specifications, the risk-adjusted profits generated by our trading strategies are economically significant.

### Appendix A. A cross-sectional model with return extrapolation

There are a number of extrapolation models that try to explain empirical facts about the aggregate stock market. However, few extrapolation models have been developed for the cross-section of individual stocks.<sup>18</sup> In this section, we study the asset pricing implications of a simple cross-sectional model that features some investors who extrapolate from a stock's recent past returns when forming beliefs about its future return.

We consider a finite-horizon economy with  $T + 1$  dates,  $t = 0, 1, \dots, T$ . There are  $N + 1$  assets: a risk-free asset whose interest rate is normalized to zero and  $N$  risky assets. Risky asset  $i$  is a claim to a single dividend payment at the terminal date, and the payment equals

$$D_{i,T} = D_{i,0} + \varepsilon_{i,1} + \dots + \varepsilon_{i,T}, \quad (\text{A.1})$$

where

$$\begin{aligned} \varepsilon_{i,t} &= \zeta_i \cdot \varepsilon_{M,t} + \eta_{i,t}, \\ \varepsilon_{M,t} &\sim \mathcal{N}(0, \sigma_M^2), \\ \eta_{i,t} &\sim \mathcal{N}(0, \sigma_i^2), \quad \text{i.i.d. over time and across stocks.} \end{aligned} \quad (\text{A.2})$$

The value of  $D_{i,0}$  is public information at time 0. Both the market-wide news  $\varepsilon_{M,t}$  and the firm-specific news  $\eta_{i,t}$  become public at time  $t$ . The fundamental news of risky asset  $i$  has a loading of  $\zeta_i$  on the market-wide news. The price of this asset,  $P_{i,t}$ , is endogenously determined in equilibrium, and its supply is fixed at  $Q_i$ .

There are two types of agent, fundamental traders and extrapolators. Fundamental traders make up a population fraction  $\mu^f$  of the economy, and extrapolators make up a population fraction  $\mu^e$  of the economy;  $\mu^e = 1 - \mu^f$ . Both types of agent maximize their expected utility defined over the next period's wealth with constant absolute risk aversion  $\gamma$ . The key behavioral assumption of the model is

<sup>18</sup> Barberis and Shleifer (2003) develop a cross-sectional extrapolation model to study comovement within and across investment styles. The focus of our model, however, is to study expectation formation and its asset pricing implications at the individual stock level.

that, for risky asset  $i$ ,

$$\mathbb{E}_t^e[\tilde{P}_{i,t+1} - P_{i,t}] = \lambda_{i,0} + \lambda_{i,1}S_{i,t}, \quad (\text{A.3})$$

where  $\lambda_{i,1} > 0$ ,  $\lambda_{i,2} \in (0, 1)$  and

$$S_{i,t} \equiv (1 - \lambda_{i,2}) \sum_{k=0}^{\infty} (\lambda_{i,2})^k (P_{i,t-k} - P_{i,t-k-1}). \quad (\text{A.4})$$

That is, extrapolators' time- $t$  expectation about the price change of risky asset  $i$  over the next period is a linear function of the (normalized) weighted average of all past price changes; we call this weighted average of past price changes "sentiment"  $S_{i,t}$ .<sup>19</sup> The parameter  $\lambda_{i,1}$  measures the overall effect of past price changes on extrapolator beliefs. The parameter  $\lambda_{i,2}$  measures the weight an extrapolator puts on recent price changes relative to distant price changes. Empirically, the Forcerank data allow us to estimate the belief parameters  $\lambda_{i,1}$  and  $\lambda_{i,2}$ .<sup>20</sup> We provide a detailed discussion of these parameters in Sections 3 and 4 of the main text.

