Extrapolative Beliefs in the Cross-Section: What Can We Learn from the Crowds?*

Zhi Da†, Xing Huang‡, Lawrence Jin§

April 9, 2019

ABSTRACT

Using novel data from a crowdsourcing platform for ranking stocks, we investigate how individuals form expectations about future stock returns in the cross-section. We find that investors extrapolate from past returns, with more weight on more recent returns, especially when recent returns are negative or salient. Such an extrapolative bias is stronger among non-professionals. Moreover, consensus rankings negatively predict future stock returns in the cross-section, more so among stocks with low institutional ownership and a high degree of extrapolative bias, consistent with the asset pricing implications of extrapolative beliefs.

JEL classification: G4, G12
Keywords: Return Extrapolation, Beliefs in the Cross-Section, Expectation Formation

*We thank Nicholas Barberis, Julien Cujean, Peter Cziraki, Michael Ewens, Samuel Hartzmark, Burton Hollifield, Lisa Kramer, Juhani Linnainmaa, Yueran Ma, Abhiroop Mukherjee, Marina Niessner, Terrance Odean, Cameron Peng, Jesse Shapiro, David Solomon, and seminar and conference participants at Baylor University, Boston College, Caltech, CKGSB, Florida International University, SWUFE, UC Berkeley, Washington University's Olin Business School, the 2019 American Finance Association Annual Meeting, the 15th Annual Conference in Financial Economics Research (IDC), CICF 2018, EFA 2018, FIRS 2018, the LA Finance Day Conference, MFA 2019, 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 Dorgen, Josh Dulberger, and Aram Balian from Forcerank.com for their generous support.

†Mendoza College of Business at University of Notre Dame. zda@nd.edu, +1 (574) 631-0354.
‡Olin Business School at Washington University in St. Louis. xing.huang@wustl.edu, +1 (314) 935-7848.
§Division of the Humanities and Social Sciences at Caltech. lawrence.jin@caltech.edu, +1 (626) 395-4558.

I. Introduction

A central question in finance is how investors form expectations about future asset returns. Recent works by Vissing-Jorgensen (2004), Amromin and Sharpe (2013), Greenwood and Shleifer (2014), and Kuchler and Zafar (2016) provide convincing evidence of return extrapolation, or 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, Greenwood, Jin, and Shleifer (2015) and Jin and Sui (2018) 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 been tested primarily with market-level data so far. For example, by using survey expectations of investors about future stock market returns, Cassella and Gulen (2018a) estimate the degree of extrapolation bias (DOX) and find that market return predictability is significant only during high-DOX periods. In the cross-section, however, analyzing extrapolative beliefs and testing the asset pricing implications of these beliefs are challenging; this is in part due to the lack of data that directly measure investors' expectations on future returns of individual stocks.[1]

The emergence of financial technology (FinTech) makes it easy to survey a large number of individual investors on their beliefs about a cross-section of stocks. In this paper, we analyze a novel dataset from a crowdsourcing platform (called Forcerank) for ranking stocks. In each contest on this platform, participants rank 10 stocks based on their perceived future performance of these stocks over the course of the contest (usually one week). Compared to alternative data sources, Forcerank data have a number of unique advantages. It contains precise ranking information among a set of pre-determined stocks. Moreover, the rankings correspond to a clearly-specified horizon and are solicited from highly diverse and geographically distributed individuals in a blind setting that rules out herding or cross-learning.[2]

[1]Cassella and Gulen (2018b) analyze the relation between investor expectations about aggregate stock market returns and the relative pricing of stocks in the cross-section. Bordalo, Gennaioli, La Porta, and Shleifer (2017) examine analyst expectations about earnings growth in the cross-section. However, these two studies do not directly analyze cross-sectional data of *return* expectations.

[2]Some social media platforms (e.g., StockTwits; Seeking Alpha) collect user opinions in textual form. These opinions can be viewed as a beliefs measure. However, textual information may not 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 biases, we find suggestive evidence for extrapolative beliefs in the Appendix even among equity analysts, after removing the very illiquid penny

Taking advantage of the Forcerank data, we investigate how individuals form their expectations about future returns on individual stocks and how these return expectations affect asset prices. We first estimate, across stocks, a linear regression of investor expectations on past stock returns; here we use rankings data from Forcerank as a proxy for investor expectations. 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, for returns from recent to more distant past, the coefficients decline in general: quantitatively, returns four weeks earlier are only about 9% as important as returns in the most recent week. Moreover, this extrapolative pattern remains almost identical after controlling for past fundamental and risk measures and is very robust to alternative regression model specifications. As an external validation, we find very similar extrapolative pattern when we examine the initial buys of a large group of short-term retail traders in the brokerage account data used in Odean (1998). In other words, our findings do not seem to be driven by the unique game setting on Forcerank.

To parsimoniously capture the extrapolative pattern, we further apply an exponential decay function as the weighting scheme on past returns.[3] In doing so, we summarize the *degree of extrapolation* from investor expectations into two parameters. The first parameter λ_1—it is a scaling factor multiplied to all past returns—captures a "level" effect—that is, the overall extent to which individuals respond to past returns. The second parameter λ_2—this is 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 λ_1 and λ_2: when λ_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 λ_1 is high. We find that λ_1 estimated from the Forcerank expectations data is significantly positive and λ_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 allows us to study the determinants of investors' extrapolative expec-

stocks.) Finally, individual investors' trading decisions are sometimes used as a measure for investor beliefs (Barber, Odean, and Zhu (2009)), although they may also be driven by factors other than beliefs such as liquidity shocks and preferences (Odean (1998); Barberis and Xiong (2012); Ingersoll and Jin (2013)). Furthermore, short-sale constraints may prevent an investor from expressing negative return expectations through trading, and the investor' choice set may be limited to stocks that recently caught her attention (Barber and Odean (2008)).

[3]Early works of Greenwood and Shleifer (2014), Barberis et al. (2015), and Cassella and Gulen (2018a) have used this specification to study investors' return expectations about the aggregate stock market.

tations as captured by λ_1 and λ_2. We find that extrapolation is *asymmetric* between positive and negative past returns: investors put more weights (a larger λ_1) on negative past returns, and these weights decay more slowly into the past (a higher λ_2) for negative returns. Similarly, we find that investor expectations respond more strongly to salient past returns (a larger λ_1), and salient returns from both the recent past and the distant past affect investor expectations (a higher λ_2).

We further examine how user and firm characteristics affect expectation formation. We show that, financial professionals, compared to non-professionals, display a lower degree of extrapolation. Specifically, the λ_1 estimate for professionals is significantly lower than that for the non-professionals, suggesting that overall professionals rely less on past stock returns when forming expectations about the next-week return. Moreover, the λ_2 estimate for professionals is significantly higher than that for the non-professionals, suggesting that professionals' expectations rely on past returns over a longer history. At the firm level, we find that λ_1 is positively related to firm size, but negatively related to the firm's average volatility of weekly returns. We also find that λ_2 is positively related to firm size and turnover, but 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 findings provide empirical regularities for future theoretical works 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 not. We use Forerank scores—the average ranking across all participants who make a forecast of an individual stock—as a proxy for extrapolative expectations. We find that a higher Forcerank score significantly predicts a *lower* next-week return in Fama-MacBeth cross-sectional regressions. We then decompose the Forcerank score into a predicted score (as a weighted average of the stock's past twelve week returns where the weights are calibrated to the beliefs of Forcerank participant) and a residual. We find significant negative return predictability on both components. These return predictability results survive the control of returns over the past one-week, one-month, and one-quarter, so we are not simply rediscovering the well-documented short-term return reversals. Altogether, our results suggest that the beliefs of Forcerank participants are systematically biased.

To clarify, we do not claim that Forcerank users alone move market 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 their impact on asset prices, in the Appendix, we present a cross-sectional model of return extrapolation. The model consists of two types of agents, extrapolators and fundamental traders. Consistent with the beliefs of Forcerank users, extrapolators form expectations about future returns of individual stocks by extrapolating 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 returns as we documented.

Importantly, the stylised model 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 traded by extrapolators whose degree of extrapolative bias—measured by $\lambda_1(1 - \lambda_2)$ in the model—is higher. 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 7 basis points per day (or, equivalently, almost 18 percent per year) in our sample, after controlling for the Fama-French five-factors and a short-term reversal factor. As an external validation of our findings, we extend our analysis to the sample of stocks that are *not* covered by the Forcerank platform. To do this, we compute predicted Forcerank scores for non-Forcerank stocks as a weighted average of their past twelve-week returns where the weights are calibrated to the Forcerank data. We find the predicted score to negatively forecast the next-week return 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 the predicted score 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); Koijen, Schmeling, and Vrugt (2015); Kuchler and Zafar (2016)). 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, Li, and Yu (2015); Barberis et al. (2015); Barberis, Greenwood, Jin, and Shleifer (2018); Cassella and Gulen (2018a); Jin and Sui (2018); Bordalo, Gennaioli, and Shleifer (2018); Greenwood, Hanson, and Jin (2018); Nagel and Xu (2018); Gennaioli and Shleifer (2018); Bordalo et al. (2017); Cassella and Gulen (2018b)). Finally, our paper contributes to the voluminous literature on short-term reversal starting from Fama (1965), Jegadeesh (1990), and Lehmann (1990). The fact that we document significant return predictability on the largest and most liquid stocks suggest that extrapolative belief (in addition to liquidity shock) can also be an important contributor to the short-term return reversal phenomenon.

In what follows, we first describe in details the crowdsourcing platform and our sample in Section II. Section III presents the empirical results that analyze investors' expectation formation. Section IV shows the predictive power of Forcerank scores for future stock returns as well as the results from trading strategies. Section V concludes. The Appendix contains a stylized asset pricing model of extrapolative belief and additional results using equity analysts' price targets.

II. Data and Summary Statistics

In this section, we provide description for the data from Forcerank.com. Forcerank is a crowdsourcing platform that organizes weekly competitions in which participants enter thematic games, and in each game, rank a list of 10 stocks according to their *perceived* expected performance (% gain) of these stocks over the course of the game (usually one week).

There are two main types of games. Most games are comprised within an industry group. For example, in one game, contestants may be asked to rank 10 stocks from the same E-commerce industry based on their expectations of the stocks' next-week returns. Occasionally, the industry group is further partitioned by the market capitalizations of the stocks. For example, one game may contain only large stocks from the Biotech industry. The other type of games is based on special themes, such as most heavily shorted stocks, or ETFs. We focus on individual firms in our study and therefore exclude games involving ETFs. Table 1 lists the types of games in our final sample, which covers a period from February 2016 to December 2017.

[Place Table 1 about here]

In addition, the same game is repeatedly conducted every week on the platform, resulting in multiple weekly contests for the same game. The goal of the participants is to correctly match up the rankings with the actual performances at the end of the contest period. Figure 2 illustrates an example of one such contest.

[Place Figure 2 about here]

Upon completion of each game, Forcerank assigns points to participants based on the accuracy of their own rankings exclusively, rather than benchmarking to the performance of other participants. Points do not result in monetary compensation due to the legal risks involved.[4] Instead, Forcerank maintains weekly leader boards where participants are ranked based on their cumulative points in the past 13 weeks. Ranking participants based on cumulative points helps alleviate strategic "anti-herding" behavior that can arise from tournament with a convex payoff structure. In addition, herding or anti-herding is difficult on the platform for two other reasons. First, during each game, 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, among users who choose to participate on Forcerank, they are likely to truthfully reveal their relative 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 from individual investors. Over time, Forcerank expands its game coverage to also include industries such as fast food, investment banking, airlines, and semiconductors. The only non-industry game we study involves heavily shorted stocks (78 weekly contests that span a period from March 2016 to December 2017).

[4] Monetary compensation could turn the Forcerank game into illegal security-based swap in the eye of SEC (see http://dodd-frank.com/sec-says-mobile-phone-game-is-an-illegal-security-based-swap). The lack of monetary compensation may explain the slow growth in user participation and Forcerank's decision to temporarily shut down the platform since April 2018, in order to focus on developing another crowdsourcing platform called Estimize.

