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

Zhi Da†, Xing Huang‡, Lawrence Jin§

May 5, 2018

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. 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). We find that, when forming expectations, investors extrapolate from past returns, with more weight on more recent returns, especially when recent returns are negative. The extrapolation bias is stronger among Forcerank users who are not financial 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. This return predictability extends to large stocks that are not covered on the platform and is not driven by liquidity-shock-induced price reversals. Finally, the residual component of the consensus rankings orthogonal to past stock returns also negatively predicts future returns, suggesting that investor sentiment is above and beyond return extrapolation.

JEL classification: G4, G12

Keywords: Return Extrapolation, Beliefs in the Cross-section, Expectation Formation

*We thank Nicholas Barberis, Michael Ewens, Samuel Hartzmark, Juhani Linnainmaa, Terrance Odean, Jesse Shapiro, David Solomon, and seminar participants at Boston College, Caltech, LA Finance Day Conference, Washington University's Olin Business School, the Yale Junior Finance Conference, and the WAPFIN conference at NYU Stern 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 returns. Recent works by Greenwood and Shleifer (2014), Koijen, Schmeling, and Vrugt (2015), and Kuchler and Zafar (2016) provide convincing evidence of return extrapolation, the notion that investors' expectations about an asset's future return are a positive function of the asset's recent past returns. Recent models of Barberis, 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 (2017) estimate the degree of extrapolation bias (DOX) and find that market return predictability is significant only during high-DOX periods. However, in the cross-section, analyzing extrapolative beliefs and more generally expectation formation 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 for us 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 for ranking stocks (Forcerank.com). Forcerank collects expectations on future stock performance over specified horizons from highly diverse and geographically distributed individuals. Given this dataset, we investigate how individuals form their expectations about future returns on individual stocks and how these return expectations affect asset prices.

We begin our analysis by developing a cross-sectional model of return extrapolation. The model consists of two types of agents, extrapolators and fundamental traders. Extrapolators form expectations about future returns of individual stocks by extrapolating recent past returns of these stocks, and they trade stocks based on these extrapolative beliefs. Fundamental traders, on the other hand, serve as arbitrageurs who correct for mispricing. The model makes specific predictions

[1] Cassella and Gulen (2018) 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 analyze cross-sectional data of *return* expectations.

on how extrapolators' beliefs affect return predictability in the cross-section; we test and confirm these predictions later in the paper. The model also highlights conceptual differences between extrapolation in the aggregate market and extrapolation in the cross-section.

With the model at hand, we turn to the data to study how investors form expectations in the cross-section. 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 past returns when forming expectations about its future return, with more weight on more recent returns: returns four weeks earlier are only about 9% as important as returns in the most recent week. In addition, extrapolation is asymmetric: investors put more weight on *negative* past returns, and they display a longer memory span for these negative returns.

To parsimoniously study the determinants of investor expectations, we further apply an exponential decay function as the weighting scheme on past returns.[2] In doing so, we summarize the *degree of extrapolation bias* from investor expectations into two parameters. The first parameter λ_1—it is a scaling factor multiplied to all the 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 bias 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 bias 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 show that Forcerank participants have a strong degree of extrapolation bias.

In addition, the λ_1 parameter estimated using *rational* expectations is significantly negative. We interpret the magnitude of this parameter as the *degree of mispricing*; it is an equilibrium outcome that incorporates both extrapolator beliefs and how arbitrageurs trade against mispricing. Our model predicts that stocks associated with a larger degree of extrapolation bias also have a higher

[2]Earlier works of Greenwood and Shleifer (2014), Barberis et al. (2015), and Cassella and Gulen (2017) use this specification to study investors' return expectations about the aggregate stock market.

degree of mispricing. In the cross-section, we present empirical evidence that is consistent with this model prediction: we find that stocks with a higher λ_1 estimated from the Forcerank expectations data—these stocks tend to have a higher degree of extrapolation bias—indeed have a more negative λ_1 estimated using rational expectations, suggesting a higher degree of mispricing associated with the extrapolation bias.

We further exploit the cross-sectional heterogeneity of stocks by examining how user characteristics and firm characteristics affect expectation formation. We show that, financial professionals, compared to non-professionals, display a lower degree of extrapolation bias. Specifically, the λ_1 estimate for professionals is significantly lower than that for the non-professionals, suggesting that 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 have a longer memory span for past returns. 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.

Expectation formation from Forcerank users represents the thinking process of many behavioral investors who trade in the market and affect asset prices. We support this argument in several ways. First, we use Forcerank scores as a proxy for extrapolative beliefs and find these scores to negatively predict next-week stock returns in the cross-section. A trading strategy that buys stocks with low scores and sells stocks with high scores generates a significant profit of 7 basis points per day (or, equivalently, almost 18 percent per year) after controlling for the Fama-French five-factors and a short-term reversal factor.

We also compute a predicted score for each stock as a weighted average of its past twelve week returns where the weights are calibrated to the beliefs of the Forcerank users. Interestingly, such a predicted score forecasts the stock's next-week return *better* than its past return over any specific horizon. This result, combined with the fact that our sample contains mostly large stocks, suggests that our results are not driven by liquidity-shock-induced short-term return reversals.

Our findings therefore suggest that extrapolative beliefs also play an important role in explaining the well-documented short-term price reversal. In other words, our paper also contributes to the voluminous literature on short-term reversal starting from Fama (1965), Jegadeesh (1990), and Lehmann (1990).

Furthermore, we find that, consistent with the prediction of our model, return predictability is stronger among stocks whose clienteles are dominated behavioral extrapolators; we use institutional ownership obtained from the Thomson-Reuters Institutional Holdings (13F) Database as an instrument for investor clienteles.

To further confirm the external validity of our findings, we conduct an out-of-sample test: 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, even though the standard short-term return reversal strategies fail to do so.

Interestingly, the Forcerank score negatively forecasts a stock's next-week return even after controlling for *all* past returns of the stock. That is, the residual component of the consensus ranking that is orthogonal to the stock's past returns also negatively predicts its future return. This finding suggests that investor sentiment contains an important component above and beyond the return extrapolation component.

Forcerank data have a number of unique advantages compared to alternative data sources. For example, several other social media platforms (e.g., StockTwits; Seeking Alpha) also collect cross-sectional opinions or expectations about stock performance. However, these platforms mostly collect textual information from contributors that may not be easily converted to precise quantitative information. Another data source for quantitative cross-sectional expectations is the equity analyst one-year-ahead target prices. Different from target prices, our data cover a more diverse group of individuals and therefore allow us to better understand how heterogeneous investors form beliefs in

the cross-section. In addition, our data is not affected by the potential "selection bias" that target prices are subject to (Brav and Lehavy (2003)). It is also not affected by biases that arise from analysts' career concerns and investment banking relations. Notwithstanding these biases, we find suggestive evidence for extrapolative beliefs even among equity analysts, especially after removing the very illiquid penny stocks.