Next, we derive the share demand for the two types of agent. We begin with fundamental traders. As mentioned above, each fundamental trader has a constant absolute risk aversion (CARA) utility function defined over her next period's wealth. At time  $t$ , she chooses her per capita share demand  $N_t^f$  on the risky assets to maximize

$$\mathbb{E}_t^f \left[ -e^{-\gamma (W_t^f + (\tilde{P}_{t+1} - P_t)' N_t^f)} \right], \quad (\text{A.5})$$

which implies

$$N_t^f = \frac{1}{\gamma} (\Sigma_t^f)^{-1} (\mathbb{E}_t^f[\tilde{P}_{t+1}] - P_t), \quad (\text{A.6})$$

where  $\Sigma_t^f$  is the variance-covariance matrix of the next period's price changes perceived by fundamental traders at time  $t$  and  $P_t = (P_{1,t}, P_{2,t}, \dots, P_{N-1,t}, P_{N,t})'$ . We assume

$$(\Sigma_t^f)_{i,j} = \Sigma_{i,j} \equiv \begin{cases} \zeta_i^2 \sigma_M^2 + \sigma_i^2 & i = j \\ \zeta_i \zeta_j \sigma_M^2 & i \neq j \end{cases}. \quad (\text{A.7})$$

That is, for simplicity, we assume that fundamental traders believe that the covariance for changes in price is the same as the covariance for changes in fundamentals.

Applying Eq. (A.6) at time  $T - 1$ , and noting that at the terminal date  $T$ ,  $P_T = D_T \equiv (D_{1,T}, D_{2,T}, \dots, D_{N-1,T}, D_{N,T})'$ , we obtain

$$N_{T-1}^f = \frac{1}{\gamma} \Sigma^{-1} (D_{T-1} - P_{T-1}). \quad (\text{A.8})$$

Eq. (A.8) and market clearing together imply

$$\mu^f \frac{1}{\gamma} \Sigma^{-1} (D_{T-1} - P_{T-1}) + \mu^e N_{T-1}^e = Q, \quad (\text{A.9})$$

<sup>19</sup> Since our economy begins at  $t = 0$ , we can also write sentiment as

$$S_{i,t} = (1 - \lambda_{i,2}) \sum_{k=0}^{t-1} (\lambda_{i,2})^k (P_{i,t-k} - P_{i,t-k-1}) + (\lambda_{i,2})^t S_{i,0},$$

where  $S_{i,0}$  represents the initial level of sentiment at  $t = 0$ , summarizing the weighted average of past price changes from  $t = -\infty$  to  $t = 0$ .

<sup>20</sup> When we use a consensus Forcerank score instead of  $\mathbb{E}_t^e[\tilde{P}_{i,t+1} - P_{i,t}]$  as the dependent variable in Eq. (A.3), we will be able to estimate  $\lambda_{i,1}$  up to an affine transformation.

where  $Q \equiv (Q_1, Q_2, \dots, Q_{N-1}, Q_N)'$  and  $N_{T-1}^e$  is extrapolators' per capita share demand on the risky assets at time  $T - 1$ . Rearranging terms gives

$$P_{T-1} = D_{T-1} - (\mu^f)^{-1} \gamma \Sigma (Q - \mu^e N_{T-1}^e). \quad (\text{A.10})$$

We further impose

$$\mathbb{E}_t^f(N_{t+1}^f) = Q. \quad (\text{A.11})$$

This is a bounded rationality assumption, which says that fundamental traders do not directly compute investors' future demands. Instead, they expect that all market participants will demand the per capita supply of the risky assets in the next period.

Eqs. (A.6), (A.10), and (A.11) together give

$$\begin{aligned} N_{T-2}^f &= \frac{1}{\gamma} \Sigma^{-1} (\mathbb{E}_{T-2}^f[\tilde{P}_{T-1}] - P_{T-2}) \\ &= \frac{1}{\gamma} \Sigma^{-1} (D_{T-2} - \gamma \Sigma Q - P_{T-2}). \end{aligned} \quad (\text{A.12})$$

Recursively, the time- $t$  per capita share demand of fundamental traders is

$$N_t^f = \frac{1}{\gamma} \Sigma^{-1} (D_t - \gamma (T - t - 1) \Sigma Q - P_t), \quad (\text{A.13})$$

where  $D_t \equiv (D_{1,t}, D_{2,t}, \dots, D_{N-1,t}, D_{N,t})'$  and  $D_{i,t} \equiv D_{i,0} + \varepsilon_{i,1} + \dots + \varepsilon_{i,t}$  for risky asset  $i$ .