Our final sample contains 293 unique stock tickers. It contains 12,798 contributions submitted by 1,045 distinct users. A breakdown of stocks and users can be found in Table 2.

[Place Table 2 about here]

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, our sample stocks are less likely to be subject to the short-term return reversals induced by liquidity shocks. Our sample also gears towards growth stocks. The average stock has a book-to-market ratio B/M of 0.37 (the median is 0.26). The average B/M quintile rank is 2.20.

The contest participation of users 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.

We observe self-reported user professional background among a fraction of users who registered before March 2017. Specifically, among 606 users who registered before March 2017, 244 of them chose to report their professional backgrounds. Panel B of Table 2 breaks down these 244 users. Among them, 72 are financial professionals. We conjecture that the extrapolation bias is less pronounced among financial professionals. In our empirical analysis, we confirm this conjecture.

III. Expectation Formation

In this section, we study the formation of investor expectations using the Forcerank data. We first look at the average beliefs of Forcerank users, and relate these beliefs to past variables such as past stock returns. We then examine how user and firm characteristics affect expectation formation.

To start, we analyze how past stock returns affect Forcerank users' average expectations on future stock returns. For each week t, individuals are asked to submit rankings of 10 stocks according to their *perceived* expected performance of these stocks over week $t+1$. For each stock in each contest, we measure the investor expectation by the consensus Forcerank score averaged across all individuals. For each individual, their highest ranked stock receives a score of 10; and the second highest ranked stock receives a score of 9. Similarly, the lowest ranked stock receives a score of 1; and the second lowest ranked stock receives a score of 2.

III.1. Linear model

We start with a simple linear model with the consensus rank score as the dependent variable and past stock returns as the independent variables:

$$\text{Forcerank}_{i,t} = \gamma_0 + \sum\nolimits_{s=0}^{n} \beta_s \cdot R_{i,t-s} + \varepsilon_{i,t}, \tag{1}$$

where $\text{Forcerank}_{i,t}$ is the week-t consensus (average) rank based on investors' expectations about the performance of stock i over week $t+1$; $R_{i,t-s}$ represents the lagged return (or the contest-adjusted return we define below) of stock i over week $t-s$, and $s = 0$ to 11 weeks.

[Place Table 3 about here]

The results are reported in Table 3. Column (1) uses the raw level of past returns. The results show clear evidence that individuals extrapolate past returns. The coefficients on the past twelve weekly returns are all positive and mostly significant. More important, for returns from recent to more distant past, the coefficients decline in general, meaning that investors put higher weight on more recent returns.[5]

Given that individuals submit relative rankings on Forcerank, it is possible that the relative levels of returns within a contest are more relevant to form beliefs. In Column (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 contest average return. The regression results remain similar in Column (2). The coefficients on past contest-adjusted returns and the R-squared all increase, indicating a better fit of the data.

The positive relation between the current return expectation and recent past returns is robust to different estimation methods and sampling periods. For example, in Column (3), we estimate an ordered logit model that accounts for the ranking nature of our dependent variables. In Column (4), we also convert the independent variables (past weekly returns) to rankings. The return

[5] The regression evidence is supported by results from a survey of 20 Caltech undergraduate students. When asked about how they come up with their rankings of stocks, the typical response is that the ranking is based on "last week and last month's performance," "a quick look of past month returns," or "roughly on last week's ranks."

extrapolation pattern remains similar. The coefficients on the past twelve weekly returns (or return rankings) are all positive and significant and decline with the lag.

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. While we observe return extrapolative pattern in both sub-periods, the pattern seems to be stronger in the second half. A potential reason is that the increased user participation on Forcerank over time makes the consensus ranking less noisy.

Finally, Column (7) includes other potential drivers of future return expectations. These additional controls include past fundamental news as measured by earnings surprises over the past four quarters, the tone of news coverage, and the expected return according to the CAPM. These additional controls do not alter the basic return extrapolation pattern. The coefficients on past contest-adjusted returns remain positive and significant. Moreover, the decay pattern among coefficients for past returns remains strong and quantitatively similar with and without controls. The additional data requirement reduces the sample size by almost 40%. For this reason, we do not include these additional controls in most of our remaining analyses.

III.2. External Validation

An immediate concern is that our findings so far are driven by features unique to Forcerank: game specification, the interface, and the characteristics and incentives of users who self-selected to participate, etc. As an external validation, we now examine the trading behavior of a large group of retail investors. Specifically, we focus on the initial buys of these investors. Compared to other types of trades—such as additional purchase of the same stock and sales, which could be driven by non-belief factors such as preferences and 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 1991 to 1996 (as in Odean (1998)). 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. Finally, we remove individuals who have less than ten sales from the sample.

[Place Figure 3 about here]

We run a linear regression similar to (1) of initial buys on past twelve week returns. The dependent variable is an indicator variable that equals one if, during week t, there is at least one individual who purchases the stock for the first time in the sample. Figure 3 then plots the coefficient estimates on past twelve week returns using initial buys (right y-axis) against those from the Forcerank sample (left y-axis). The solid lines correspond to the coefficient estimates, and the dashed lines correspond to the 95% confidence interval. We find that the extrapolative belief patterns are very similar across these two settings, as the two sets of coefficient estimates are almost proportional to each other.

III.3. Exponential decay model

By using a simple linear specification that allows for independent weights on different past returns, we observe a clear and robust decay pattern in the relation between the investor's current return expectation and recent past returns. To capture this pattern parsimoniously, we now estimate a parametric extrapolation model which assumes an exponential decay of weights on past returns. Specifically, we examine an empirical version of the key behavioral assumption (A.3) made in our stylised model about investor beliefs:

$$\text{Forcerank}_{i,t} = \lambda_0 + \lambda_1 \cdot \sum\nolimits_{s=0}^{n} w_s R_{i,t-s} + \varepsilon_{i,t}, \qquad \text{where} \;\; w_s = \frac{\lambda_2^s}{\sum_{j=0}^{n} \lambda_2^j}. \tag{2}$$

This exponential decay specification has been previously estimated by Greenwood and Shleifer (2014), Barberis et al. (2015), and Cassella and Gulen (2018a), using aggregate stock market data. As discussed previously, it allows us to characterize the relation between the investor's current return expectation and recent past returns by two parameters. The first parameter λ_1—it is a scaling factor multiplied to all past returns of stock i—captures a "level" effect—that is, the overall extent to which individuals respond to these past returns. The second parameter λ_2—it governs how past returns are relatively weighted in forming the expectation—captures a "slope" effect: a λ_2 significantly smaller than one means that investors put much higher weights on recent past returns as opposed to distant past returns. When an investor puts more weights on all past returns of stock i and, furthermore, assigns more weight on more recent returns, her beliefs are more extrapolative. That is, a higher λ_1 and a lower λ_2 estimated using the Forcerank data jointly lead to a higher degree of extrapolation bias. Here, we estimate the two parameters by assuming them to be constant in the full sample across all stocks and individuals. In Section IV, we let them to be game specific—we estimate $\lambda_{i,1}$ and $\lambda_{i,2}$ for each game i. We then derive and test the cross-sectional relation between these two parameters with the predictability of future returns.

[Place Figure 4 and Table 4 about here]

We include various number of lags (n in equation (2)) in the estimation of λ_1 and λ_2. Figure 4 shows that both λ_1 and λ_2 become stable when we include more than twelve past weekly returns in the estimation. We therefore use $n = 11$ for the rest of our analysis.

Table 4 confirms the extrapolation bias using the nonlinear specification in (2). Specifically, Column (1) uses the raw level of past returns. Columns (2) and onwards focus on contest-adjusted returns. Column (2) reports λ_1 of 34.12 and λ_2 of 0.55. These joint estimates of λ_1 and λ_2 suggest that Forcerank participants exhibit a strong degree of extrapolation bias. To check robustness of these estimates, Column (3) uses the rankings of past returns. Columns (4) and (5) break up the results for the first-half (before March 1, 2017) and the second-half (after March 1, 2017) of our sample period. Finally, Column (6) further includes controls of past fundamentals, the tone of news coverage, and the CAPM expected return. Across all columns, we find λ_1 to be significantly positive and λ_2 to be positive and significantly smaller than one. For the rest of the paper, we focus primarily on the nonlinear specification in equation (2) when analyzing investor expectations, as it succinctly summarizes the extrapolation bias by two parameters.

One concern about our Forcerank data is that it only measures return expectations over one week, a very short horizon, and as a result, our data may not be helpful for understanding investor beliefs over longer horizons (such as six months or one year). However, one observation can be helpful for addressing this concern: investors tend to look at past few weeks' returns when forecasting the next-week return; and they tend to look at past few years' returns when forecasting the next-year return. That is, when forecasting the future return over a time length of t, investors tend to look at past returns over a time length of $N \cdot t$, where N tends to be independent of the expectation horizon t, capturing a more fundamental psychological factor such as the degree of a recency bias or the speed of memory decay. Indeed, a λ_2 of 0.55 estimated in Column (2) of Table 4 suggests that, when forming expectations about the next-week return, investors put about

9% weight to returns four weeks earlier relative to returns in the most recent week. In comparison, Barberis et al. (2015) report a similar estimation of 0.49 for λ_2 using Gallap data in which investors make longer-term forecasts, suggesting that when forming expectations about the next-year return, investors put about 6% weight to returns four years earlier relative to returns in the most recent year. This comparison indicates stability in the estimation of belief parameters after adjusting for time horizons; the psychology literature has documented some evidence for this type of "time scale invariance" (see Maylor, Chater, and Brown (2001)). As a result, our findings on expectation formation may also have direct implications for expectations over longer horizons.

III.4. Past return characteristics

To develop a deeper understanding of expectation formation, we further generalize the regression in (2) by separately estimating λ_1 and λ_2 for past returns of different characteristics. Recent empirical, experimental, and neuroscience evidence suggests that expectation formation may differ depending on whether past outcomes are positive or negative.[6] To test this hypothesis, we first separate past returns to positive returns and negative returns, and run a generalized non-linear regression of the form

$$\begin{aligned} \text{Forcerank}_{i,t} &= \lambda_0 + \lambda_{1,pos} \cdot \sum_{s=0}^{n} \mathbb{1}_{\{R_{i,t-s} \geq 0\}} \cdot w_{s,pos} R_{i,t-s} \\ &\quad + \lambda_{1,neg} \cdot \sum_{s=0}^{n} \mathbb{1}_{\{R_{i,t-s} < 0\}} \cdot w_{s,neg} R_{i,t-s} + \varepsilon_{i,t}, \end{aligned} \tag{3}$$

where $w_{s,pos} = \frac{\lambda_{2,pos}^{s}}{\sum_{j=0}^{n} \lambda_{2,pos}^{j}}$ and $w_{s,neg} = \frac{\lambda_{2,neg}^{s}}{\sum_{j=0}^{n} \lambda_{2,neg}^{j}}$.

Column (1) of Table 5 reports the empirical estimates of $\lambda_{1,pos}$, $\lambda_{2,pos}$, $\lambda_{1,neg}$, and $\lambda_{2,neg}$. The results show that return extrapolation is *asymmetric*. In particular, individuals seem to put more weights on negative past returns—$\lambda_{1,neg}$ is much larger than $\lambda_{1,pos}$—and these weights declay more slowly into the past for negative past returns—$\lambda_{2,neg}$ is much higher than $\lambda_{2,pos}$, 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

[6]For a review of this evidence, see Kuhnen (2015).

in the most recent week in determining the current expectation about future returns.

[Place Table 5 about here]

It is worth noting that Cassella and Gulen (2018a) document a closely related finding that λ_2 estimated from return expectations about the aggregate stock market is significantly higher in bear markets than in bull markets. Our result complements their finding by showing that 1) a similar asymmetric pattern in λ_2 holds in the cross-section, and 2) the asymmetry in the degree of extrapolation bias also comes from the difference between $\lambda_{1,pos}$ and $\lambda_{1,neg}$.

