

Investor Composition and Overreaction

Michael Blank, Spencer Yongwook Kwon, and Johnny Tang*

January 14, 2023

[\(click here for the latest version\)](#)

Abstract

Do stock price run-ups predictably revert? We develop a model of financial markets with two types of investors: rational investors and “oversensitive” investors who react excessively to salient public news. The model yields a summary statistic for the degree to which a stock price has overreacted to news: the gap in holdings between oversensitive and rational investors. We compute this measure empirically using quarterly institutional holdings data. We first measure each investor’s news sensitivity using their tendency to purchase stocks that have experienced positive earnings announcements. Consistent with our model’s premise, we find that news sensitivity is a persistent investor characteristic. We next aggregate our investor-level measure to the stock level to compute the asset-level holdings gap between oversensitive and rational investors. A larger holdings gap forecasts less continuation in stock prices, and greater reversals in the long-run, especially for extreme price run-ups. Furthermore, our holdings gap aggregates several distinct channels of overreaction, including both price extrapolation and overreaction to non-price information.

*Michael Blank: Harvard University, mblank@g.harvard.edu. Spencer Yongwook Kwon: Harvard University, yongwookkwon@g.harvard.edu. Johnny Tang: Harvard University, jtang01@g.harvard.edu. We thank Francesca Bastianello, Malcolm Baker, Pedro Bordalo, John Campbell, John Conlon, Mark Egan, Paul Fontanier, Xavier Gabaix, Nicola Gennaioli, Robin Greenwood, Sam Hanson, David Laibson, Yueran Ma, Dev Patel, Giorgio Saponaro, Evan Soltas, Andrei Shleifer, Suproteem K. Sarkar, Josh Schwartzstein, Jeremy Stein, Adi Sunderam, David Thesmar, Luis Viceira, and seminar participants at Harvard (Finance and Behavioral) for their thoughtful comments and discussions. Spencer Y. Kwon is grateful for support from the Alfred P. Sloan Foundation Pre-doctoral Fellowship in Behavioral Macroeconomics, awarded through the NBER.

1 Introduction

Large fluctuations in asset prices are a prominent feature of financial markets, from the tulip mania to dot-com stocks to cryptocurrencies (Mackay, 1841; Shiller, 2015). These fluctuations are often interpreted as outcomes of overreaction to salient news and price extrapolation, with short-run rises in prices leading to overshooting and long-run reversals (De Bondt and Thaler, 1985; Lakonishok et al., 1994).¹ However, the overall correlation between short-run price increases and long-run reversals is relatively weak in the data (Fama, 2014; Greenwood et al., 2019). The weak empirical link between run-ups and reversals does not invalidate theories that link short-run momentum to long-run reversals (Barberis et al., 1998; Hong and Stein, 1999; Daniel et al., 1998; Scheinkman and Xiong, 2003), but highlights that overreaction and large price increases are not equivalent. Investors react correctly to some news events, which lead to price run-ups that are sustained, and overreact to others, which lead to run-ups that revert in the long-run. To distinguish between these two cases and sharpen the connection between past run-ups and future reversals, one requires a measure of overreaction that goes beyond price increases.

In this paper, we develop such a measure using investor holdings data. Our approach is based on the assumption, which we validate empirically, that some investors respond more strongly to news than others. This assumption aligns with narrative accounts of market euphoria (Mackay, 1841; Kindleberger, 1978; Brooks, 1999) and studies of institutional investor behavior (Greenwood and Nagel, 2009; Chernenko et al., 2016), which show that some investors systematically participate more in episodes of overreaction.² If an investor’s responsiveness to news, what we call her “news-sensitivity,” is stable across different stocks and time periods, the holdings of highly news-sensitive investors are informative of the degree of overreaction.

Our approach circumvents two major challenges associated with directly analyzing news events to detect overreaction. The first issue is that it is often difficult to identify the specific news event that drove a particular price increase (Cutler et al., 1988). Second, even if we can identify such an event, it is difficult to know how it impacted investor expectations. Specifically, the “salience” of a

¹La Porta (1996) and Bordalo et al. (2019) examine the link between asset price reversals and direct expectations of fundamentals. The broad pattern of short-run momentum and long-run reversals holds across a broad set of asset classes (Asness et al., 2013).