In what follows, we first derive and discuss in Section II a canonical cross-sectional extrapolation model which serves as a guidance for many of our empirical analyses. In Section III, we then describe in details the crowdsourcing platform and our sample. Section IV presents the empirical results that analyze investors' expectation formation. Section V shows the predictive power of Forcerank scores for future stock returns as well as the results from an out-of-sample test. Section VI concludes. Technical details are in the Appendix.

II. Cross-sectional Predictions of an Extrapolative Model

In behavioral finance, there exists a number of extrapolation models that try to explain facts about the aggregate stock market. However, not many extrapolation models have been developed for the cross-section of individual stocks. This is due to two reasons. First, there is lack of direct empirical support for extrapolative expectations in the cross-section.[3] Second, the micro-foundation for how investors form expectations when facing multiple assets remains unclear.[4]

In this section, we develop a simple cross-sectional model of return extrapolation and study its asset pricing implications. We consider a finite-horizon model 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

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

[3] Lakonishok, Shleifer, and Vishny (1994) provide empirical evidence in the cross-section that is consistent with extrapolation. However, they do not directly use expectations data. Brav, Lehavy, and Michaely (2005) examine analysts' expectations and find little evidence for extrapolation.

[4] The possible sources of extrapolative expectations include "belief in the law of small numbers" (Barberis, Shleifer, and Vishny (1998); Rabin (2002)), availability heuristic (Jin (2015); Gennaioli, Shleifer, and Vishny (2012)), and experience effect (Malmendier and Nagel (2011, 2016)), among others. However, these studies of expectation formation have primarily focused on a single risky asset.

where

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

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,

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

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" $S_{i,t}$. The parameter $\lambda_{i,1}$ measures the overall effect of past price changes on extrapolator beliefs. The parameter $\lambda_{i,2}$ measures the weight an extrapolator puts on recent price changes relative to distant price changes. Empirically, the Forcerank data allow us to estimate the 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 IV and V.

The time-t per-capita share demand of fundamental traders is

$$N^F_t = \frac{1}{\gamma} (\Sigma^F)^{-1} \left(D_t - \gamma (T-t-1) \Sigma^F Q - P_t \right), \tag{4}$$

where

$$(\Sigma^F)_{i,j} \equiv \begin{cases} \zeta_i^2 \sigma_{\varepsilon,M}^2 + \sigma_{\eta,i}^2 & i = j \\ \zeta_i \zeta_j \sigma_{\varepsilon,M}^2 & i \neq j \end{cases}, \tag{5}$$

$D_t \equiv (D_{1,t}, D_{2,t}, \ldots, D_{N-1,t}, D_{N,t})'$, $P_t \equiv (P_{1,t}, P_{2,t}, \ldots, P_{N-1,t}, P_{N,t})'$, and $Q \equiv (Q_1, Q_2, \ldots, Q_{N-1}, Q_N)'$.

On the other hand, the time-t per-capita share demand of extrapolators is

$$N_t^E = \frac{1}{\gamma}(\Sigma^F)^{-1} X_t, \tag{6}$$

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

We derive the share demand from fundamental traders and extrapolators in the Appendix. Intuitively, equation (4) 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 (6) 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

$$P_{i,t} = \frac{D_{i,t} + (\mu^F)^{-1} \mu^E [\lambda_{i,0} + \lambda_{i,1}(1-\lambda_{i,2}) \sum_{k=1}^{\infty} (\lambda_{i,2})^k (P_{i,t-k} - P_{i,t-k-1}) - \lambda_{i,1}(1-\lambda_{i,2}) P_{i,t-1}] + \alpha_{i,t}}{1 - (\mu^E/\mu^F)\lambda_{i,1}(1-\lambda_{i,2})}, \tag{7}$$

where $\alpha_{i,t} \equiv -(\gamma(T-t-1)\Sigma^F Q + (\mu^F)^{-1}\gamma\Sigma^F Q)_i$.[5] The pricing equation (7) further implies that, in the context of the model, running the predictability regression of the following

$$P_{i,t+1} - P_{i,t} = \hat{\alpha}_i + b_i \cdot S_{i,t} + \hat{\varepsilon}_{i,t+1}, \tag{8}$$

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

$$b_i = -\frac{(\mu^E/\mu^F)\lambda_{i,1}(1-\lambda_{i,2})}{1 - (\mu^E/\mu^F)\lambda_{i,1}(1-\lambda_{i,2})}. \tag{9}$$

[5] 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 (7) only depends on its own past prices, but not on the past prices of other stocks.

The empirical analogy to equation (8) 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 (7) demonstrates a role for an amplification factor $[1-(\mu^E/\mu^F)\lambda_{i,1}(1-\lambda_{i,2})]^{-1}$: 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

$$(\mu^E/\mu^F)\lambda_{i,1}(1-\lambda_{i,2}) < 1. \tag{10}$$

For condition (10) 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 (10) satisfied. The analytical result for the regression coefficient b_i in (9) 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, a higher $\lambda_{i,1}$, or a lower $\lambda_{i,2}$, the magnitude of the regression coefficient b_i in (9) 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 $\lambda_{i,1}$ and a lower $\lambda_{i,2}$), 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 (2017). 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 V.

We complete the description of the model by discussing the conceptual differences between

extrapolation in the aggregate market and extrapolation in the cross-section. According to (8), 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 $\eta_{i,t}$. Consider a symmetric case in which

$$\beta_i \equiv \beta, \quad \sigma_{\eta,i} \equiv \sigma_\eta, \quad \lambda_{i,1} \equiv \lambda_1, \quad \lambda_{i,2} \equiv \lambda_2. \tag{11}$$

For an equal-weighted market portfolio, its return is

$$P_{M,t+1} - P_{M,t} = \frac{1}{N} \sum\nolimits_{i=1}^{N} P_{i,t+1} - P_{i,t}. \tag{12}$$

We then have

$$\begin{aligned} P_{M,t+1} - P_{M,t} =& [1 - (\mu^E/\mu^F)\lambda_1(1-\lambda_2)]^{-1} \\ & \left(\varepsilon_{M,t+1} + \eta_{M,t+1} - (\mu^F)^{-1}\mu^E\lambda_1(1-\lambda_2)^2 \sum\nolimits_{k=0}^{\infty} (\lambda_2)^k (P_{M,t-k} - P_{M,t-k-1}) + \alpha\right), \end{aligned} \tag{13}$$

where

$$\eta_{M,t} \sim \mathcal{N}(0, \sigma_\eta^2/N). \tag{14}$$

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. Equations (8) and (13) highlight the difference between extrapolation in the aggregate market and extrapolation in the cross-section: firm-specific shocks have a direct impact on firm-level extrapolation but have a much less impact on market-wide extrapolation.

III. Data and Summary Statistics

In this section, we provide description for the data from Forcerank.com. Forcerank is a crowd-sourcing 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]

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).

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.

IV. Empirical Analysis: Expectation Formation

In this section, we study the formation of investor expectations using the Forcerank data. To start, we analyze how past stock returns affect investors' 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.

IV.1. Expectations and past returns: 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:

$$\text{Forcerank}_{i,t} = \gamma_0 + \sum\nolimits_{s=0}^{n} \beta_s \cdot R_{i,t-s} + \varepsilon_{i,t}, \tag{15}$$

where Forcerank$_{i,t}$ is the week-t consensus rank based on investors' expected 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 Panel A of 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 12 weekly returns are all positive and mostly significant. More important, the coefficients are declining in general, meaning that investors put higher weight on more recent returns.