One potential explanation for the results in Column (1) is that negative returns are more salient than positive returns, and therefore affect expectations to a larger extent.[7] To directly test this hypothesis, we now separately estimate λ_1 and λ_2 for salient and non-salient past returns using

$$\begin{aligned} \text{Forcerank}_{i,t} \quad = \quad & \lambda_0 + \lambda_{1,sal} \cdot \sum\nolimits_{s=0}^{n} \mathbb{1}_{\{R_{i,t-s} \text{is salient}\}} \cdot w_{s,sal} R_{i,t-s} \\ & + \lambda_{1,nonsal} \cdot \sum\nolimits_{s=0}^{n} \mathbb{1}_{\{R_{i,t-s} \text{is non-salient}\}} \cdot w_{s,nonsal} R_{i,t-s} + \varepsilon_{i,t}, \end{aligned} \tag{4}$$

where $w_{s,sal} = \frac{\lambda_{2,sal}^s}{\sum_{j=0}^n \lambda_{2,sal}^j}$ and $w_{s,nonsal} = \frac{\lambda_{2,nonsal}^s}{\sum_{j=0}^n \lambda_{2,nonsal}^j}$.

To measure the level of salience associated with the return in a week, we count the number of news articles on that firm in that week. The news coverage data is obtained from Ravenpack. To ensure extreme returns are not driving variation in λ_1 and λ_2 mechanically, we orthogonalize news coverage against absolute returns. Specifically, for each week in the sample period, we run a cross-sectional regression of the total number of news coverage on a stock on its absolute return in the same week. We then define a stock return as salient (non-salient) if the residual from the cross-sectional regression is above (below) median.

Column (2) of Table 5 reports the empirical estimates of $\lambda_{1,sal}$, $\lambda_{2,sal}$, $\lambda_{1,nonsal}$, and $\lambda_{2,nonsal}$. Our data confirm the hypothesis that more salient returns are more important in shaping investor expectations. Indeed, we find that investor expectations respond more strongly to salient returns—$\lambda_{1,sal}$ is much larger than $\lambda_{1,nonsal}$—and even salient returns from the distant past remain important for current expectations—$\lambda_{2,sal}$ is significantly higher than $\lambda_{2,nonsal}$.

[7] Consistent with this explanation, Brown, Neath, and Chater (2007) provide a review of experimental evidence showing past experienced outcomes that are more distinctive from their nearby psychological neighbors—outcomes experienced around the same time—are more likely to get retrieved from memory.

III.5. User and stock characteristics

So far we have been estimating λ_1 and λ_2 for different past return characteristics. At the same time, our cross-sectional setting also allows us to link return expectations to different user characteristics. Panel A of Table 6 examines the extrapolation bias separately for professional and non-professional users.

[Place Table 6 about here]

Interestingly, between professional and non-professional users, the extrapolation parameters are quite different. Focusing on the results with contest-adjusted returns, professionals have a λ_1 of 26.35, which is lower than that of the non-professionals (33.77), suggesting that they rely less on past stock returns when forming expectations about the next-week return. Moreover, professionals have a λ_2 of 0.773, which is higher than that of the non-professionals (0.552). The result suggests that non-professionals display stronger extrapolation bias as they overweight recent returns more strongly. The weight that non-professionals put on returns decays by about 90% one month into the past, while the weight applied by professionals takes more than two months to decay by 90%. The longer decay pattern for professionals is also suggested by the data on equity analyst target prices.

We then examine how extrapolation parameters vary across different stocks. First, we estimate 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 larger firms, as well as their past returns, are more visible or salient to investors. As a result, these returns have a stronger impact on investor expectations. Another related explanation is that data from larger firms are more accessible to investors (Begenau, Farboodi, and Veldkamp (2017)). As a result, availability heuristic implies that information about these firms is more important for investors when they form expectations. In addition to the effect of firm size on $\lambda_{i,1}$ and $\lambda_{i,2}$, a firm's turnover—a measure that is positively related to the firm's size—also positively affects $\lambda_{i,2}$. Moreover, a firm's return volatility averaged across all weeks in our sample period is negatively related to $\lambda_{i,1}$: when past returns of a firm are more volatile, it becomes more difficult for investors to identify a price trend, therefore reducing their degree of extrapolation bias. Finally, $\lambda_{i,2}$ is higher for value stocks relative to growth stocks.

To conclude this section, we find strong evidence using Forcerank data that individuals extrapolate on past returns when forming expectations about future individual stock returns. Such extrapolative beliefs are stronger when recent returns are negative or salient, and among non-professionals and stocks of certain characteristics. Evidence from retail investors' initial buys provides important external validation to our findings. A natural follow-up question to examine is whether these expectations from Forcerank users are accurate or not. We address this question in the next section.

IV. Return Predictability

In this section, we examine cross-sectional return predictability associated with investor expectations expressed in their Forcerank scores. First, we examine the return predictability of the original Forcerank score, its predicted component explained by past returns, and the residual component that is orthogonal to past returns. Second, we repeat the analysis in subsamples and discuss how the result fits into a stylized asset pricing model with extrapolative belief. Finally, we evaluate the economic magnitude associated with the return predictability using trading strategies, both in-sample among Forcerank stocks and out-of-sample among all stocks over a longer period.

IV.1. Cross-sectional regressions

We first examine return predictability using Fama-MacBeth regressions where the dependent variable is the individual stock's daily return in week $t + 1$. The regression results are reported in Table 7.

[Place Table 7 about here]

As Column (1) shows, Forcerank scores negatively and significantly predict stock returns over the next week. To isolate the sentiment component of the Forcerank score which is a function of past returns, we consider a predicted Forcerank score. The predicted score is computed as the fitted value from the nonlinear regression in Table 4 (Column (2)). In other words, it is a weighted average of past twelve week 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 negatively and significantly predicts stock returns over the next week. The coefficient on the predicted score is even slightly greater in (absolute) magnitude to that on the raw Forcerank score in Column (1). It is interesting that, although past returns together explain only around 6% of the variation in the Forcerank score, they contribute significantly to Forcerank's return predictive power.

Of course, a large literature on short-term return reversal has already shown that the past return itself negatively predicts the future return and such a reversal maybe driven by liquidity shocks unrelated to extrapolation bias (see, Jegadeesh and Titman (1995); Campbell, Grossman, and Wang (1993), among others). In addition, return reversal tends to be stronger among similar stocks in the same industry (see, for example, Da, Liu, and Schaumburg (2013)). Since contests in our sample mostly include similar stocks in the same industry, a natural question is whether the predictive power associated with the Forcerank score simply reflects liquidity-shock-induced return reversal. A prior, we do not expect liquidity shocks to affect our sample stocks since they tend to be large stocks as evident in Table 2.

To address this question more directly, we examine short-term return reversal explicitly in the regressions. For each stock, we assign a quintile score based on either its contest-adjusted past one-week return (Ret(t)), or contest-adjusted past one-month return (Ret($t-3,t$)), or contest-adjusted past one-quarter return (Ret($t-11,t$)). Columns (4) and (6) show that neither the past one-week return nor the past one-quarter return has significant predictive power on the future one-week return, even after contest adjustment. Column (5) shows that the past one-month return has significant predictive power on the future one-week return, after contest adjustment. Overall, the evidence suggests that only a weak standard short-term return reversal is present in our sample.

More importantly, Columns (7) and (8) show that Forcerank score and predicted score both drive out past return measures when they are included in the same regression. Recall that the predicted score is simply a weighted average of past twelve weekly returns. The fact that the weighted average return, calibrated to extrapolative beliefs, drives out both the recent one-week return and the recent one-month return (an equal-weighted average) supports the predictions of the extrapolation model in the Appendix.

We also examine the return predictive power of the residual score, which by construction is orthogonal to past returns. Interestingly, Columns (3) and (9) show that the residual score also negatively and significantly predicts stock returns over the next week, with or without controlling for past returns. The finding suggests that the return predictive power of Forcerank score is not completely driven by its association with past returns. In other words, the Forcerank score may reveal additional investor "sentiment." We leave it for future studies.

In sum, Table 7 shows that the Forcerank score, its component related to past returns, and the residual component all negatively and significantly predict stock returns over the next week. In other words, Forcerank expectations are amazingly terrible!

Given our sample only includes about 1,000 Forcerank users, we do not claim that these users move market prices nor they represent all the market participants. Instead, we interpret our evidence as suggesting that Forcerank scores represent the thinking process of behavioral investors in the market. To understand the return predictability result better in a theoretical framework, in the Appendix, we sketch a simple cross-sectional model of return extrapolation where behavioral extrapolators and fundamental traders interact and jointly determine the asset prices. In the model, risk-adverse fundamental traders cannot completely undo the extrapolative bias, and as a result, extrapolative belief negatively predicts future return as we documented.

Importantly, the stylised model makes predictions on how extrapolators' beliefs affect return predictability in the cross section. For example, return predictability should stronger among stocks whose clienteles are dominated by behavioral extrapolators, and among stocks that experience stronger extrapolative bias as measured by $\lambda_1(1-\lambda_2)$. We now turn to these two testable predictions in the next subsection.

IV.2. Sub-sample analyses

The theoretical model in the Appendix predicts a clientele effect. Specifically, the return predictability should be stronger among stocks where extrapolators account for a bigger fraction of investor population (a bigger μ^E). To the extent that institutional investors are more likely to be rational investors, we use the level of institutional ownership as a proxy for $1 - \mu^E$.

[Place Table 8 about here]

Panel A of Table 8 runs the Fama-MacBeth return predictive regressions (as in Table 7) separately for stocks with below-median institutional ownership (more extrapolators) and for stocks with above-median institutional ownership (less extrapolators). The results confirm that the return predictability of the Forcerank score and the predicted score are only present when more extrapolators trade on the stocks.

The theoretical model in the Appendix also predicts stronger return predictability among stocks displaying higher degree of extrapolative bias. In the model, the extrapolative bias is directly measured as $\lambda_1(1-\lambda_2)$. For each stock, we estimate λ_1 and λ_2 from the in equation (2) and compute its extrapolative bias. Panel B of Table 8 then runs the Fama-MacBeth return predictive regressions separately for stocks with below-median extrapolative bias and for stocks with above-median extrapolative bias. The results confirm that the Forcerank score and the predicted score both have much stronger return predictive power among stocks with larger extrapolative bias.

IV.3. Trading strategies

To evaluate the economic magnitude associated with return predictability, we form trading strategies. At the beginning of each week, we sort the stocks on different variables into five quintiles in each contest. The portfolio is rebalanced every week. Stocks whose prices are below five dollars at the beginning of each week are removed. The results are shown in Panel A of Table 9.

[Place Table 9 about here]

Row (1) sorts stocks on the consensus Forcerank scores. It shows that Forcerank scores predict future stock returns: there is a monotonic negative relationship between Forcerank scores and stock returns over week $t+1$. The high-score-minus-low-score return spread is –8.11 bps per day (t-value of –2.33). The 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 and the short-term reversal factor.

Row (2) sorts stocks on the predicted Forcerank scores. It again shows a monotonic negative relationship between the predicted scores and stock returns of week $t+1$. The high-score-minus-low-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 and the short-term reversal factors. The six-factor alpha is still –5.47 bps per day (t-value of –1.70).

Row (3) shows that the return predictive power of the residual score is even slightly larger than that of the predicted score in economic magnitude. The high-score-minus-low-score return spread is –6.89 bps per day (t-value of –2.07). The return spread remains significant even after controlling for the Fama-French five factors, the momentum and the short-term reversal factors. The seven-factor alpha is still –6.67 bps per day (t-value of –2.01). In other words, the investor sentiment revealed by the Forcerank score adds incremental return predictive power above and beyond past returns.

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 significant return spreads. In particular, the negative relationship between the past one-month returns and the future one-week returns is not monotonic, explaining the lack of significant trading strategy return, even though past one-month return is marginally significant from the regression in Table 7.

To further address the concern regarding the generalizability of the extrapolative beliefs extracted from Forcerank, we conduct an out-of-sample validation test. We study return predictability among all stocks over a longer period starting from April 9, 2001. If the belief structure estimated from Forcerank data represents the belief formation process of the extrapolators who trade in the market, we would expect that the predicted Forcerank scores for non-Forcerank stocks also have predictive power for future returns. We choose the starting date of the out-of-sample period to be the date of full implementation of decimalization for all equities and options on exchanges, to alleviate the concern that we are simply capturing the short-term return reversal due to the bid-ask bounce and other liquidity issues.