²More generally, persistent heterogeneity in news sensitivity can also be driven by differences in investment styles (Barberis and Shleifer, 2003), relative skill, and clientele (Frazzini and Lamont, 2008; Guercio and Reuter, 2014).

given news event – its tendency to attract investor attention and trigger overreaction – is difficult to measure and depends on a complex array of forces.³ For example, the news event may be featured prominently in the media, or may trigger associations with past successful companies or forward-looking narratives about the new economy. Aggregating these factors and assessing their relative strengths is highly challenging. By exploiting systematic differences in investor news sensitivity, in particular an investor’s tendency to overreact to salient news, our approach yields a stock-level measure of overreaction at all times, even in settings where directly measuring news is difficult.

We begin with a simple model to fix ideas. We modify a static rational expectations equilibrium framework to incorporate heterogeneity in how investors respond to the same public news. Each stock experiences a news event, modeled as a common noisy signal of fundamentals. While rational investors respond to the news in a Bayesian manner, oversensitive investors respond excessively when the news is salient. Each news event varies exogenously in its salience, and hence its propensity to attract excess demand from oversensitive investors. Lastly, to allow for momentum in asset prices, we assume that the asset is supplied by a set of inattentive investors who do not respond to the news event. Under these assumptions, the expected future returns after a price increase is given by the product between the recent price increase and difference in holdings between oversensitive and rational investors, which we call the “holdings gap” of the asset. If the holdings gap is sufficiently low, there is momentum: the current price increase and future returns are positively correlated. Conversely, if the holdings gap is sufficiently high, price increases are expected to revert. Intuitively, the current price increase combines two exogenous quantities: the objective value of the public signal and its salience. The holdings gap reveals the latter to the econometrician and forecasts future returns.

Our model’s prediction that the holdings gap strongly predicts future returns *conditional* on a large price increase separates it from models in which some investors are better informed than others ([Grossman and Stiglitz, 1980](#)). In such models, investor holdings have predictive power for returns – stocks held by uninformed investors will tend to underperform; however this predictability will not be stronger following large price run ups. Thus, we can test whether the holdings gap

³Our use of the term salience broadly follows the notion of salience in the economics literature as the tendency of a stimulus to attract attention and be overweighted in decision making. [Bordalo et al. \(2022b\)](#) review three determinants of salience: a stimulus’ contrast with surroundings, its surprise relative to expectations, and its general prominence. [Li and Camerer \(2022\)](#) and [Bose et al. \(2022\)](#) measure visual salience and show how salient inputs are overweighted in decision-making.

uniquely isolates overreaction by comparing the performance of high holdings gap stocks that have experienced large price run ups with those that have not.

In the full model, we extend our static framework to incorporate dynamics and endogenous learning from prices. Following a news event, investors gradually obtain a more precise estimate of fundamentals from news and prices. We endogenize how rational and oversensitive investors learn from prices by introducing noisy supply (Diamond and Verrecchia, 1981; Grossman and Stiglitz, 1980; Hellwig, 1980) and diagnostic expectations (Bordalo et al., 2018, 2021). A key insight from the dynamic model is that oversensitive investors are *late-stage price extrapolators*: instead of buying stocks that rise immediately following the news event, as momentum traders would (Hong and Stein, 1999), oversensitive investors buy winners long after the news event and thereby earn lower returns. This is because the degree to which oversensitive investors extrapolate from prices is proportional to how informative prices are about fundamentals, which increases over time. In the cross-section, stocks with higher holdings gaps have lower short-run momentum and greater long-run reversals.

In the second part of the paper, we empirically implement the holdings-gap methodology. Our model suggests a two-step procedure to measure overreaction in asset prices: we first measure each investor's sensitivity to public news. For each stock, we then aggregate the investor level measure based on each investor's holdings of the stock. We apply this procedure for large institutional investors that report quarterly holdings data in their 13F filings. We measure an investor's news sensitivity by how her holdings have reacted to past earnings surprises. Our approach is based on two key assumptions, which we validate empirically: an investor's sensitivity to earnings surprises is not only persistent over time, but also correlated with how she reacts to other news beyond earnings. The holdings gap for each stock is then the holdings-weighted average of news-sensitivity of investors who hold the stock.