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), 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. The coefficients on past contest-adjusted returns and the R-squared all increase, indicating a better fit of the data.

To take into account that extrapolation of past returns may be nonlinear, Columns (3) and (4) consider different coefficients on the positive contest-adjusted returns and the negative contest-adjusted returns. The results show that extrapolation is *asymmetric*: it is stronger on the negative side. In particular, individuals seem to put more weight on negative past returns, and they display a longer memory span for negative past returns (relative to contest averages). 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 all past twelve weeks we examine. In other words, a negative return from twelve weeks earlier still affects current expectation formation.

IV.2. Expectations and past returns: Exponential decay model

By using a simple linear specification that allows for independent weights on different past returns, we observe a clear decay pattern in the relation between investor expectations and past returns. To capture this pattern parsimoniously, we now follow our stylized model in Section II and proceed to estimating a parametric extrapolation model which assumes an exponential decay

of weights on past returns. Specifically, we examine an empirical version of equation (3)

$$\text{Forcerank}_{i,t} = \lambda_0 + \lambda_1 \cdot \sum\nolimits_{s=0}^{n} w_s R_{i,t-s} + \varepsilon_{i,t}, \qquad \text{where} \quad w_s = \frac{\lambda_2^s}{\sum_{j=0}^{n} \lambda_2^j}. \tag{16}$$

This exponential decay specification has been previously estimated by Greenwood and Shleifer (2014), Barberis et al. (2015), and Cassella and Gulen (2017) using aggregate stock market data. As discussed previously, it allows us to characterize the relation between investor expectations and past returns by two parameters. The first parameter λ_1—it is a scaling factor multiplied to all the 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 weighted in forming the "sentiment" variable $S_{i,t}$ constructed in (3)—captures a "slope" effect: a lower λ_2 means that investors put higher weights on more recent past returns of stock i. 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. We first estimate the two parameters by assuming them to be constants in the full sample across all stocks and individuals. In Section V, 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 3 about here]

We include various number of lags (n in equation (16)) in the estimation of λ_1 and λ_2. Figure 3 shows that both λ_1 and λ_2 become stable when we include more than 12 past weekly returns in the estimation. We therefore use $n = 12$ for the rest of our analysis.

Panel B of Table 3 confirms the extrapolation bias using the nonlinear specification in (16). For example, Column (2) presents the contest-adjusted results. We find a positive and significant λ_1 of 34.12. At the same time, λ_2 is estimated to be 0.549, which is clearly smaller than one. We focus primarily on the nonlinear specification for the rest of our empirical analysis as it succinctly summarizes the extrapolation bias by two parameters.[6]

[6]Column (3) of Panel B also implement a nonlinear regression using a *rational* ranking of stock i over week t as

IV.3. Expectations and past returns: Professional vs. Non-professional

Our cross-sectional setting uniquely allows us to link extrapolation bias to various user characteristics. To shed some light on the determinants of extrapolation bias, Table 4 examines the extrapolation bias separately for professional and non-professional users.

[Place Table 4 about here]

The results from using raw returns (Columns (1) and (3)) and contest-adjusted returns (Columns (2) and (4)) are very similar. 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 have a shorter memory span and overweight more recent returns. 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 memory span of professionals is also suggested by the data on equity analyst target prices.

In the Appendix, 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. After excluding illiquid stocks, there is suggestive evidence that equity analysts also seem to extrapolate past returns, and the weight only becomes statistically insignificant for returns three years into the past.

IV.4. Determinants of extrapolation parameters

To further study the determinants of extrapolation parameters, we now 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. Table 5 presents these results.

[Place Table 5 about here]

the dependent variable. This ranking is computed based on the stock's realized return over week $t+1$. We provide discussion related to this regression in Section V.

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 to investors. As a result, these returns have a stronger impact on investor expectations.[7] Another related explanation is that data from larger firms are more accessible to investors (**?**). As a result, availability heuristic implies that information about these firms is more important for investors' expectation formation. 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 $\lambda_{i,2}$, suggesting that trading volume-induced salience leads to a longer memory span for investors on past returns. Furthermore, we find that 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, Table 5 also suggests that investors' memory span on past returns is shorter for value stocks compared to growth stocks.

V. Empirical Analysis: Return Predictability in the Cross-section

In this section, we examine cross-sectional return predictability associated with investor expectations. 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 out-of-sample. Finally, we derive, in the cross-section, the theoretical relation between belief-based extrapolative parameters and coefficients estimated from the regression of future (realized) returns on past returns. This relation is then confirmed by utilizing cross-game variations.

V.1. Return predictability of the Forcerank score

We 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 6.

[7] Visibility can be interpreted as a persistently salient feature of an object. Therefore, it is broadly related to the literature on salience theories (**?**; **?**).

[Place Table 6 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 Panel B of Table 3 (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 ($\text{Ret}(t)$), or contest-adjusted past one-month return ($\text{Ret}(t-3,t)$), or contest-adjusted past one-quarter return ($\text{Ret}(t-11,t)$). 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 Section II.

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.

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 Table 7.

[Place Table 7 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 a 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 and the short-term reversal factor. 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 and the short-term reversal factor. The six-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 6.

In sum, Tables 6 and 7 show 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. Furthermore, trading strategies formed on Forcerank scores generate significant abnormal returns. These observations allow us to believe that the expectation formation from Forcerank users represents the thinking process of many behavioral investors who trade in the market and affect asset prices. In other words, even though we have only examined a sample of 293 unique stocks and 1,045 users on the Forcerank platform, our findings have direct implications for a larger set of stocks and a broader group of investors.

V.2. Return predictability: Subsamples and Out-of-sample

The theoretical model in Section II 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]

Table 8 runs the Fama-MacBeth return predictive regressions (as in Table 6) 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 predictabil-

ity of the Forcerank score and the predicted score are only present when more extrapolators trade on the stocks.

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 stocks that are not covered by Forcerank. 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 has predictive power for future returns.

To do this, we compute predicted Forcerank scores for non-Forcerank stocks as weighted averages of their past 12 week returns where the weights are calibrated to the Forcerank data. Specifically, the predicted Forerank score is computed as the fitted value from the nonlinear regression in Panel B of Table 3 (Column (2)) using lagged industry Fama-French 10 industry adjusted returns from week $t-11$ to week t. To evaluate the economic magnitude associated with the return predictability, we again consider trading strategies, similar to those in Table 7. The trading strategy results are reported in Table 9. Stocks whose prices are below five dollars at the beginning of each week are removed.

[Place Table 9 about here]

Panel A includes all stocks not covered by Forcerank. Row (1) sorts them on the predicted Forerank scores every week and report the quintile portfolio performance in 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 –7.4 bps per day (t-value of –5.28). The spread remains highly significant after various risk adjustments.