We now derive the share demand of extrapolators. Each extrapolator has a CARA utility function defined over her next period's wealth. At time  $t$ , she chooses her per capita share demand  $N_t^e$  on the risky assets to maximize

$$\mathbb{E}_t^e \left[ -e^{-\gamma (W_t^e + (\tilde{P}_{t+1} - P_t) N_t^e)} \right], \quad (\text{A.14})$$

which implies

$$N_t^e = \frac{1}{\gamma} (\Sigma_t^e)^{-1} (\mathbb{E}_t^e[\tilde{P}_{t+1}] - P_t), \quad (\text{A.15})$$

where  $\Sigma_t^e$  is the variance-covariance matrix of the next period's price changes perceived by extrapolators at time  $t$ . We further assume<sup>21</sup>

$$\Sigma_t^e = \Sigma_t^f = \Sigma. \quad (\text{A.16})$$

Eqs. (A.3), (A.4), (A.15), and (A.16) together imply that the time- $t$  per capita share demand of extrapolators is

$$N_t^e = \frac{1}{\gamma} \Sigma^{-1} X_t, \quad (\text{A.17})$$

where  $X_t \equiv (\lambda_{1,0} + \lambda_{1,1} S_{1,t}, \lambda_{2,0} + \lambda_{2,1} S_{2,t}, \dots, \lambda_{N-1,0} + \lambda_{N-1,1} S_{N-1,t}, \lambda_{N,0} + \lambda_{N,1} S_{N,t})'$ .

Intuitively, Eq. (A.17) shows that extrapolator demand is positively related to the levels of sentiment: when stocks' recent past returns are high, extrapolators expect high stock returns moving forward, pushing up their current share demand. On the other hand, Eq. (A.13) indicates that fundamental traders serve as arbitrageurs who correct for mispricing: their share demand is positively related to the fundamental value of the risky assets but is negatively related to the risky asset prices.

The share demands (A.13) and (A.17), together with market clearing, imply that the equilibrium price of risky asset  $i$  is<sup>22</sup>

$$\begin{aligned} P_{i,t} &= \frac{1}{1 - (\mu^e/\mu^f) \lambda_{i,1} (1 - \lambda_{i,2})} D_{i,t} \\ &+ \frac{\mu^e/\mu^f}{1 - (\mu^e/\mu^f) \lambda_{i,1} (1 - \lambda_{i,2})} \\ &\times (\lambda_{i,0} + \lambda_{i,1} \lambda_{i,2} S_{i,t-1} - \lambda_{i,1} (1 - \lambda_{i,2}) P_{i,t-1}) \\ &- \frac{(\gamma (T - t - 1) \Sigma Q + (\mu^f)^{-1} \gamma \Sigma Q)_i}{1 - (\mu^e/\mu^f) \lambda_{i,1} (1 - \lambda_{i,2})}. \end{aligned} \quad (\text{A.18})$$

Eq. (A.18) demonstrates an amplification mechanism: the good fundamental news at time  $t$ —an increase from  $D_{i,t-1}$  to  $D_{i,t}$ —pushes up the price  $P_{i,t}$  for risky asset  $i$ , causing extrapolators to become more optimistic about the asset's future return and hence increasing their share demand. This in turn pushes the price  $P_{i,t}$  further up. The amplification mechanism implies that equilibrium only exists if

$$(\mu^e/\mu^f) \lambda_{i,1} (1 - \lambda_{i,2}) < 1. \quad (\text{A.19})$$

This inequality holds under two conditions. First, there needs to be a sufficient population fraction of fundamental traders in the economy who trade against mispricing (i.e.,  $\mu^f/\mu^e$  needs to be sufficiently large). Second,  $\lambda_{i,1} (1 - \lambda_{i,2})$ —what we define as extrapolators' degree of extrapolation for stock  $i$ —needs to be sufficiently small.