We then test the core asset-pricing predictions of our model. In all specifications, we lag our holdings gap measure to ensure that our results are truly predictive and not mechanically generated by the overlap between periods we measure news sensitivity and periods we measure returns. First, we find that stocks with a high holdings gap have lower short-run momentum and greater long-run reversals. We find that the positive 12 month auto-correlation in returns is reduced by two-thirds for stocks with high holdings gaps. Our results are stronger when we control for past 3-year returns,

or refine our holdings gap to be based on recent 1-year inflows. To assess whether our measure delivers excess predictability beyond standard factors, we form two momentum portfolios, overreactive and non-overreactive momentum, by double-sorting stocks based on past returns and the most recent holdings gap.⁴ While both momentum portfolios have positive short-run cumulative abnormal returns, only the overreactive momentum portfolio sharply reverses in the long-run, with cumulative 3-year abnormal returns of -20% . By contrast, the cumulative abnormal returns of non-overreactive momentum are both greater in the short-run and more sustained in the long-run. By focusing our attention on stocks with greater investor overreaction, we are able to strengthen the link between short-run momentum and long-run reversals.

We also show that our measure is a *conditional* predictor of returns: stocks with high holdings gap experience large negative returns conditional on significant run-ups, but much less so unconditionally. We define a run-up episode as a stock experiencing more than 100% returns over a one-year period (Greenwood et al., 2019). Among these episodes, those in the highest quintile of holdings gap revert by 17% on average, relative to 7% for episodes in the lowest quintile. Outside of these episodes, we find that high holdings gap stocks do not underperform unconditionally. For comparison, we construct a measure of how skilled the investors holding a stock are, proxying for skill with their past returns. We find that this measure predicts returns regardless of whether a stock has experienced a large price run up or not. Taken together, these results are consistent with the idea that our measure captures overreaction, while other measures based on holdings do not. Finally, we confirm that oversensitive investors are late-stage price extrapolators: across the run-up episodes, we find that the holdings gap systematically increases, indicating inflows from oversensitive investors at the later stages of the run-up. Unconditionally, we find that an investor's news sensitivity is positively correlated with her tendency to buy stocks with recent price increases and lower future benchmarked returns, which tend to have already been bought by other 13F investors.⁵

In the final section of the paper, we test our measure's ability to aggregate investor overreaction to price and non-price information. We analyze the change in a stock's holdings gap around

⁴Precisely, we define overreactive winners (losers) as stocks with high (low) recent returns and high (low) holdings gap. Analogously, non-overreactive winners (losers) are stocks with high (low) returns and low (high) holdings gap. The overreactive (non-overreactive) momentum portfolio is formed by going long overreactive (non-overreactive) winners and going short overreactive (non-overreactive) losers.

⁵While this implies that stocks with high holdings gap underperform on average, we show that the underperformance is concentrated in run-ups. This is consistent with news-sensitive investors more likely to buy the stock conditional on a price increase.

positive earnings announcements, and find that positive news is associated with a multi-quarter increase in the holdings gap, indicating persistent inflow from oversensitive investors. We then study how the increase in the holdings gap varies with announcement characteristics, by regressing post-announcement changes in the holdings gap on past returns and non-price variables.⁶ We consider two salient examples of non-price variables: firm fundamentals, measured by high sales and earnings growth (Bordalo et al., 2019; De La O and Myers, 2021), and industry developments, in particular the performance of top firms in the same industry (Shiller, 2015). We find that the holdings gap not only increases with returns, but also with fundamentals and industry variables holding fixed the price path. This is consistent with our measure capturing investor reaction to non-price information, and rules out the possibility that our measure purely reflects mechanical price extrapolation. Finally, for the same sample, we compare the relative explanatory power of our variables in predicting post-announcement returns. Relative to predicting returns only using past returns and the Fama-French factors, adding the holdings gap further raises the explanatory power by 20%. The gain in explanatory power is greater than that of adding any individual non-price variable or all three variables together, consistent with the holdings gap being an aggregate measure of overreaction.