For comparison, Rows (2) and (3) 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 much smaller to those in Row (1). Extrapolative beliefs, by applying declining weights to past weekly 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 B, we repeat the trading strategies among the largest non-Forcerank stocks (those in the top CRSP size quintile) that are least likely to be subject to illiquidity. Not surprisingly, the standard short-term return reversal strategies are no longer profitable among these largest stocks as evident in Rows (2) and (3). In sharp contrast, predicted Forerank score based on extrapolative beliefs still generates a statistically and economically significant return spread of –3.1 bps per day (t-value of –2.19). These findings therefore suggest that extrapolative beliefs also play an important role in explaining the well-documented short-term price reversal, especially among large stocks.

V.3. Extrapolation parameters across games

In this section, we allow the extrapolation parameters to vary across different games. In other words, we estimate $\lambda_{g,1}$ and $\lambda_{g,2}$ for each game g by the following regression:

$$\text{Forcerank}_{i,g,t} = \lambda_{g,0} + \lambda_{g,1} \cdot \sum\nolimits_{s=0}^{n} w_{g,s} R_{i,g,t-s} + \varepsilon_{i,g,t}, \qquad \text{where} \quad w_{g,s} = \frac{\lambda_{g,2}^{s}}{\sum_{j=0}^{n} \lambda_{g,2}^{j}}. \tag{17}$$

The estimates vary widely across games. Interestingly, games with the lowest $\lambda_{g,1}$ estimates—these estimates are even negative in some cases—mainly include stocks that are "most heavily shorted." This indicates that, for stocks with high short interest, investors do not put much weight on past returns or even hold a contrarian view by putting negative weight on these past returns. Moreover, the $\lambda_{g,2}$ estimates for this type of games are close to zero, indicating that investors have a very short memory span when forming expectations about returns on these most heavily shorted stocks. Among the industry-level games, investors put high weight on past returns (high $\lambda_{g,1}$ estimates) with a long memory span (high $\lambda_{g,2}$ estimates) for industries such as investment banks, fast food, enterprise software, and semiconductor. On the other hand, investors put high weight on past returns (high $\lambda_{g,1}$ estimates) with a short memory span (low $\lambda_{g,2}$ estimates) for industries such as oil services and chemicals.

To further understand the extent to which investor beliefs are biased, we replace the Forcerank rankings data on the left hand side of regression (17) by rankings computed using *realized* stock returns. That is, for stock i over week t, we compute its ranking by its realized return over week $t+1$. In doing so, we effectively impose *rational* expectations. Specifically, we use the following

regression:

$$\text{Rational Ranking}_{i,g,t} = \lambda^r_{g,0} + \lambda^r_{g,1} \cdot \sum_{s=0}^{n} w^r_{g,s} R_{i,g,t-s} + \varepsilon^r_{i,g,t}, \qquad \text{where} \quad w^r_{g,s} = \frac{(\lambda^r_{g,2})^s}{\sum_{j=0}^{n} (\lambda^r_{g,2})^j}, \tag{18}$$

to estimate $\lambda^r_{g,1}$ and $\lambda^r_{g,2}$. Here superscript "r" is an abbreviation for "rational" expectation. By comparing (18) with (8), we find that our extrapolation model in Section II predicts

$$\lambda^r_{g,1} \approx \xi \cdot b_g = -\xi \cdot \frac{(\mu^E/\mu^F)\lambda_{g,1}(1-\lambda_{g,2})}{1-(\mu^E/\mu^F)\lambda_{g,1}(1-\lambda_{g,2})}. \tag{19}$$

That is, the $\lambda^r_{g,1}$ parameter estimated using rational rankings is proportional to the slope coefficient for a regression of future returns of stock i in game g on its past returns; the proportionality coefficient ξ is the scaling multiplier from returns to rankings. The approximation in (19) is due to the fact that the return-to-ranking conversion is not completely linear.

Furthermore, (19) unveils that $\lambda^r_{g,1}$ is jointly affected by the population fraction of extrapolators μ^E, the population fraction of fundamental traders μ^F, as well as extrapolators' beliefs about stock i's returns characterized by $\lambda_{g,1}$ and $\lambda_{g,2}$. Indeed, $\lambda^r_{g,1}$ is estimated using realized returns, which are an equilibrium outcome that incorporates both extrapolator beliefs and how arbitrageurs trade against mispricing. Given these observations, we interpret $\lambda^r_{g,1}$ as the *degree of mispricing*, as it intuitively captures the return predictability of past returns via the extrapolative bias. Our model predicts that stocks associated with a larger degree of extrapolation bias—these stocks typically have a higher λ_1 estimated from Forcerank scores—also have a higher degree of mispricing, and this is precisely where past returns should better predict the future return.

[Place Figure 4 about here]

The top graph in Figure 4 confirms this prediction in the data. We find that, across stocks, the λ_1 estimates from Forcerank scores are indeed strongly negatively correlated with the λ^r_1 estimates using realized returns.

The comparison between (18) and (8) also leads to the following model prediction

$$\lambda^r_{g,2} = \lambda_{g,2}. \tag{20}$$

In our model, return predictability is completely due to behavioral agents extrapolating past returns. As a result, the regression of realized returns on past returns completely recovers the $\lambda_{g,2}$ parameter of extrapolator beliefs, giving rise to (20). The bottom graph in Figure 4 shows that, empirically, $\lambda^r_{g,2}$ and $\lambda_{g,2}$ are strongly positively correlated, consistent with the model prediction. Furthermore, in the cross-section, $\lambda^r_{g,2}$ tends to be higher than $\lambda_{g,2}$. This finding suggests that some other investors in the market may have a longer memory span than the Forcerank participants in our sample.[8]

VI. 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. Our cross-sectional setting allows us to link such an extrapolation bias to stock and user characteristics. For example, we find a stronger extrapolation bias among users who are not financial professionals. We also link the degree of extrapolation bias to 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 asset pricing implications of extrapolative beliefs.

Furthermore, we compute a predicted ranking for each stock as a weighted average of its past twelve week returns where the weights are calibrated to Forcerank users' extrapolative beliefs. Interestingly, such a ranking predicts the stock's future return better than its past return over any specific horizon and even out-of-sample. This finding, combined with the fact that our sample contains mostly large stocks, suggests that our results are not driven by liquidity-shock-induced short-term return reversals. Instead, the return predictability we have documented supports the asset pricing implications of a canonical extrapolation model.

[8] A longer memory span does not directly imply that investors in the market are more rational. A rational investor who is fully aware of the belief structure of extrapolators tends to have a $\lambda^r_{g,2}$ that equals to $\lambda_{g,2}$; this investor is contrarian whenever extrapolators are extrapolative. See Barberis et al. (2015) as an example in which equation (20) holds.

Finally, the residual component of the consensus rankings orthogonal to past stock returns also negatively predicts future returns, suggesting that investor sentiment is above and beyond return extrapolation. We leave it to future research to study the nature and determinant of such investor sentiment.

Appendices

A. Proofs

A micro-foundation for fundamental trader demand in Equation (4).