For each stock in each week, we first compute a predicted Forcerank score as the fitted value from the nonlinear regression in Table 4 (Column (2)) using the stock's lagged industry adjusted returns from week $t-11$ to week t. We also compute a second predicted Forcerank score, allowing for asymmetry between positive and negative past returns, as the fitted value from the nonlinear regression in Table 5 (Column (1)). To evaluate the economic magnitude associated with the return predictability, we again consider trading strategies, similar to those in Panel A. The trading strategy results are reported in Panels B and C of Table 9. Stocks whose prices are below five dollars at the beginning of each week are removed to further reduce the impact of illiquidity.

Panel B includes all stocks. Row (1) sorts them on the predicted Forerank scores every week and report the top and bottom decile portfolio performance in daily return over the next week. As in the Forcerank sample, we again observe a monotonic negative relation between the predicted scores and stock returns. The high-score-minus-low-score return spread is –24.4 bps per day (t-value of –14.29). The spread remains highly significant after various risk adjustments. Allowing the asymmetry between positive and negative past returns in Predicted score PN (Row (2)) increases the spread and alphas even more.

For comparison, Rows (3) and (4) 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. While they also produce statistically significant trading strategy returns, the magnitude of the return spreads are smaller to those in Rows (1) and (2). Extrapolative beliefs, by applying declining weights to past weekly returns and allowing different weights on positive and negative returns, predict future return better than past weekly returns over any specific horizons.

Could the return predictability be a simple manifestation of liquidity shocks that cause initial price pressure and subsequent price reversal? To address this concern, in Panel C, we repeat the trading strategies among the largest stocks (those in the top CRSP size quintile) that are least likely to be subject to illiquidity. We find predicted scores to still outperform simple past returns even among this subset of stocks.

V. Conclusion

Taking advantage of novel data from a crowdsourcing platform (Forcerank.com) for ranking stocks, we provide strong empirical evidence that investors extrapolate recent past returns of individual stocks when forming expectations about their future returns. Extrapolation is asymmetric between positive and negative past returns: investors put more weight on negative past returns, and they display a longer memory span for these negative returns. Similarly, investor expectations respond more strongly to salient past returns, and investors have a longer memory span for these salient returns. Furthermore, we link the extrapolation bias to user and firm characteristics. We find a stronger extrapolation bias among users who are not financial professionals. We also find that the extrapolation parameters are affected by firm characteristics including size, book-to-market ratio, return volatility, and turnover.

To examine how investor expectations affect asset prices, we find that consensus rankings negatively predict future stock returns in the cross-section, more so among stocks with low institutional ownership and a high degree of extrapolative bias, consistent with the predictions from a simple asset pricing model of extrapolative beliefs.

Appendices

A. A Cross-Sectional Model with Return Extrapolation

In this section, we study the asset pricing implications of a simple cross-sectional model that features some investors who extrapolate from past returns when forming beliefs about stocks' future returns. We consider a finite-horizon economy with $T+1$ periods, $t = 0, 1, \ldots, T$. There are $N+1$ assets: one risk-free asset with its interest rate normalized to zero; and N risky assets. Risky asset i is a claim to a single dividend payment at the terminal date that is equal to

$$D_{i,T} = D_{i,0} + \varepsilon_{i,1} + \ldots + \varepsilon_{i,T}, \tag{A.1}$$

where

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

The value of $D_{i,0}$ is public information at time 0. Both the market 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 ζ_i on the market 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 traders, fundamental traders and extrapolators. Fundamental traders make up a population fraction μ^F of the economy, and extrapolators make up a population fraction μ^E of the economy; $\mu^E = 1 - \mu^F$. Both types of traders maximize their expected utility defined over next period's wealth with a constant absolute risk aversion coefficient of γ. The key behavioral assumption of the model is that, for risky asset i,

$$\begin{aligned} \mathbb{E}^E_t[\tilde{P}_{i,t+1} - P_{i,t}] &= \lambda_{i,0} + \lambda_{i,1}(1-\lambda_{i,2}) \sum_{k=0}^{\infty} (\lambda_{i,2})^k (P_{i,t-k} - P_{i,t-k-1}) \\ &\equiv \lambda_{i,0} + \lambda_{i,1} S_{i,t}, \end{aligned} \tag{A.3}$$

where $\lambda_{i,1} > 0$ and $\lambda_{i,2} \in (0,1)$. That is, extrapolators' time-t expectation about changes in the price of the 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" $S_{i,t}$.[8] 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

[8] Since our economy begins at $t = 0$, we can write

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

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 $t = 0$.

distant price changes. Empirically, the Forcerank data allow us to estimate the extrapolative belief parameters $\lambda_{i,0}$, $\lambda_{i,1}$, and $\lambda_{i,2}$, up to an affine transformation. We provide a detailed discussion about these parameters in Sections III and IV.

Next, we derive the share demand of traders. We begin with fundamental investors. As mentioned above, each fundamental investor has constant absolute risk aversion (CARA) preferences defined over next period's wealth with risk aversion γ. At time t, she maximizes

$$\max_{N_t^F} \mathbb{E}_t^F \left[-e^{-\gamma(W_t^F + N_t^F(\tilde{P}_{t+1} - P_t))} \right], \tag{A.4}$$

which implies

$$N_t^F = \frac{1}{\gamma}(\Sigma_t^F)^{-1}(\mathbb{E}_t^F[\tilde{P}_{t+1}] - P_t), \tag{A.5}$$

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

$$(\Sigma_t^F)_{i,j} = (\Sigma^F)_{i,j} \equiv \begin{cases} \beta_i^2 \sigma_{\varepsilon,M}^2 + \sigma_{\eta,i}^2 & i = j \\ \beta_i \beta_j \sigma_{\varepsilon,M}^2 & i \neq j \end{cases}. \tag{A.6}$$

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. At time $T-1$, knowing $P_T = D_T \equiv (D_{1,T}, D_{2,T}, \ldots, D_{N-1,T}, D_{N,T})'$,

$$N_{T-1}^F = \frac{1}{\gamma}(\Sigma^F)^{-1}(D_{T-1} - P_{T-1}). \tag{A.7}$$

Market clearing implies

$$\mu^F \frac{1}{\gamma}(\Sigma^F)^{-1}(D_{T-1} - P_{T-1}) + \mu^E N_{T-1}^E = Q, \tag{A.8}$$

where $Q \equiv (Q_1, Q_2, \ldots, Q_{N-1}, Q_N)'$. Rearranging terms gives

$$P_{T-1} = D_{T-1} - (\mu^F)^{-1}\gamma\Sigma^F(Q - \mu^E N_{T-1}^E). \tag{A.9}$$

Imposing that $\mathbb{E}_t^F(N_{t+1}^E) = Q$, a bounded rationality assumption that fundamental traders expect that other people in the market will demand the per-capita supply of the risky assets over the next period,

$$N_{T-2}^F = \frac{1}{\gamma}(\Sigma^F)^{-1}(\mathbb{E}_{T-2}^F[\tilde{P}_{T-1}] - P_{T-2}) = \frac{1}{\gamma}(\Sigma^F)^{-1}(D_{T-2} - \gamma\Sigma^F Q - P_{T-2}). \tag{A.10}$$

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

$$N_t^F = \frac{1}{\gamma}(\Sigma^F)^{-1}(D_t - \gamma(T-t-1)\Sigma^F Q - P_t), \tag{A.11}$$

where $D_t \equiv (D_{1,t}, D_{2,t}, \ldots, D_{N-1,t}, D_{N,t})'$ and $P_t \equiv (P_{1,t}, P_{2,t}, \ldots, P_{N-1,t}, P_{N,t})'$.

Same as fundamental traders, each extrapolator also has constant absolute risk aversion (CARA) preferences defined over next period's wealth with risk aversion γ. At time t, she maximizes

$$\max_{N_t^E} \mathbb{E}_t^E \left[-e^{-\gamma(W_t^E + N_t^E(\tilde{P}_{t+1} - P_t))}\right], \tag{A.12}$$

which implies

$$N_t^E = \frac{1}{\gamma}(\Sigma_t^E)^{-1}(\mathbb{E}_t^E[\tilde{P}_{t+1}] - P_t). \tag{A.13}$$

We further make the assumptions that

$$\Sigma_t^E = \Sigma_t^F = \Sigma^F, \tag{A.14}$$

and note

$$\mathbb{E}_t^E[\tilde{P}_{i,t+1} - P_{i,t}] = \lambda_{i,0} + \lambda_{i,1}(1-\lambda_{i,2}) \sum\nolimits_{k=0}^{\infty} (\lambda_{i,2})^k (P_{i,t-k} - P_{i,t-k-1}) \equiv \lambda_{i,0} + \lambda_{i,1} S_{i,t}. \tag{A.15}$$

The first-order condition of (A.12) then gives rise to the time-t per-capita share demand of extrapolator

$$N_t^E = \frac{1}{\gamma}(\Sigma^F)^{-1} X_t, \tag{A.16}$$

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

Intuitively, equation (A.11) suggests 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 negatively related to the risky asset prices. On the other hand, equation (A.16) suggests that extrapolator demand is positively related to the levels of sentiment.

Market clearing conditions imply that the price of the risky asset i is

$$P_{i,t} = \frac{D_{i,t} + (\mu^F)^{-1}\mu^E[\lambda_{i,0} + \lambda_{i,1}(1-\lambda_{i,2})\sum_{k=1}^{\infty}(\lambda_{i,2})^k (P_{i,t-k} - P_{i,t-k-1}) - \lambda_{i,1}(1-\lambda_{i,2})P_{i,t-1}] + \alpha_{i,t}}{1 - (\mu^E/\mu^F)\lambda_{i,1}(1-\lambda_{i,2})}, \tag{A.17}$$

where $\alpha_{i,t} \equiv -(\gamma(T-t-1)\Sigma^F Q + (\mu^F)^{-1}\gamma\Sigma^F Q)_i$.[9] The pricing equation (A.17) further implies

[9] Our model makes two simplifying assumptions. First, it assumes CARA preferences and therefore eliminates the wealth effect and hence any rebalancing motives. Second, it assumes bounded rationality on the part of fundamental investors—these investors always expect mispricing to be corrected over just one period for all stocks—and therefore further eliminates 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 (A.17) only depends on its own past prices, but not on the past prices of other stocks.

that, in the context of the model, running the predictability regression of the following

$$P_{i,t+1} - P_{i,t} = \hat{\alpha}_i + b_i \cdot S_{i,t} + \hat{\varepsilon}_{i,t+1}, \tag{A.18}$$

where $\hat{\alpha}_i = [1 - (\mu^E/\mu^F)\lambda_{i,1}(1-\lambda_{i,2})]^{-1}\gamma(\Sigma^F Q)_i$ and $\hat{\varepsilon}_{i,t+1} = [1 - (\mu^E/\mu^F)\lambda_{i,1}(1-\lambda_{i,2})]^{-1}\varepsilon_{i,t+1}$, the slope coefficient is

$$b_i = -\frac{(\mu^E/\mu^F)\lambda_{i,1}(1-\lambda_{i,2})}{1-(\mu^E/\mu^F)\lambda_{i,1}(1-\lambda_{i,2})}. \tag{A.19}$$

The empirical analogy to equation (A.18) is to regress realized cumulative returns over the subsequent period on the current level of sentiment constructed from a weighted average of past returns.

The pricing equation (A.17) demonstrates an amplification mechanism: good fundamental news at time t push up the risky asset price $P_{i,t}$, causing extrapolators to increase their share demand on the asset and hence pushing the price further up. Given this, equilibrium only exists if

$$(\mu^E/\mu^F)\lambda_{i,1}(1-\lambda_{i,2}) < 1. \tag{A.20}$$

For condition (A.20) to hold, there needs to be a significant population fraction of fundamental traders who serve as arbitrageurs against mispricing; μ^E/μ^F needs to be sufficiently low. Moreover, a lower degree of extrapolation bias among extrapolators—this leads to a lower $\lambda_{i,1}(1-\lambda_{i,2})$—helps keep condition (A.20) satisfied. The analytical result for the regression coefficient b_i in (A.19) links the stock-level belief-based parameters $\lambda_{i,1}$ and $\lambda_{i,2}$ from extrapolators and the population fraction of these agents μ^E to the degree of return predictability. Specifically, Figure 1 below shows that, for a higher μ^E or a higher $\lambda_{i,1}(1-\lambda_{i,2})$ (a higher $\lambda_{i,1}$ or a lower $\lambda_{i,2}$), the magnitude of the regression coefficient b_i in (A.19) is larger.