Related literature Our work relates to four main strands of literature. First, it is part of a large theoretical (Barberis et al., 1998; Hong and Stein, 1999; Daniel et al., 1998; Scheinkman and Xiong, 2003; Rabin and Vayanos, 2010; Bordalo et al., 2021) and empirical (Jegadeesh and Titman, 1993; Lakonishok et al., 1994; Lee and Swaminathan, 2000; Daniel and Moskowitz, 2016; Daniel et al., 2022) literature on short-run momentum and long-run reversals in asset prices. Theoretically, our paper nests both overreaction driven by price increases, which can reflect momentum trading or price extrapolation (Hong and Stein, 1999; Barberis et al., 2018; Cassella and Gulen, 2018), and overreaction driven by news regarding fundamentals (Lakonishok et al., 1994; La Porta, 1996; Bordalo et al., 2022c; De La O and Myers, 2021). We show that our measure responds to both prices and fundamentals, suggesting the importance of both channels. Empirically, a large body of work has sought to find predictors of future returns of momentum and run-ups, such as volatility and issuance (Greenwood et al., 2019), volume (Lee and Swaminathan, 2000), short-

⁶While we do not include announcement quarter returns as they are directly affected by announcement quarter inflows, our results are robust to including them.

sale constraints (Daniel et al., 2022), and market downturns (Daniel and Moskowitz, 2016). We contribute to this literature by proposing a new measure using investor holdings, a theory-driven measure of the degree of overreaction to information.

Our work also relates to the growing literature exploring the determinants of overreaction. Both theoretically and empirically, the literature documents that overreaction depends on a multitude of factors, ranging from fundamental growth (Lakonishok et al., 1994; La Porta, 1996; De La O and Myers, 2021) and its persistence (Bordalo et al., 2020b; Afrouzi et al., 2020), industry developments (Shiller, 2015), and more broadly expectations (Greenwood and Shleifer, 2014). Investor overreaction is also influenced by features of news and information that are harder to measure, such as media sentiment (Tetlock, 2007), narratives (Shiller, 2017), past reference points (Gagnon-Bartsch and Bushong, 2022), and a news' association with extreme fundamentals (Kwon and Tang, 2020). Our methodology provides a measure that aggregates these potential channels.

Our focus on investor composition as a measure of overreaction also relates our work to the empirical literature that predicts returns using investor composition. Koijen and Yogo (2019) show that investor demand driven by asset characteristics can have a powerful ability to predict returns,⁷ and Campbell et al. (2009) show that institutional inflow predicts future announcement returns. While these investor composition measures mix a variety of investor preferences, skill, and biases, our measure specifically captures overreaction to public news. Consistent with our interpretation that we are capturing overreaction, we find that our measure is most predictive of returns after large price increases, which is a property unique to our measure. Closer to our work, Frazzini and Lamont (2008) and Lou (2012) show that mutual fund inflows from retail investors predict negative future returns. Similarly, Lou and Polk (2022) show that momentum portfolios revert in the long-run if it is crowded. Our paper goes further by showing that our measure broadly reflects overreaction to news, where we measure overreaction to a richer set of non-price information beyond extrapolation of past returns.

Lastly, our work relates to the large literature on institutional investor skill. The evidence on the persistence of institutional performance is mixed. Several papers find little persistence in performance beyond those largely explained by style (Daniel et al., 1997; Carhart, 1997), while other papers have found some evidence of persistence in performance (Grinblatt and Titman, 1992), with

⁷Cella et al. (1993) and Coppola (2022) similarly show that investor characteristics predict returns in fire sales.

more recent work showing that investor skill may be more strongly exhibited in a subset of their trades, such as their purchases (Akepanidaworn et al., 2021) and holdings with high conviction (Antón et al., 2021).⁸ Our work finds suggestive evidence that a subset of investors conditionally underperform by overreacting to salient public news. However, due to widespread benchmarking and other passive holdings (Antón et al., 2021), this may only explain a small fraction of their overall performance, which may only be weakly persistent.