In this model, there are two types of investors, fundamental traders and extrapolators. Fundamental traders make up a fraction μ^E of the total population, and extrapolators make up the remaining μ^F $(= 1 - \mu^E)$. 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

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

which implies

$$N_t^F = \frac{1}{\gamma}(\Sigma_t^F)^{-1}(\mathbb{E}_t^F[\tilde{P}_{t+1}] - P_t), \tag{A.2}$$

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

$$(\Sigma_t^F)_{i,j} = (\Sigma^F)_{i,j} \equiv \begin{cases} \beta_i^2 \sigma_{\varepsilon,M}^2 + \sigma_{\eta,i}^2 & i = j \\ \beta_i \beta_j \sigma_{\varepsilon,M}^2 & i \neq j \end{cases}. \tag{A.3}$$

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

$$N_{T-1}^F = \frac{1}{\gamma}(\Sigma^F)^{-1}(D_{T-1} - P_{T-1}). \tag{A.4}$$

Market clearing implies

$$\mu^F \frac{1}{\gamma}(\Sigma^F)^{-1}(D_{T-1} - P_{T-1}) + \mu^E N_{T-1}^E = Q, \tag{A.5}$$

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

$$P_{T-1} = D_{T-1} - (\mu^F)^{-1}\gamma\Sigma^F(Q - \mu^E N^E_{T-1}). \tag{A.6}$$

Imposing that $\mathbb{E}^F_t(N^E_{t+1}) = Q$, 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,

$$N^F_{T-2} = \frac{1}{\gamma}(\Sigma^F)^{-1}(\mathbb{E}^F_{T-2}[\tilde{P}_{T-1}] - P_{T-2}) = \frac{1}{\gamma}(\Sigma^F)^{-1}(D_{T-2} - \gamma\Sigma^F Q - P_{T-2}). \tag{A.7}$$

Recursively, demands from fundamental traders are

$$N^F_t = \frac{1}{\gamma}(\Sigma^F)^{-1}(D_t - \gamma(T-t-1)\Sigma^F Q - P_t), \tag{A.8}$$

which is Equation (4) in the main text.

A micro-foundation for extrapolator demand in Equation (6).

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

$$\max_{N^E_t} \mathbb{E}^E_t \left[-e^{-\gamma(W^E_t + N^E_t(\tilde{P}_{t+1} - P_t)} \right], \tag{A.9}$$

which implies

$$N^E_t = \frac{1}{\gamma}(\Sigma^E_t)^{-1}(\mathbb{E}^E_t[\tilde{P}_{t+1}] - P_t). \tag{A.10}$$

We further make the assumptions that

$$\Sigma^E_t = \Sigma^F_t = \Sigma^F, \tag{A.11}$$

and note

$$\mathbb{E}^E_t[\tilde{P}_{i,t+1} - P_{i,t}] = \lambda_{i,0} + \lambda_{i,1}(1-\lambda_{i,2})\sum\nolimits_{k=0}^{\infty}(\lambda_{i,2})^k(P_{i,t-k} - P_{i,t-k-1}) \equiv \lambda_{i,0} + \lambda_{i,1}S_{i,t}. \tag{A.12}$$

The first-order condition of (A.9) then gives rise to Equation (6) in the main text.

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 et al. (2005) regress the implied next-year 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

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, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.

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.

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.

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, 2017, Extrapolation bias and the predictability of stock returns by price-scaled variables, *Review of Financial Studies* forthcoming.

Cassella, Stefano, and Huseyin Gulen, 2018, 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, Andrei Shleifer, and Robert Vishny, 2012, Neglected risks, financial innovation, and financial fragility, *Journal of Financial Economics* 104, 452–468.

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

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., 2015, A speculative asset pricing model of financial instability, Working paper, California Institute of Technology.

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.

Lakonishok, Josef, Andrei Shleifer, and Robert Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541–1578.

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

Malmendier, Ulrike, and Stefan Nagel, 2011, Depression-babies: Do macroeconomic experiences affect risk-taking?, *Quarterly Journal of Economics* 126, 373–416.

Malmendier, Ulrike, and Stefan Nagel, 2016, Learning from inflation experiences, *Quarterly Journal of Economics* 131, 53–87.

Rabin, Matthew, 2002, Inference by believers in the law of small numbers, *Quarterly Journal of Economics* 117, 775–816.

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 λ_1. Panel C plots this regression coefficient as a function of λ_2. The default parameter values are $\mu^E = 0.5$, $\lambda_1 = 1$, and $\lambda_2 = 0.4$.

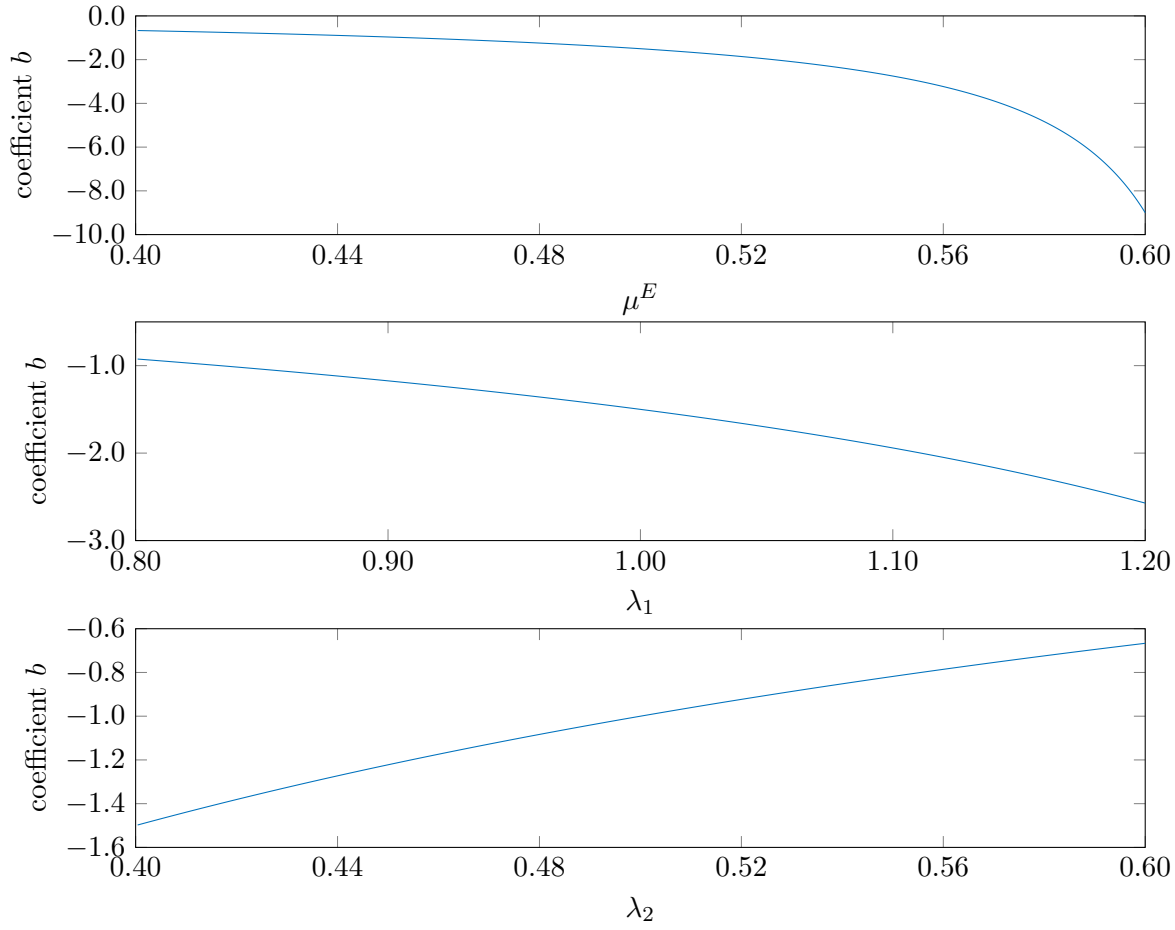


Figure 2. Illustration of the Forcerank interface

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.