[Place Figure 1 about here]

Intuitively, when there are more extrapolators in the economy (a higher μ^E), and when each extrapolator exhibits a higher degree of extrapolation bias (a higher $\lambda_{i,1}(1-\lambda_{i,2})$, the stock is more overvalued, hence giving rise to a stronger degree of return predictability. This model implication is consistent with the empirical findings of Cassella and Gulen (2018a). Using time series data of investor surveys and aggregate stock market prices, they estimate the degree of extrapolation bias (DOX) as an empirical proxy for the aggregate value of λ_2. They find that market return predictability is high during high-DOX periods.

Our individual-stock-level data of investors' return expectations uniquely allow for testing the relation between return predictability and μ^E, $\lambda_{i,1}$, and $\lambda_{i,2}$ in the cross-section. Specifically, our theoretical analysis suggests that return predictability by the sentiment level should be stronger among stocks associated with higher participation of extrapolators and a higher degree of extrapolation bias. We discuss our empirical tests on return predictability in Section IV.

We complete the description of the model by making two remarks. First, there exist conceptual differences between extrapolation in the aggregate market and extrapolation in the cross-section.

According to (A.18), the return for risky asset i is driven by realizations of $\varepsilon_{i,t}$, which have a systematic component $\varepsilon_{M,t}$ and a firm-specific component $\eta_{i,t}$. Given that extrapolators form beliefs about asset i's return based on its total returns in the past (as assumed in (A.3)), these investors extrapolate, to the same extent, both the systematic component and the idiosyncratic component of the past returns. After aggregation, however, firm-specific shocks would have a much less impact on market-wide extrapolation.[10]

The second remark concerns the model's ability to generate momentum. Some extrapolation models (for example, Barberis and Shleifer (2003) and Barberis et al. (2018)) give rise to both momentum and longer-term reversals. Some other extrapolation models (for example, Barberis et al. (2015) and Jin and Sui (2018)), however, only generate reversals. The key difference lies in the models' assumption on the relationship between extrapolators' current return expectation and past returns. If this relationship is assumed to be hump-shaped—that is, if we regress the current expectations on all past returns, the coefficients, when plotted against the passage time between the current time and the time when the past return takes place, display a hump shape—the model generates both momentum and reversals. If, however, this relationship is assumed to be monotonically decreasing, then the model only generates reversals. Ultimately, the relationship between the current expectation and past returns needs to be measured empirically. As we show in Section III, our weekly expectations data find this relationship to be monotonically decreasing. This in turn justifies our belief assumption in (A.3). Moreover, consistent with this documented monotonic relationship, we only observe reversals at the weekly horizon.

B. Evidence from equity analyst target price

We acknowledge the existence of other data source on stock-level investor return expectations. For example, Value Line provides three-to-five year target price on individual stocks at quarterly frequency. The implied long-term expected return forecasts are mostly driven by measures of systematic risk such as the CAPM beta. The sell-side equity analysts provide one-year-ahead target price on individual stocks. Brav, Lehavy, and Michaely (2005) regress the implied next-year

[10] To see this result, consider a special case of the model with symmetry

$$\zeta_i \equiv \zeta, \quad \sigma_{\eta,i} \equiv \sigma_\eta, \quad \lambda_{i,1} \equiv \lambda_1, \quad \lambda_{i,2} \equiv \lambda_2. \tag{A.21}$$

For an equal-weighted market portfolio, its return is

$$P_{M,t+1} - P_{M,t} = \frac{1}{N} \sum\nolimits_{i=1}^{N} P_{i,t+1} - P_{i,t}. \tag{A.22}$$

We then have

$$\begin{aligned} P_{M,t+1} - P_{M,t} =& [1 - (\mu^E/\mu^F)\lambda_1(1-\lambda_2)]^{-1} \\ & \left(\varepsilon_{M,t+1} + \eta_{M,t+1} - (\mu^F)^{-1}\mu^E\lambda_1(1-\lambda_2)^2 \sum\nolimits_{k=0}^{\infty} (\lambda_2)^k (P_{M,t-k} - P_{M,t-k-1}) + \alpha\right), \end{aligned} \tag{A.23}$$

where

$$\eta_{M,t} \sim \mathcal{N}(0, \sigma_\eta^2/N). \tag{A.24}$$

As N goes to infinity, the idiosyncratic component $\eta_{M,t}$ goes to zero. In other words, market-wide sentiment negatively affects future market returns, but its movement becomes independent of the idiosyncratic component of firm returns.

expected return on the past one-year return and find little evidence for extrapolation.

On the surface of it, the equity analyst target price directly measures the investor expectation of the next-year stock price. An important caveat, however, is that these target prices are subject to a "selection bias." Brav and Lehavy (2003) document that analysts are more likely to issue target prices in support of a buy/strong buy recommendation. Consistent with this upward bias, they find the consensus target price to be 32.9% higher than the current market price. Indeed, Da and Schaumburg (2011) show that only the *relative* valuation implied by the price targets of similar stocks is informative. Forcerank focuses on such relative valuation directly: in sharp contrast to the case of equity analyst target price coverage, users on Forcerank.com need to cover all stocks in the contests to form their rankings forecasts.

Nevertheless, to the extent that equity analyst target prices reveal the expectation of sophisticated institutional investors while users on Forcerank.com are more likely to be individual investors, comparing and contrasting these two sets of expectation data could be informative.

In Table A1, we analyze the target price implied next-year stock expected return using consensus target prices collected from I/B/E/S at the end of each year from 1999 to 2015. We regress the target price implied expected returns (TPER) on lagged annual returns. Similar to previous literature, we find little evidence that supports extrapolative expectation. Columns (1) and (2) include all returns in the form of levels. The coefficient on the past year returns, Ret(t), is significantly negative, which could be mechanical since the end-of-year price shows up in TPER via denominator while in Ret(t) via numerator and it is not perfectly synchronized with the consensus target prices used in computing TPER. Interestingly, the coefficients on the lagged returns of year $t-1$ and $t-2$ are sensitive to the sample and become significantly positive after removing illiquid stocks with low prices ($\leq$ \$5). The results are similar when we measure all lagged returns in relative terms (Columns (3) and (4)). Overall, after excluding illiquid stocks, there is suggestive evidence that equity analysts also seem to extrapolate past returns.

[Place Table A1 about here]

REFERENCES

Amromin, Gene, and Steven A. Sharpe, 2013, Expectation of risk and return among household investors: Are their sharpe ratios countercyclical?, Working paper, Federal Reserve Board, Washington DC.

Barber, Brad M., and Terrance Odean, 2008, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *Review of Financial Studies* 21, 785–818.

Barber, Brad M., Terrance Odean, and Ning Zhu, 2009, Systematic noise, *Journal of Financial Markets* 12, 547–569.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, 2015, X-capm: An extrapolative capital asset pricing model, *Journal of Financial Economics* 115, 1–24.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, 2018, Extrapolation and bubbles, *Journal of Financial Economics* 129, 203–227.

Barberis, Nicholas, and Andrei Shleifer, 2003, Style investing, *Journal of Financial Economics* 68, 161–199.

Barberis, Nicholas, and Wei Xiong, 2012, Realization utility, *Journal of Financial Economics* 104, 251–271.

Begenau, Juliane, Maryam Farboodi, and Laura Veldkamp, 2017, Big data in finance and the growth of large firms, Working paper, Stanford University, Princeton University, and New York University.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer, 2017, Diagnostic expectations and stock returns, Working paper, University of Oxford, Bocconi University, Brown University, and Harvard University.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2018, Diagnostic expectations and credit cycles, *Journal of Finance* 73, 199–227.

Brav, Alon, and Reuven Lehavy, 2003, An empirical analysis of analysts' target prices: Short-term informativeness and long-term dynamics, *Journal of Finance* 58, 1933–1967.

Brav, Alon, Reuven Lehavy, and Roni Michaely, 2005, Using expectations to test asset pricing models, *Financial Management* 34, 31–64.

Brown, Gordon D. A., Ian Neath, and Nick Chater, 2007, A temporal ratio model of memory, *Psychological Review* 114, 539–576.

Campbell, John Y., Sanford J. Grossman, and Jiang Wang, 1993, Trading volume and serial correlation in stock returns, *Quarterly Journal of Economics* 108, 905–939.

Cassella, Stefano, and Huseyin Gulen, 2018a, Extrapolation bias and the predictability of stock returns by price-scaled variables, *Review of Financial Studies* 31, 4345–4397.

Cassella, Stefano, and Huseyin Gulen, 2018b, Return expectations, sentiment, and the stock market, Working paper, Tilburg University and Purdue University.

Da, Zhi, Qianqiu Liu, and Ernst Schaumburg, 2013, A closer look at the short-term return reversal, *Management Science* 60, 658–674.

Da, Zhi, and Ernst Schaumburg, 2011, Relative valuation and analyst target price forecasts, *Journal of Financial Markets* 14, 161–192.

Fama, Eugene, 1965, The behavior of stock market prices, *Journal of Business* 38, 34–105.

Gennaioli, Nicola, and Andrei Shleifer, 2018, *A Crisis of Beliefs: Investor Psychology and Financial Fragility* (Princeton University Press).

Greenwood, Robin, Sam Hanson, and Lawrence Jin, 2018, Reflexivity in credit markets, Working paper, Harvard University and California Institute of Technology.

Greenwood, Robin, and Andrei Shleifer, 2014, Expectations of returns and expected returns, *Review of Financial Studies* 27, 714–746.

Hirshleifer, David, Jun Li, and Jianfeng Yu, 2015, Asset pricing in production economies with extrapolative expectations, *Journal of Monetary Economics* 76, 87–106.

Ingersoll, Jonathan E., and Lawrence J. Jin, 2013, Realization utility with reference-dependent preferences, *Review of Financial Studies* 26, 723–767.

Jegadeesh, Narasimhan, 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881–898.

Jegadeesh, Narasimhan, and Sheridan Titman, 1995, Overreaction, delayed reaction, and contrarian profits, *Review of Financial Studies* 8, 973–993.

Jin, Lawrence J., and Pengfei Sui, 2018, Asset pricing with return extrapolation, Working paper, California Institute of Technology.

Koijen, Ralph S. J., Maik Schmeling, and Evert B. Vrugt, 2015, Survey expectations of returns and asset pricing puzzles, Working paper, London Business School.

Kuchler, Theresa, and Basit Zafar, 2016, Personal experiences and expectations about aggregate outcomes, Working paper, New York University and Federal Reserve Bank of New York.

Kuhnen, Camelia M., 2015, Asymmetric learning from financial information, *Journal of Finance* 70, 2029–2062.

Lehmann, Bruce, 1990, Fads, martingales, and market efficiency, *Quarterly Journal of Economics* 105, 1–28.

Maylor, Elizabeth A., Nick Chater, and Gordon D. A. Brown, 2001, Scale invariance in the retrieval of retrospective and prospective memories, *Psychonomic Bulletin and Review* 8, 162–167.

Nagel, Stefan, and Zhengyang Xu, 2018, Asset pricing with fading memory, Working paper, University of Chicago and University of Michigan.

Odean, Terrance, 1998, Are investors reluctant to realized their losses?, *Journal of Finance* 53, 1775–1798.

Piazzesi, Monika, and Martin Schneider, 2009, Momentum traders in the housing market: Survey evidence and a search model, *American Economic Review* 99, 406–411.

Vissing-Jorgensen, Annette, 2004, Perspectives on behavioral finance: Does "irrationality" disappear with wealth? Evidence from expectations and actions, *NBER Macroeconomics Annual 2003* 18, 139–194.