To understand the asset pricing implications of the model, we run the following predictive regression of the future price change  $P_{i,t+1} - P_{i,t}$  on the current sentiment  $S_{i,t}$ :

$$P_{i,t+1} - P_{i,t} = \alpha_i + b_i \cdot S_{i,t} + \xi_{i,t+1}. \quad (\text{A.20})$$

The price equation in Eq. (A.18) implies  $\alpha_i = (1 - (\mu^e/\mu^f) \lambda_{i,1} (1 - \lambda_{i,2}))^{-1} \gamma (\Sigma Q)_i$  and  $\xi_{i,t+1} = (1 - (\mu^e/\mu^f) \lambda_{i,1} (1 - \lambda_{i,2}))^{-1} \varepsilon_{i,t+1}$ . More importantly, the slope coefficient in Eq. (A.20) equals

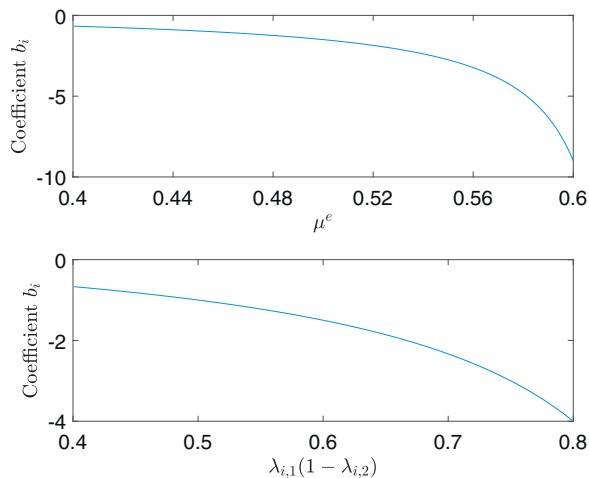
$$b_i = -\frac{(\mu^e/\mu^f) \lambda_{i,1} (1 - \lambda_{i,2})}{1 - (\mu^e/\mu^f) \lambda_{i,1} (1 - \lambda_{i,2})}, \quad (\text{A.21})$$

and  $b_i < 0$  if  $0 < (\mu^e/\mu^f) \lambda_{i,1} (1 - \lambda_{i,2}) < 1$ .

Note that extrapolators' time- $t$  sentiment  $S_{i,t}$  for risky asset  $i$  is an expectation measure. Up to an affine transformation, the empirical analog of  $S_{i,t}$  is the predicted score of stock  $i$ : it is the fitted value from the nonlinear regression in Eq. (2) of the main text using stock  $i$ 's consensus Forcerank score as the dependent variable and using the stock's contest-adjusted past returns as the explanatory variables. The model predicts that the coefficient from regressing the future stock return on the

<sup>21</sup> Early work of Barberis and Shleifer (2003) and Barberis et al. (2018) has also made the simplifying assumptions of Eqs. (A.11) and (A.16).

<sup>22</sup> Cross-asset extrapolation does not arise in our model. There are two reasons for this. First, we assume CARA preferences for fundamental traders and extrapolators; these preferences eliminate any wealth effect and rebalancing motives. Second, we assume bounded rationality on the part of fundamental traders—these traders always expect mispricing for all risky assets to be corrected over the next period—and therefore further eliminate any hedging motives. Given these two assumptions, our cross-sectional model of return extrapolation reduces to a model of return extrapolation on individual stocks: the price of stock  $i$  in Eq. (A.18) only depends on its own past prices but not on the past prices of other stocks.