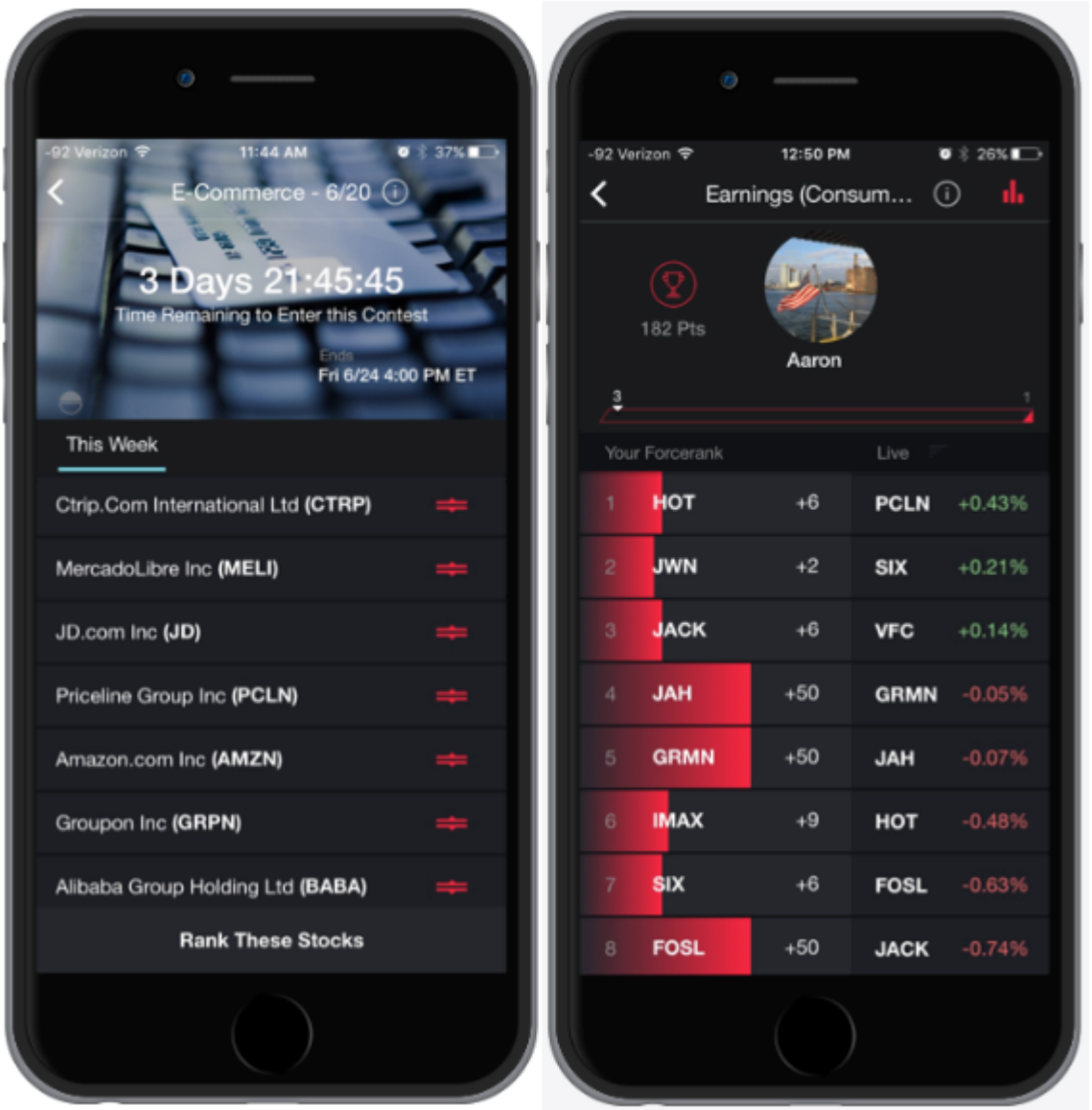


Figure 3. Estimated λ_1 and λ_2 and lagged returns included

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 (16) of the main text:

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

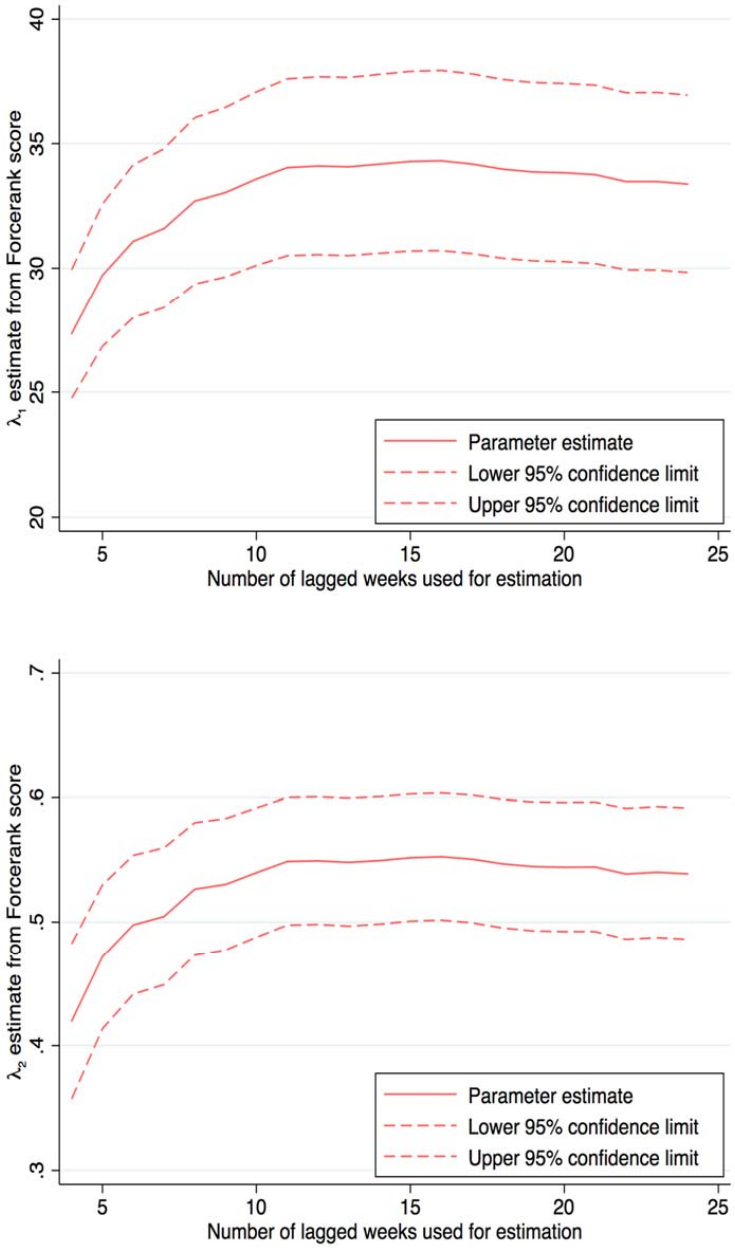


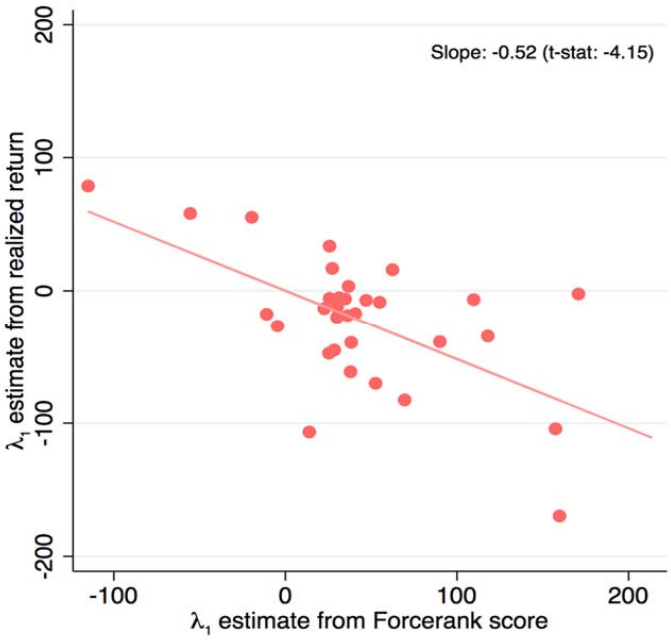
Figure 4. λ_1 and λ_2: Forcerank rankings vs. Realized return-based rankings

The top figure plots, across different games, $\lambda^r_{g,1}$ against $\lambda_{g,1}$; the bottom figure plots $\lambda^r_{g,2}$ against $\lambda_{g,2}$. The estimates of $\lambda_{g,1}$ and $\lambda_{g,2}$ are from (17) of the main text:

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

The estimates of $\lambda^r_{i,1}$ and $\lambda^r_{i,2}$ are from (18) of the main text:

$$\text{Rational Ranking}_{i,g,t} = \lambda^r_{g,0} + \lambda^r_{g,1} \cdot \sum\nolimits_{s=0}^{n} w^r_{g,s} R_{i,g,t-s} + \varepsilon^r_{i,g,t}, \qquad \text{where} \quad w^r_{g,s} = \frac{(\lambda^r_{g,2})^s}{\sum_{j=0}^{n} (\lambda^r_{g,2})^j}.$$



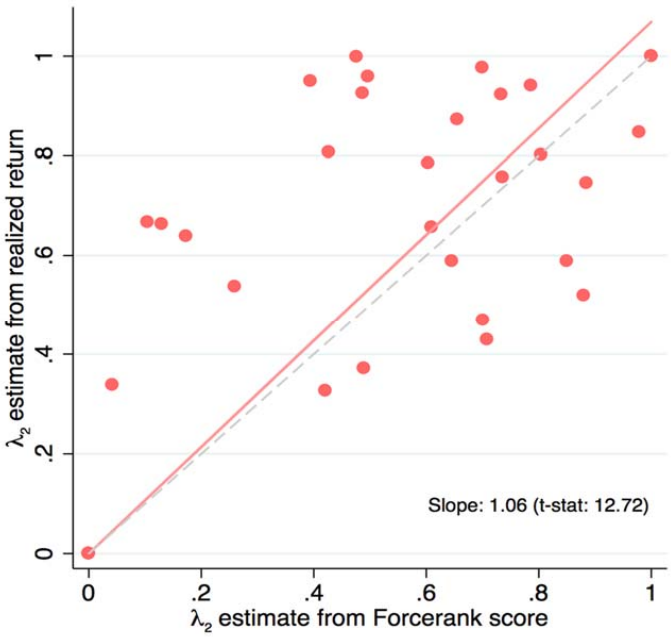


Table 1. Games and contests in the sample

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 belief: Full sample

The table presents the results of contest-level regression. 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 consensus ranking (1 to 10)—a stock's average ranking across all individuals—the 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. The explanatory variables include lagged returns from week $t-12$ to week t. Panel A reports the results of a linear regression specified in equation (15) of the main text. Column (1) uses the raw level of past returns. Columns (2)-(4) use contest-adjusted returns (that is, the raw return in excess of the contest average return). Column (3) reports slope coefficients for a regression in which contest-adjust returns in the negative region are set to zero. Similarly, Column (4) reports slope coefficients for a regression in which contest-adjusted returns in the positive region are set to zero. Panel B implements a nonlinear regression specified in equation (16) of the main text:

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

In addition, Column (3) of Panel B implements a nonlinear regression

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

$\text{Rational Ranking}_{i,t}$ is the a ranking of stock i over week t computed based on its realized return over week $t+1$. We provide discussion related to this regression in Section V. The standard errors are in parenthesis. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A: Linear specification

	(1)	(2)	(3)	(4)
Dependent variable:		Forcerank score		
Lagged return in:	level	contest-adj	pos contest-adj	neg contest-adj
Ret(t)	11.21***	16.98***	8.905***	14.10***
	(0.559)	(0.684)	(0.881)	(1.046)
Ret($t-1$)	3.298***	5.139***	0.712	5.954***
	(0.555)	(0.679)	(0.894)	(1.022)
Ret($t-2$)	3.150***	4.327***	0.497	6.135***
	(0.560)	(0.688)	(0.899)	(1.038)
Ret($t-3$)	2.025***	2.821***	0.718	3.841***
	(0.565)	(0.694)	(0.910)	(1.042)
Ret($t-4$)	2.590***	3.703***	0.394	6.111***
	(0.564)	(0.695)	(0.894)	(1.066)
Ret($t-5$)	1.911***	2.259***	–0.139	4.582***
	(0.559)	(0.689)	(0.898)	(1.035)
Ret($t-6$)	0.785	1.188*	–0.778	2.818***
	(0.542)	(0.668)	(0.851)	(1.032)
Ret($t-7$)	2.146***	3.669***	0.105	4.998***
	(0.541)	(0.668)	(0.845)	(1.024)
Ret($t-8$)	0.651	1.503**	–0.651	2.582**
	(0.542)	(0.679)	(0.865)	(1.017)
Ret($t-9$)	1.575***	2.466***	–0.381	3.726***
	(0.535)	(0.669)	(0.864)	(0.996)
Ret($t-10$)	1.339**	1.529**	0.170	2.179**
	(0.520)	(0.659)	(0.845)	(0.951)
Ret($t-11$)	0.838	0.812	–0.844	2.227**
	(0.516)	(0.652)	(0.842)	(0.936)
Observations	12,010	12,010	12,010	
R-squared	0.042	0.064	0.064	