Figure 1. The relation between return predictability and parameters μ^E, λ_1, and λ_2

Panel A plots the slope coefficient for a regression of the future one-week change in price $P_{t+1} - P_t$ on the current sentiment S_t as a function of μ^E, the fraction of extrapolators in the economy. Panel B plots the same regression coefficient as a function of $\lambda_1(1 - \lambda_2)$. The default parameter value is $\mu^E = 0.5$.

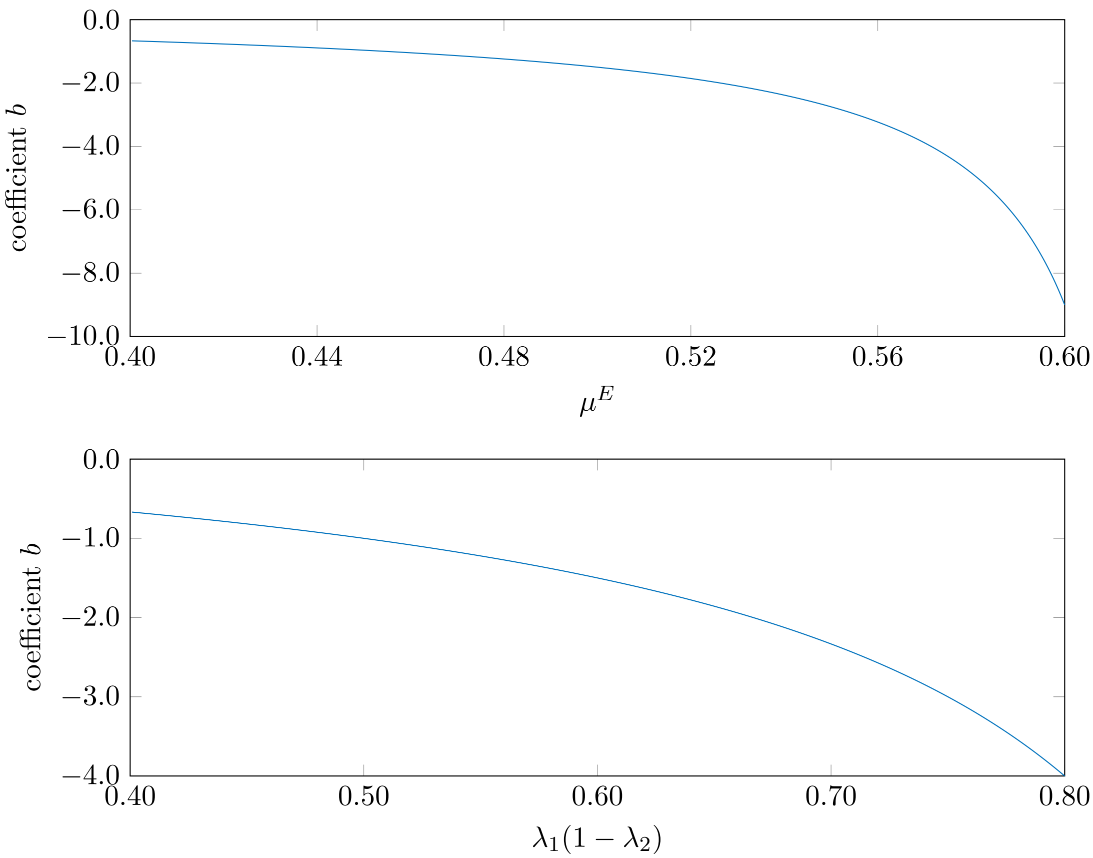


Figure 2. Illustration of the Forcerank interface

The figure on the left presents a screenshot of the interface for a contest for the E-commerce industry which begins at 9:30am June 20th, 2016 and ends at 4:00pm June 24th, 2016. The current time is 11:44am and the time remaining to enter the contest is 3 days 21 hours 45 minutes and 45 seconds. A user could drag the bars next to the company names to rank these stocks. The figure on the right presents a screenshot of the scoring page. The right column under "Live" displays the actual ranking of stocks based on the realized returns during the contest period. The left column under "Your Forcerank" shows the ranking submitted by the user "Aaron" with the corresponding live scores earned for this contest. The scores are based on the difference between the user's rank and the actual rank with a ranking multiplier. More weights are applied to rankings at the top and the bottom.

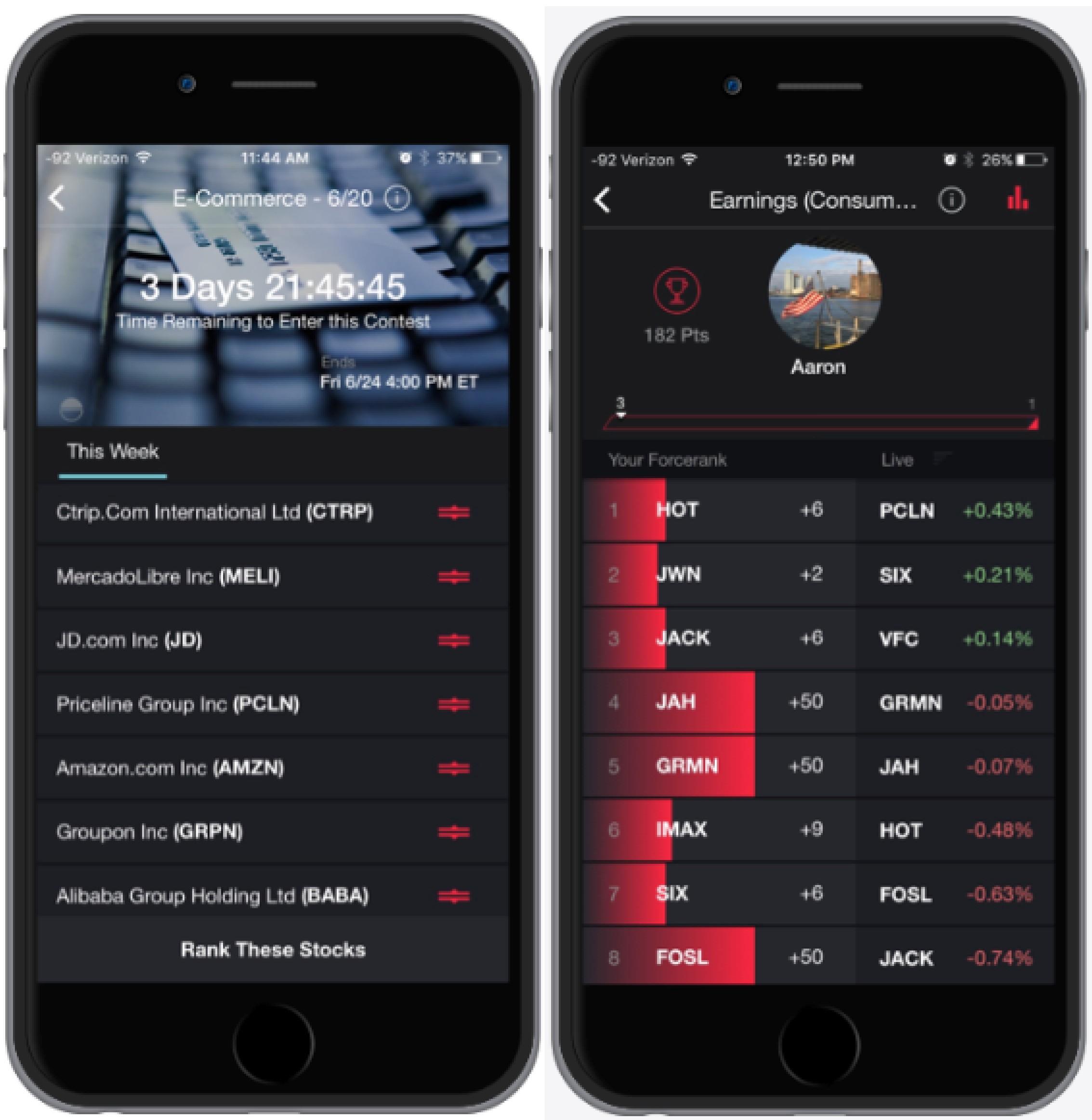


Figure 3. Extrapolative beliefs: Forcerank vs. initial buys

The figure plots the coefficient estimates from regressing the Forcerank consensus rank score on past twelve week returns (left y-axis) and the coefficient estimates from regressing initial buys of retail investors on past twelve week returns (right y-axis). The solid lines correspond to the coefficient estimates, and the dashed lines correspond to the 95% confidence interval. Initial buys are measured using individual-level transaction records from a large discount brokerage firm over the period 1991 to 1996 (as in Odean (1998)). In the initial buy regression, the dependent variable is an indicator variable that equals one if, during week t, there is at least one individual who purchases the stock for the first time in the sample. The initial buys are based on the trades of frequent traders whose median round trip time is less than ten days. Round trip time is defined as the gap between each sale and its most recent purchase. Individuals who have less than ten sales are removed from the sample.

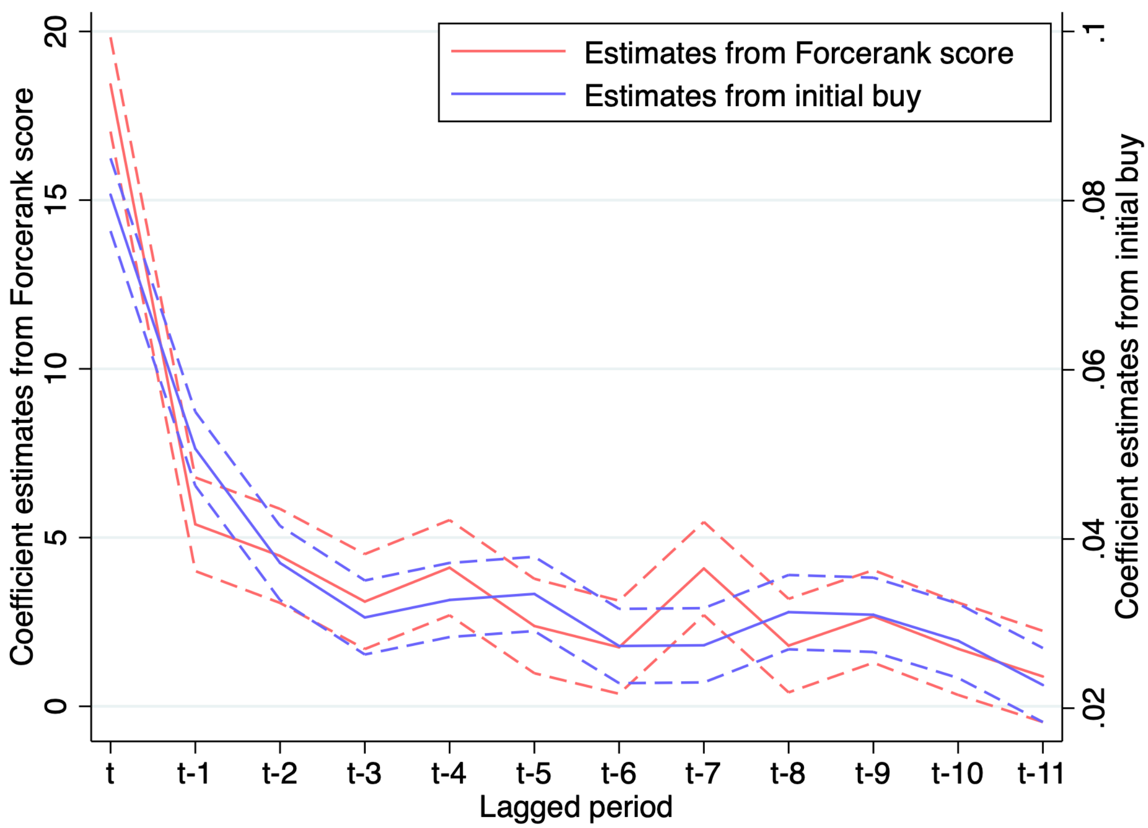
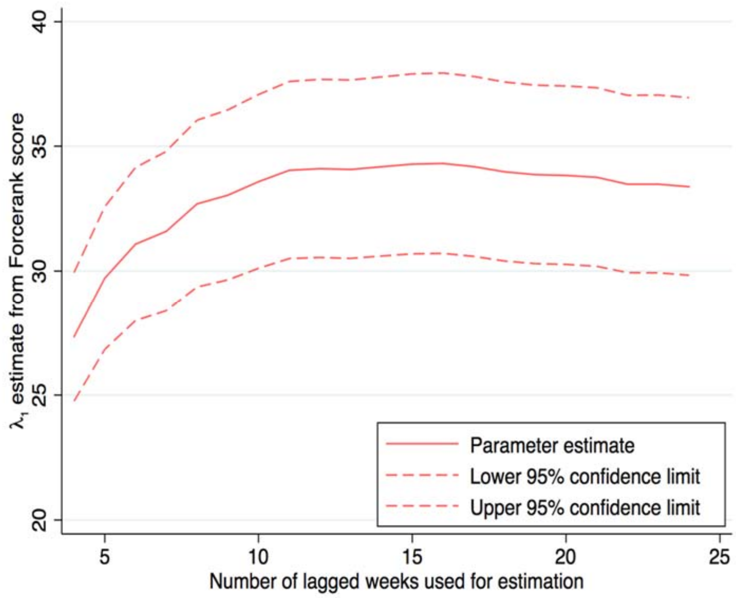