**Fig. A1.** Relation between return predictability and parameters  $\mu^e$ ,  $\lambda_{i,1}$ , and  $\lambda_{i,2}$ . The figures analyze the slope coefficient  $b_i$  in the predictive regression  $P_{i,t+1} - P_{i,t} = \alpha_i + b_i \cdot S_{i,t} + \xi_{i,t+1}$  specified in Eq. (A.20) of the Appendix, where  $S_{i,t}$ , as defined in Eq. (A.4), is extrapolators' time- $t$  "sentiment" on stock  $i$ . Panel A plots  $b_i$  as a function of  $\mu^e$ , the population fraction of extrapolators in the economy. Panel B plots  $b_i$  as a function of  $\lambda_{i,1}(1 - \lambda_{i,2})$ , our model-implied measure of extrapolators' degree of extrapolation for stock  $i$ . The default parameter values are  $\lambda_{i,1} = 1$ ,  $\lambda_{i,2} = 0.4$ , and  $\mu^e = 0.5$ .

stock's current predicted score—or on the stock's current Forcerank score that contains the predicted score as a component—should be negative. Moreover, the expression of  $b_i$  in Eq. (A.21) connects stock  $i$ 's return predictability with the belief parameters  $\lambda_{i,1}$  and  $\lambda_{i,2}$  from extrapolators and the population fraction of extrapolators  $\mu^e$ .

Fig. A1 shows that, for a higher  $\mu^e$  or a higher  $\lambda_{i,1}(1 - \lambda_{i,2})$ , the magnitude of the regression coefficient  $b_i$  in Eq. (A.21) is larger. In other words, the model generates two predictions regarding the heterogeneity of return predictability. First, return predictability should be stronger among stocks whose clienteles are dominated by behavioral extrapolators—this is when  $\mu^e$  is higher. Second, return predictability should also be stronger among stocks traded by extrapolators whose degree of extrapolation is higher—this is when  $\lambda_{i,1}(1 - \lambda_{i,2})$  is higher. In Section 4 of the main text, we use our cross-sectional expectation data to test and confirm these two model predictions.

We complete the discussion of the model by making a remark on the model's ability to generate momentum. Some extrapolation models—e.g., Barberis and Shleifer (2003) and Barberis et al. (2018)—give rise to both momentum and longer-term return reversals. Some other extrapolation models—e.g., Barberis et al. (2015) and Jin and Sui (2019)—however, only generate return reversals. The key difference lies in the models' assumption on the relation between extrapolators' current return expectation and past returns. If this relation is assumed to be hump-shaped—i.e., if we regress the current model-implied return expectation on all past returns, the coefficients, when plotted against the passage time between the current time and the time when the past return took place, display a hump shape—then the model generates both momentum and reversals. If, on the

other hand, this relation is assumed to be monotonically decreasing, and furthermore, if there is no gap in time between belief formation and trading, then the model only generates reversals. In the end, the relation between the current return expectation and past returns needs to be measured empirically. As we show in Section 3 of the main text, our weekly expectations data indeed find this relation to be monotonically decreasing. This in turn justifies our key belief assumption in Eqs. (A.3) and (A.4). Moreover, consistent with the asset pricing implication of this documented monotonic relation, we only observe return reversals, instead of both momentum and reversals, at the weekly horizon.

## Appendix B. Additional discussion about time-scale invariance

Time-scale invariance refers to the hypothesis that memory retrieval is invariant across different time scales. The psychology literature has documented experimental evidence for this hypothesis (see, e.g., Maylor et al., 2001 and Moreton and Ward, 2010). Survey data about investor expectations provide further empirical support for time-scale invariance. From Column (1) of Table 4, an estimate of 0.59 for  $\lambda_2$  suggests that, when forming expectations about the next week's return, investors put 12% weight on returns four weeks earlier relative to returns in the most recent week; the ratio of the forecasting horizon (one week) to the backward-looking time window of expectation formation (four weeks) is about one over four. In comparison, Barberis et al. (2015) report a similar estimate of 0.61 for  $\lambda_2$  using Gallup data in which investors make longer-term forecasts, suggesting that when forming expectations about the next year's return, investors put 14% weight on returns four years earlier relative to returns in the most recent year; while the forecasting horizon is now one year rather than one week, the ratio of the forecasting horizon to the backward-looking time window remains to be one over four.<sup>23</sup> This numerical comparison shows stability in the estimation of belief parameters after adjusting for time horizons. It suggests that our findings about expectation formation are not restricted to a short forecasting horizon; it also has direct implications for the formation of expectations over longer forecasting horizons.