Panel B: Nonlinear specification

	(1)	(2)	(3)
Dependent variable:	Forcerank score		Realized return
Lagged return in:	level	contest-adj	contest-adj
λ_0	5.401***	5.498***	5.504***
	(0.027)	(0.025)	(0.026)
λ_1	23.83***	34.12***	–5.430**
	(1.586)	(1.820)	(2.263)
λ_2	0.590***	0.549***	0.705***
	(0.030)	(0.026)	(0.157)
Observations	12,010	12,010	12,010
R-squared	0.037	0.056	0.001

Table 4. Extrapolative belief: Professional vs. Non-professional users

The table presents the results of contest-level regression for professional users vs. non-professional users. 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 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 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. The explanatory variables include lagged returns from week $t-11$ to week t. The regression is based on a nonlinear regression specified in equation (16) of the main text. The standard errors are in parenthesis. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)
Dependent variable:	Professional		Non-professional	
Lagged return in:	level	contest-adj	level	contest-adj
λ_0	5.431***	5.498***	5.403***	5.494***
	(0.031)	(0.029)	(0.029)	(0.028)
λ_1	17.68***	26.35***	23.61***	33.77***
	(2.273)	(2.784)	(1.820)	(2.055)
λ_2	0.776***	0.773***	0.608***	0.552***
	(0.042)	(0.035)	(0.034)	(0.030)
Observations	9,658	9,658	10,261	10,261
R-squared	0.010	0.015	0.032	0.050

Table 5. Determinants of extrapolation parameters

The table analyzes how firm-level extrapolation parameters vary with firm characteristics. For each firm, we estimate $\lambda_{i,1}$ and $\lambda_{i,2}$ from the following specification:

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

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; (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.

	(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 6. 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 Panel B of Table 3 (Column (2)). The residual of this regression is labeled as the residual score. 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 7. Return predictability: Trading strategy

This table shows daily calendar-time portfolio returns. 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 Panel B of Table 3 (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 short-term reversal factor. 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)
	$Q1$ (Low)	$Q2$	$Q3$	$Q4$	$Q5$ (High)	$H-L$ Raw	$H-L$ CAPM	$H-L$ FF5	$H-L$ FF5 + Rev
Forcerank score	0.155***	0.0954*	0.120**	0.0837*	0.0740*	–0.0811**	–0.0680*	–0.0706**	–0.0705**
	(3.07)	(1.94)	(2.35)	(1.86)	(1.67)	(–2.33)	(–1.96)	(–2.16)	(–2.14)
Predicted score	0.137***	0.123***	0.120**	0.0858	0.0666	–0.0651**	–0.0675**	–0.0679**	–0.0547*
	(2.88)	(2.62)	(2.52)	(1.64)	(1.39)	(–2.01)	(–2.07)	(–2.08)	(–1.70)
Residual score	0.154***	0.103**	0.0900*	0.0911**	0.0853*	–0.0689**	–0.0635*	–0.0670**	–0.0667**
	(3.03)	(2.18)	(1.75)	(2.08)	(1.81)	(–2.07)	(–1.89)	(–2.03)	(–2.01)
Contest-adjusted Ret(t)	0.128***	0.112**	0.117**	0.101**	0.0803	–0.0367	–0.0441	–0.0429	–0.0351
	(2.70)	(2.42)	(2.47)	(2.20)	(1.54)	(–1.09)	(–1.31)	(–1.28)	(–1.05)
Contest-adjusted Ret($t-3,t$)	0.117**	0.127***	0.135***	0.0733	0.0806	–0.0445	–0.0455	–0.0462	–0.0301
	(2.41)	(2.75)	(2.85)	(1.56)	(1.56)	(–1.25)	(–1.26)	(–1.30)	(–0.86)

Table 8. Return predictability: Subsamples

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 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 Panel B of Table 3 (Column (2)). 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. The t-statistics are in parenthesis. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	(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

Table 9. Return predictability: Out-of-sample

This table shows daily calendar-time portfolio returns for out-of-sample stocks which are not covered on Forcerank platform. Panel A includes all non-Forcerank stocks; Panel B includes the top size quintile non-Forcerank stocks based on CRSP cap-based portfolio assignment. At the beginning of every calendar week $t+1$, all stocks are ranked in ascending order on the basis of their predicted Forcerank score (which is computed as the fitted value from the nonlinear regression in Panel B of Table 3 (Column (2)) using lagged FF10 industry adjusted returns from week $t-11$ to week t), the lagged industry-adjusted return of week t, and the lagged industry-adjusted monthly return. Stocks with a price below five dollars per share at the beginning of each calendar week are removed from the sample. 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 short-term reversal factor. 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)
	$Q1$ (Low)	$Q2$	$Q3$	$Q4$	$Q5$ (High)	$H-L$ Raw	$H-L$ CAPM	$H-L$ FF5	$H-L$ FF5 + Rev
Panel A: Out of sample firms (all)									
Predicted score	0.112***	0.0832***	0.0650***	0.0516**	0.0377	–0.0744***	–0.0652***	–0.0674***	–0.0656***
	(3.70)	(3.29)	(2.79)	(2.23)	(1.48)	(–5.28)	(–4.74)	(–4.92)	(–6.05)
Industry-adjusted Ret(t)	0.0864***	0.0850***	0.0718***	0.0592***	0.0445*	–0.0419**	–0.0307*	–0.0317*	–0.0298**
	(2.72)	(3.48)	(3.28)	(2.66)	(1.65)	(–2.49)	(–1.87)	(–1.93)	(–2.15)
Industry-adjusted Ret($t-3,t$)	0.0923***	0.0784***	0.0689***	0.0604***	0.0437*	–0.0486***	–0.0343*	–0.0366**	–0.0337***
	(2.77)	(3.16)	(3.12)	(2.74)	(1.69)	(–2.67)	(–1.96)	(–2.11)	(–3.04)
Panel B: Out of sample firms (top quintile)									
Predicted score	0.0751***	0.0692***	0.0566**	0.0600***	0.0441*	–0.0310**	–0.0229*	–0.0246*	–0.0227**
	(2.69)	(2.77)	(2.41)	(2.59)	(1.85)	(–2.19)	(–1.65)	(–1.78)	(–2.12)
Industry-adjusted Ret(t)	0.0615**	0.0746***	0.0714***	0.0618***	0.0503**	–0.0112	–0.00201	–0.00195	-7.69×10^{-5}
	(2.08)	(3.06)	(3.22)	(2.75)	(1.98)	(–0.62)	(–0.11)	(–0.11)	(–0.00)
Industry-adjusted Ret($t-3,t$)	0.0603*	0.0730***	0.0647***	0.0693***	0.0385	–0.0218	–0.00782	–0.00919	–0.00606
	(1.94)	(2.97)	(2.89)	(3.08)	(1.61)	(–1.12)	(–0.41)	(–0.49)	(–0.51)

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