Figure 4. Estimated λ_1 and λ_2 and lagged returns included

The figures plot the estimated λ_1 and λ_2 each as a function of n, the number of lagged returns included in the nonlinear regression specified in equation (2) of the main text:

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



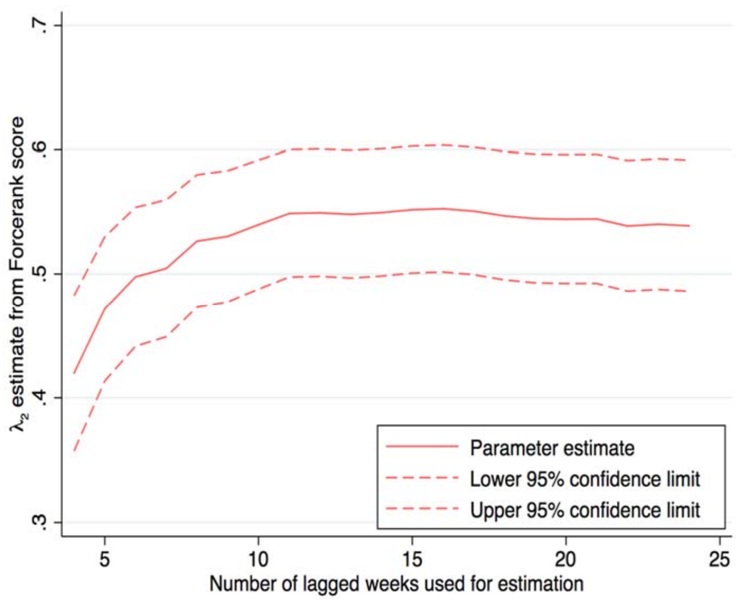


Table 1. Games and contests in the sample

The table presents the number of contests for different types of games in our sample. Each game consists of 10 stocks which all share one or multiple specific characteristics. The same game is repeated conducted every week on the platform, resulting in multiple weekly contests for the same game. There are two main types of games in our sample: (1) industry games which include stocks in a specific sector or industry; (2) most heavily shorted games which include stocks with high short interest in the past month. Other industries include chemicals, pharmaceuticals, oil services, and academics.

Types of games		Number of contests
Industry		1,318
	Enterprise software (Large: 69; Small/Mid: 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
	Other	160
Most heavily shorted (March 2016 to December 2017)		78
Total		1,396

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 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 25 Fama-French size and B/M portfolios based on market capitalization at the end of June of year t , and B/M, 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 self-reported user professional background among a fraction of users who registered before March 2017 (606 out of a total of 1,045 distinct users).

Panel A: Stock and user characteristics

Firm-week-level financial characteristics (Number of observations = 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 quntile	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 = 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
Non professional ($N = 172$)		
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

Table 3. Extrapolative beliefs: Linear regression model

The table presents the results of a contest-level linear regression specified in equation (1) of the main text. The dependent variable is the consensus ranking (1 to 10)—a stock's ranking based on the average ranking across all individuals—the highest ranked stock receives a score of 10; and the lowest ranked stock receives a score of 1. The explanatory variables include lagged returns from week $t-11$ to week t. Column (1) uses the raw level of past returns. Columns (2) and onwards focus on contest-adjusted returns (that is, the raw return in excess of the contest average return). Column (3) uses a ordered logit regression. Column (4) uses the rankings of past returns. Columns (5) and (6) repeat the analysis in Column (2) separately for the first-half (before 3/1/2017) and the second-half (after 3/1/2017) of our sample period. Finally, Column (7) also includes controls of past fundamental news measured by earnings surprises over the past four quarters, tones of news coverage, and the CAPM expected returns. The standard errors are in parenthesis. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep var:	Forcerank Score						
	level	adj	ordered logit	ranking	first half	second half	w controls
Ret(t)	11.21***	16.98***	11.95***	0.276***	17.83***	19.10***	20.65***
	(0.559)	(0.684)	(0.476)	(0.009)	(0.963)	(1.060)	(1.020)
Ret($t-1$)	3.298***	5.139***	3.479***	0.0723***	6.601***	3.897***	5.236***
	(0.555)	(0.679)	(0.448)	(0.009)	(0.953)	(1.058)	(1.014)
Ret($t-2$)	3.150***	4.327***	2.845***	0.0558***	3.993***	4.883***	4.710***
	(0.560)	(0.688)	(0.445)	(0.009)	(0.960)	(1.060)	(1.025)
Ret($t-3$)	2.025***	2.821***	2.013***	0.0486***	3.677***	2.363**	2.935***
	(0.565)	(0.694)	(0.453)	(0.009)	(0.968)	(1.071)	(1.034)
Ret($t-4$)	2.590***	3.703***	2.589***	0.0556***	4.150***	4.174***	5.221***
	(0.564)	(0.695)	(0.455)	(0.009)	(0.973)	(1.070)	(1.034)
Ret($t-5$)	1.911***	2.259***	1.485***	0.0457***	1.616*	3.265***	2.592**
	(0.559)	(0.689)	(0.454)	(0.009)	(0.958)	(1.072)	(1.035)
Ret($t-6$)	0.785	1.188*	1.110**	0.0363***	-0.197	3.952***	1.818*
	(0.542)	(0.668)	(0.443)	(0.009)	(0.946)	(1.054)	(1.024)
Ret($t-7$)	2.146***	3.669***	2.489***	0.0541***	2.418**	5.894***	4.873***
	(0.541)	(0.668)	(0.445)	(0.009)	(0.946)	(1.054)	(1.025)
Ret($t-8$)	0.651	1.503**	1.171***	0.0251***	0.219	3.446***	2.033**
	(0.542)	(0.679)	(0.444)	(0.009)	(0.949)	(1.062)	(1.012)
Ret($t-9$)	1.575***	2.466***	1.732***	0.0431***	1.579*	3.903***	2.394**
	(0.535)	(0.669)	(0.442)	(0.009)	(0.922)	(1.079)	(0.999)
Ret($t-10$)	1.339**	1.529**	1.125**	0.0246***	0.206	3.706***	1.826*
	(0.520)	(0.659)	(0.443)	(0.009)	(0.898)	(1.113)	(1.000)
Ret($t-11$)	0.838	0.812	0.498	0.0140	-0.126	2.258**	0.700
	(0.516)	(0.652)	(0.436)	(0.009)	(0.885)	(1.098)	(0.994)
Observations	12010	12010	12010	12050	5542	6468	7631
R-squared	0.042	0.064	0.0157	0.099	0.074	0.069	0.113

Table 4. Extrapolative beliefs: Exponential decay model

The table presents the results of a contest-level nonlinear regression. The dependent variable is the consensus ranking (1 to 10)—a stock's ranking based on the average ranking across all individuals—the highest ranked stock receives a score of 10; and the lowest ranked stock receives a score of 1. The explanatory variables include lagged returns from week $t-11$ to week t; We implement a nonlinear regression specified in equation (2) of the main text:

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

Column (1) uses the raw level of past returns. Columns (2) and onwards focus on contest-adjusted returns (that is, the raw return in excess of the contest average return). Column (3) uses the rankings of past returns. Columns (4) and (5) repeat the analysis in Column (2) separately for the first-half (before 3/1/2017) and the second-half (after 3/1/2017) of our sample period. Finally, Column (6) also includes controls of fundamentals, news tone, the CAPM expected returns. The standard errors are in parenthesis. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Dep var:	Forcerank Score					
Lagged return in:	level	adj	ranking	first half	second half	w controls
λ_0	5.401***	5.498***	2.789***	5.491***	5.503***	4.675***
	(0.027)	(0.025)	(0.118)	(0.037)	(0.035)	(0.075)
λ_1	23.83***	34.12***	0.493***	35.57***	50.74***	35.22***
	(1.586)	(1.820)	(0.021)	(2.521)	(3.460)	(2.413)
λ_2	0.590***	0.549***	0.471***	0.534***	0.739***	0.461***
	(0.030)	(0.026)	(0.023)	(0.035)	(0.025)	(0.037)
Observations	12010	12010	12010	5542	6468	7631
R-squared	0.037	0.056	0.085	0.070	0.052	0.103

Table 5. Extrapolative beliefs: Past return characteristics

The table presents the results of a contest-level nonlinear regression for various past return characteristics. The dependant variable is the consensus Forcerank score. The regression in Column (1) is specified in equation (3) of the main text:

$$\begin{aligned} \text{Forcerank}_{i,t} &= \lambda_0 + \lambda_{1,pos} \cdot \sum\nolimits_{s=0}^{n} \mathbb{1}_{\{R_{i,t-s} \geq 0\}} \cdot w_{s,pos} R_{i,t-s} \\ &\quad + \lambda_{1,neg} \cdot \sum\nolimits_{s=0}^{n} \mathbb{1}_{\{R_{i,t-s} < 0\}} \cdot w_{s,neg} R_{i,t-s} + \varepsilon_{i,t}, \end{aligned}$$

where $w_{s,pos} = \frac{\lambda_{2,pos}^{s}}{\sum_{j=0}^{n} \lambda_{2,pos}^{j}}$ and $w_{s,neg} = \frac{\lambda_{2,neg}^{s}}{\sum_{j=0}^{n} \lambda_{2,neg}^{j}}$. The regression in Column (2) is specified in equation (4) of the main text:

$$\begin{aligned} \text{Forcerank}_{i,t} &= \lambda_0 + \lambda_{1,sal} \cdot \sum\nolimits_{s=0}^{n} \mathbb{1}_{\{R_{i,t-s} \text{is salient}\}} \cdot w_{s,sal} R_{i,t-s} \\ &\quad + \lambda_{1,nonsal} \cdot \sum\nolimits_{s=0}^{n} \mathbb{1}_{\{R_{i,t-s} \text{is non-salient}\}} \cdot w_{s,nonsal} R_{i,t-s} + \varepsilon_{i,t}, \end{aligned}$$

where $w_{s,sal} = \frac{\lambda_{2,sal}^{s}}{\sum_{j=0}^{n} \lambda_{2,sal}^{j}}$ and $w_{s,nonsal} = \frac{\lambda_{2,nonsal}^{s}}{\sum_{j=0}^{n} \lambda_{2,nonsal}^{j}}$. To measure the level of salience associated with the return in a week, we count the number of news articles on that firm in that week. The news coverage data is obtained from Ravenpack. To ensure extreme returns are not driving our results mechanically, we orthogonalize news coverage on absolute return. Specifically, for each week in the sample period, we run a cross-sectional regression of the total number of news coverage on a stock on its absolute return in the same week. We then define a stock return as salient (non-salient) if the residual from the cross-sectional regression is above (below) median. The standard errors are in parenthesis. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	(1)		(2)
Dep var: Forcerank Score			
λ_0	6.213***	λ_0	5.508***
	(0.047)		(0.025)
$\lambda_{1,pos}$	18.44***	$\lambda_{1,sal}$	55.84***
	(1.674)		(3.462)
$\lambda_{2,pos}$	0.0538	$\lambda_{2,sal}$	0.765***
	(0.063)		(0.021)
$\lambda_{1,neg}$	77.93***	$\lambda_{1,nonsal}$	28.88***
	(3.115)		(2.038)
$\lambda_{2,neg}$	0.818***	$\lambda_{2,nonsal}$	0.356***
	(0.016)		(0.043)
Observations	12010		12010
R-squared	0.081		0.063