## References

- Amromin, G., Sharpe, S.A., 2013. Expectation of risk and return among household investors: are their Sharpe ratios countercyclical? Unpublished working paper. Federal Reserve Board.
- Atkins, A.B., Dyl, E.A., 1990. Price reversals, bid-ask spreads, and market efficiency. *J. Financ. Quant. Anal.* 25, 535–547.
- Bacchetta, P., Mertens, E., van Wincoop, E., 2009. Predictability in financial markets: what do survey expectations tell us? *J. Int. Money Finance* 28, 406–426.
- Barber, B.M., Odean, T., 2000. Trading is hazardous to your wealth: the common stock investment performance of individual investors. *J. Finance* 55, 773–806.

<sup>23</sup> Specifically, the estimate of  $\lambda_2$  is based on  $\lambda_2 = \exp(-0.49 \times 1) \approx 0.61$ , where 0.49 is the estimate of the extrapolation parameter, a coefficient in the exponent, from Barberis et al. (2015), and 1 corresponds to the one-year time interval between consecutive past annual returns.

- Barber, B.M., Odean, T., 2008. All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. *Rev. Financ. Stud.* 21, 785–818.
- Barber, B.M., Odean, T., Zhu, N., 2009. Systematic noise. *J. Financ. Mark.* 12, 547–569.
- Barberis, N., 2018. Psychology-based models of asset prices and trading volume. In: Bernheim, D., DellaVigna, S., Laibson, D. (Eds.), *Handbook of Behavioral Economics*. North Holland, Amsterdam, pp. 79–175.
- Barberis, N., Greenwood, R., Jin, L., Shleifer, A., 2015. X-CAPM: an extrapolative capital asset pricing model. *J. Financ. Econ.* 115, 1–24.
- Barberis, N., Greenwood, R., Jin, L., Shleifer, A., 2018. Extrapolation and bubbles. *J. Financ. Econ.* 129, 203–227.
- Barberis, N., Shleifer, A., 2003. Style investing. *J. Financ. Econ.* 68, 161–199.
- Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. *J. Financ. Econ.* 49, 307–343.
- Barberis, N., Xiong, W., 2012. Realization utility. *J. Financ. Econ.* 104, 251–271.
- Begenau, J., Farboodi, M., Veldkamp, L., 2018. Big data in finance and the growth of large firms. *J. Monetary Econ.* 97, 71–87.
- Bordalo, P., Gennaioli, N., La Porta, R., Shleifer, A., 2019. Diagnostic expectations and stock returns. *J. Finance* 74, 2839–2874.
- Bordalo, P., Gennaioli, N., Shleifer, A., 2018. Diagnostic expectations and credit cycles. *J. Finance* 73, 199–227.
- Brav, A., Lehavy, R., 2003. An empirical analysis of analysts' target prices: short-term informativeness and long-term dynamics. *J. Finance* 58, 1933–1967.
- Campbell, J.Y., Grossman, S.J., Wang, J., 1993. Trading volume and serial correlation in stock returns. *Q. J. Econ.* 108, 905–939.
- Cassella, S., Gulen, H., 2018. Extrapolation bias and the predictability of stock returns by price-scaled variables. *Rev. Financ. Stud.* 31, 4345–4397.
- Cassella, S., Gulen, H., 2019. Belief-based equity market sentiment. Unpublished working paper. *Tilburg University and Purdue University*.
- Da, Z., Liu, Q., Schaumburg, E., 2013. A closer look at the short-term return reversal. *Manag. Sci.* 60, 658–674.
- Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under- and overreactions. *J. Finance* 53, 1839–1885.
- Fama, E., 1965. The behavior of stock market prices. *J. Bus.* 38, 34–105.
- Garcia, D., 2018. The kinks of financial journalism. Unpublished working paper. *University of Colorado Boulder*.