Table 6. Extrapolative beliefs: User and stock characteristics

The table in Panel A presents the results of a contest-level nonlinear regression for professional users vs. non-professional users. The regression is specified in equation (2) of the main text. The dependent variable is a stock's consensus ranking averaged across professional users (Columns (1) and (2)) or non-professional users (Columns (3) and (4)). The highest ranked stock receives a score of 10; and the lowest ranked stock receives a score of. 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 firm, we estimate $\lambda_{i,1}$ and $\lambda_{i,2}$ from the following specification:

$$\text{Forcerank}_{i,t} = \lambda_{i,0} + \lambda_{i,1} \cdot \sum\nolimits_{s=0}^{n} w_{i,s} R_{i,t-s} + \varepsilon_{i,t}, \qquad \text{where} \;\; w_{i,s} = \frac{\lambda_{i,2}^{s}}{\sum_{j=0}^{n} \lambda_{i,2}^{j}},\; 0 \le \lambda_{i,2} < 1.$$

Firms with 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) ℓn(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 volatility during our sample period; and (4) turnover: a firm's average weekly turnover during our sample period. The standard errors are in parenthesis. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A: User background

	(1)	(2)
Dep var:	Professionals	Non-Professionals
λ_0	5.498***	5.494***
	(0.029)	(0.028)
λ_1	26.35***	33.77***
	(2.784)	(2.055)
λ_2	0.773***	0.552***
	(0.035)	(0.030)
Observations	9658	10261
R-squared	0.015	0.050

Panel B: Stock characteristics

	(1)	(2)
Dependent variable:	$\lambda_{i,1}$	$\lambda_{i,2}$
ℓn(Size)	8.173**	0.0659***
	(4.074)	(0.024)
B/M	7.034	–0.153*
	(14.299)	(0.083)
Volatility	–12.03**	–0.0142
	(5.221)	(0.030)
Turnover	1.548	0.113**
	(7.602)	(0.044)
Observations	137	137
R-squared	0.174	0.101

Table 7. Return predictability: Fama-MacBeth regression

This table presents the results of Fama-MacBeth forecasting regressions of individual stock returns. For each week t, individuals are asked to rank 10 stocks according to their perceived expected performance of these stocks over week $t+1$. The dependent variable is the daily stock return of week $t+1$. The explanatory variables include the average Forcerank score, variables related to the lagged stock returns, and the residual score orthogonal to past returns. The average Forcerank score is the average of the Forcerank ordinal consensus rankings of the same stock across contests. The predicted score is computed as the fitted value from the nonlinear regression in Table 4 (Column (2)). The residual of this regression is labeled as the residual score. Returns are in daily percent, and the t-statistics are in parenthesis. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable:	Daily return in week $t+1$								
Forcerank score	–0.0246***						–0.0293***		
	(–2.75)						(–3.22)		
Predicted score		–0.0321***						–0.0716***	
		(–4.22)						(–4.41)	
Residual score			–0.0162*						–0.0238***
			(–1.86)						(–2.71)
Ret(t) score				0.00250			–0.0221	0.0245	–0.0255
				(0.25)			(–1.02)	(1.32)	(–1.18)
Ret($t-3,t$) score					–0.0267**		0.00493	0.0264	–0.00112
					(–2.30)		(0.24)	(1.20)	(–0.06)
Ret($t-11,t$) score						–0.00510	0.00576	0.00316	0.00836
						(–0.40)	(0.44)	(0.22)	(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

Table 8. Return predictability: Subsamples

This table presents the results of Fama-MacBeth forecasting regressions of individual stock returns. The dependent variable is the daily stock return of week $t+1$. The explanatory variables include the Forcerank score and variables related to the lagged stock returns. The average Forcerank scores is the average of the Forcerank ordinal consensus rankings of the same stock across contests. The predicted score is computed as the fitted value from the nonlinear regression in Table 4 (Column (2)). In Panel A, all stocks are partitioned into two groups by institutional ownership, which is obtained from the Thomson-Reuters Institutional Holdings (13F) Database and measured at the end of December of 2015. The ownership is set to zero if there is no institution in the database reporting its ownership of the stock. Stocks with low institutional ownership have the fraction of shares owned by institutions below the median. In Panel B, all stocks are partitioned into two groups based on the measure of extrapolation bias, defined as $\lambda_1(1-\lambda_2)$. Returns are in daily percent, and the t-statistics are in parenthesis. *, **, 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***	–0.0125		
	(–3.13)	(–1.12)		
Predicted score			–0.0880***	–0.00617
			(–4.28)	(–0.30)
Ret(t) score	–0.0132	0.0230	0.0408**	0.0146
	(–0.86)	(1.50)	(2.45)	(0.77)
Ret($t-3,t$) score	–0.00234	–0.00568	0.00524	–0.0455**
	(–0.33)	(–0.66)	(0.27)	(–1.99)
Ret($t-11,t$) score	–0.00291	0.00176	0.0398**	–0.00609
	(–0.30)	(0.41)	(2.14)	(–0.34)
Observations	30,014	29,915	30,014	29,915
R-squared	0.148	0.176	0.135	0.171

Panel B: Extrapolation bias

	(1)	(2)	(3)	(4)
Dependent variable:		Ret(t+1)		
	Low bias	High bias	Low bias	High bias
Forcerank score	-0.0196*	-0.235**		
	(-1.72)	(-2.28)		
Predicted score			-0.0558***	-0.161***
			(-3.11)	(-4.89)
Ret(t) score	0.0243**	-0.0112	0.0655***	0.0962***
	(2.27)	(-0.51)	(3.95)	(3.35)
Ret(t-3, t) score	-0.0424***	-0.0154	-0.0456**	0.0626**
	(-6.85)	(-1.38)	(-2.46)	(2.24)
Ret(t-11, t) score	0.0166***	0.0359**	0.0221	-0.0603**
	(4.59)	(2.23)	(1.21)	(-2.28)
Observations	19617	18730	19617	18730
R-squared	0.160	0.234	0.150	0.237

Table 9. Return predictability: Trading strategy

This table shows daily calendar-time portfolio returns. In Panel A, at the beginning of every calendar week $t+1$, all stocks are ranked in ascending order on the basis of their average Forcerank scores, the predicted scores (which are computed as the fitted values from the nonlinear regression in Table 4 (Column (2)), the residual scores (which are the difference between the Forcerank scores and the predicted scores), the lagged contest-adjusted return of week t, and the lagged contest-adjusted monthly return. All 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 with the momentum and the short-term reversal factor. Panels B and C examine an out-of-sample period from April 9, 2001 to December 31, 2017. We also examine a predicted score (PN) from the nonlinear regression allowing for asymmetry between positive and negative past returns (equation (3) of the main text). Panel B includes all stocks; Panel C includes the top size quintile stocks based on CRSP cap-based portfolio assignment. Stocks with a price below five dollars 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 parenthesis. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	L	H	$H-L$ Raw	$H-L$ CAPM	$H-L$ FF5	$H-L$ FF5 +Mom+Rev
Panel A: In-sample						
Forcerank score	0.155***	0.0740*	–0.0811**	–0.0680*	–0.0706**	–0.0705**
	(3.07)	(1.67)	(–2.33)	(–1.96)	(–2.16)	(–2.14)
Predicted score	0.137***	0.0666	–0.0651**	–0.0675**	–0.0679**	–0.0547*
	(2.88)	(1.39)	(–2.01)	(–2.07)	(–2.08)	(–1.70)
Residual score	0.154***	0.0853*	–0.0689**	–0.0635*	–0.0670**	–0.0667**
	(3.03)	(1.81)	(–2.07)	(–1.89)	(–2.03)	(–2.01)
Contest-adjusted Ret(t)	0.128***	0.0803	–0.0367	–0.0441	–0.0429	–0.0351
	(2.70)	(1.54)	(–1.09)	(–1.31)	(–1.28)	(–1.05)
Contest-adjusted Ret($t-3,t$)	0.117**	0.0806	–0.0445	–0.0455	–0.0462	–0.0301
	(2.41)	(1.56)	(–1.25)	(–1.26)	(–1.30)	(–0.86)

	(1)	(2)	(3)	(4)	(5)	(6)
	L	H	$H-L$ Raw	$H-L$ CAPM	$H-L$ FF5	$H-L$ FF5+Mom+Rev
Panel B: Out-of-Sample						
Predicted score	0.242***	–0.00249	–0.244***	–0.238***	–0.246***	–0.215***
	(9.04)	(–0.11)	(–14.29)	(–14.34)	(–14.91)	(–16.59)
Predicted score PN	0.246***	-8.65×10^{-5}	–0.246***	–0.238***	–0.255***	–0.231***
	(8.95)	(–0.00)	(–13.78)	(–13.94)	(–15.38)	(–16.09)
Industry-adjusted Ret(t)	0.224***	0.00274	–0.221***	–0.216***	–0.223***	–0.198***
	(8.44)	(0.12)	(–12.13)	(–12.07)	(–12.55)	(–12.55)
Industry-adjusted Ret($t-3,t$)	0.200***	0.0285	–0.172***	–0.164***	–0.175***	–0.136***
	(7.13)	(1.29)	(–9.08)	(–9.02)	(–9.73)	(–11.01)
Panel C: Out-of-sample large cap						
Predicted score	0.114***	0.0106	–0.104***	–0.0967***	–0.110***	–0.0671***
	(3.29)	(0.36)	(–4.14)	(–3.92)	(–4.50)	(–3.36)
Predicted score PN	0.156***	0.0185	–0.137***	–0.120***	–0.150***	–0.121***
	(3.58)	(0.85)	(–4.18)	(–3.91)	(–5.12)	(–4.50)
Industry-adjusted Ret(t)	0.113***	0.0208	–0.0921***	–0.0843***	–0.0984***	–0.0608**
	(3.17)	(0.70)	(–3.34)	(–3.11)	(–3.68)	(–2.56)
Industry-adjusted Ret($t-3,t$)	0.0842**	0.0252	–0.0590**	–0.0489*	–0.0633**	–0.00689
	(2.29)	(0.87)	(–2.04)	(–1.74)	(–2.27)	(–0.34)

Table A1. Extrapolative expectation: Target price implied expected return

The table presents the results of Fama-MacBeth regressions. The dependent variable is the expected return of year $t+1$ implied by the consensus target price from analysts. The explanatory variables include lagged returns from year $t-3$ to year t. In Columns (1) and (2), all lagged returns are in the form of levels. In Columns (3) and (4), all lagged returns are in the form of percentile rank within the same year; returns are now in relative terms. The sample for Columns (1) and (3) includes all stocks, while the sample for Columns (2) and (4) includes stocks with prices greater than five dollars. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)
Dependent variable:	Target price implied expected return, TPER($t+1$)			
Returns in the form of:	Level	Level	Percentile rank	Percentile rank
Sample:	Full sample	Price > \$5	Full sample	Price > \$5
Ret(t)	–0.340***	–0.169***	–0.385***	–0.356***
	(0.114)	(0.051)	(0.043)	(0.041)
Ret($t-1$)	–0.00947	0.0615**	0.0386	0.0678*
	(0.035)	(0.024)	(0.034)	(0.035)
Ret($t-2$)	0.0216	0.0229**	0.00746	0.0247*
	(0.028)	(0.010)	(0.013)	(0.013)
Ret($t-3$)	–0.0381	0.00577	–0.0364	–0.0268
	(0.025)	(0.018)	(0.027)	(0.026)
Observations	20,606	19,063	20,606	19,063
R-squared	0.060	0.088	0.207	0.191