- Gennaioli, N., Shleifer, A., 2018. *A Crisis of Beliefs: Investor Psychology and Financial Fragility*. Princeton, Princeton University Press.
- Giglio, S., Maggiori, M., Stroebel, J., Utkus, S., 2020. Five facts about beliefs and portfolios. Unpublished working paper. *Yale University, Stanford University, and New York University*.
- Greenwood, R., Hanson, S., Jin, L., 2019. Reflexivity in credit markets. Unpublished working paper. *Harvard University and California Institute of Technology*.
- Greenwood, R., Shleifer, A., 2014. Expectations of returns and expected returns. *Rev. Financ. Stud.* 27, 714–746.
- Hirshleifer, D., Li, J., Yu, J., 2015. Asset pricing in production economies with extrapolative expectations. *J. Monet. Econ.* 76, 87–106.
- Ingersoll, J.E., Jin, L.J., 2013. Realization utility with reference-dependent preferences. *Rev. Financ. Stud.* 26, 723–767.
- Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. *J. Finance* 45, 881–898.
- Jegadeesh, N., Titman, S., 1995. Overreaction, delayed reaction, and contrarian profits. *Rev. Financ. Stud.* 8, 973–993.
- Jin, L.J., Sui, P., 2019. Asset pricing with return extrapolation. Unpublished working paper. *California Institute of Technology*.
- Kojien, R.S.J., Schmeling, M., Vrugt, E.B., 2015. Survey expectations of returns and asset pricing puzzles. Unpublished working paper. *London Business School*.
- Kuchler, T., Zafar, B., 2019. Personal experiences and expectations about aggregate outcomes. *J. Finance* 74, 2491–2542.
- Kuhnen, C.M., 2015. Asymmetric learning from financial information. *J. Finance* 70, 2029–2062.
- Kuhnen, C.M., Knutson, B., 2005. The neural basis of financial risk taking. *Neuron* 47, 763–770.
- Kumar, A., Ruenzi, S., Ungeheuer, M., 2019. Daily winners and losers. Unpublished working paper. *University of Miami*.
- Lehmann, B., 1990. Fads, martingales, and market efficiency. *Q. J. Econ.* 105, 1–28.
- Liu, H., Peng, C., Xiong, W.A., Xiong, W., 2020. Resolving the excessive trading puzzle: an integrated approach based on surveys and transactions. Unpublished working paper. *London School of Economics and Princeton University*.
- Maylor, E.A., Chater, N., Brown, G.D.A., 2001. Scale invariance in the retrieval of retrospective and prospective memories. *Psychonomic Bull. Rev.* 8, 162–167.
- Moreton, B.J., Ward, G., 2010. Time scale similarity and long-term memory for autobiographical events. *Psychonom. Bull. Rev.* 17, 510–515.
- Nagel, S., Xu, Z., 2019. Asset pricing with fading memory. Unpublished working paper. *University of Chicago and University of Michigan*.
- Niessner, M., So, E.C., 2018. Bad news bearers: the negative tilt of the financial press. Unpublished working paper. *Massachusetts Institute of Technology*.
- Odean, T., 1998. Are investors reluctant to realized their losses? *J. Finance* 53, 1775–1798.
- Piazzesi, M., Schneider, M., 2009. Momentum traders in the housing market: survey evidence and a search model. *Am. Econ. Rev.* 99, 406–411.
- Reyes, T., 2019. Negativity bias in attention allocation: retail investors' reaction to stock returns. *Int. Rev. Finance* 19, 155–189.
- Subrahmanyam, A., 2005. Distinguishing between rationales for short-horizon predictability of stock returns. *Financ. Rev.* 40, 11–35.
- Vissing-Jørgensen, A., 2004. Perspectives on behavioral finance: does "irrationality" disappear with wealth? Evidence from expectations and actions. *NBER Macroecon. Annual* 18, 139–194.