The remainder of the paper is organized as follows. Section 2 introduces our model in two steps, starting with a simple static model of overreaction and investor heterogeneity, and then the full model, where we add dynamics and endogenous price overreaction. Turning to the data, Section 3 introduces our measure of investor news sensitivity, and finds that it is a highly persistent investor characteristic. After constructing the overreactive holdings gap from investor news sensitivity, Section 4 tests our model’s core predictions, and in particular show that the measure can predict which short-run price increases will revert in the long-run. Section 5 further investigates the mechanism of how our measure predicts returns, and in particular its ability to capture investor overreaction to price increases and various non-price information. Section 6 concludes.

2 Model: investor composition and overreaction

In this section, we present a model of asset prices with heterogeneous investors and show that relative investor holdings measure how much prices have overreacted to news. We first build intuition by extending a standard static REE model to include price inertia and overreaction. We assume two types of investors: rational investors, who react to public news about fundamentals in a Bayesian way, and oversensitive investors, who excessively react to the same news. We show that returns conditional on a price increase are decreasing the normalized gap in holdings between oversensitive and rational investors in equilibrium, where the gap captures the degree to which oversensitive investors overreact to the public information about the asset. We then present a full version of the model with dynamics and a psychological microfoundation using diagnostic expectations (Bordalo et al., 2018, 2021). In addition to the core predictions, our final model

⁸More recent papers that show lack of persistence in skill include Griffin and Xu (2009), Fama and French (2010), and Lewellen (2011), while Bollen and Busse (2004), Kosowski et al. (2006), Fung et al. (2008), Jagannathan et al. (2010), and Gerakos et al. (2021) find evidence of persistence.

further generates the time-varying tendency of oversensitive investors to extrapolate from price increases and the resulting joint dynamics of returns and investor composition. We relegate detailed derivations and proofs to Appendix A.

2.1 Simple static model

Environment To build intuition, we begin with a simple static model. There is a stock s with a final random payoff $V \sim N(0, \tau_V^{-1})$. The stock experiences a public news event, modeled as a common signal of fundamentals, $n = V + \varepsilon$, with $\text{Var}[\varepsilon] = \tau_\varepsilon^{-1}$. The signal is observed by a continuum of investors, indexed by i , who have a constant-absolute-risk-aversion (CARA) utility with risk-aversion A . Lastly, we assume the asset is supplied by inattentive investors (Hong and Stein, 1999, 2007) that do not react to the news and yield a total supply of assets that is upward sloping in prices: $S(p) = L \cdot p$, with $L > 0$. This cause prices to partially adapt to news and lead to momentum in the absence of overreaction.

Rational-only If all investors are rational and correctly infer fundamentals from the price p and the common signal n , asset demand is given by:

$$D^{rat}(n_i, p) = \frac{1}{A} \frac{E[V|n] - p}{\text{Var}[V|n]} = \frac{1}{A} (\tau_\varepsilon \cdot n - (\tau_V + \tau_\varepsilon)p)^9, \quad (2)$$

where the last identity follows from Bayesian normal updating. Market-clearing implies:

$$L \cdot p = \frac{1}{A} (\tau_\varepsilon n - (\tau_V + \tau_\varepsilon)p) \implies p = \frac{\tau_\varepsilon}{\tau_V + AL + \tau_\varepsilon} n. \quad (3)$$

Due to inattentive investors supplying the asset, asset prices undershoot fundamentals. Averaging across realizations of n and V , one finds momentum in asset prices: conditional on a price increase,

⁹From standard Bayesian updating of normal random variables, we obtain:

$$\begin{aligned} E[V|n] &= \frac{\tau_\varepsilon}{\tau_\varepsilon + \tau_V} n \\ \text{Var}[V|n] &= \frac{1}{\tau_\varepsilon + \tau_V}. \end{aligned} \quad (2)$$

$p > 0$, expected future returns are given by

$$E[V - p|p] = \frac{AL}{\tau_\varepsilon + \tau_V} \cdot p > 0, \quad (4)$$

where future returns ($V - p$) and current returns (p) are positively correlated.

Oversensitive demand To obtain overreaction in asset prices, we introduce oversensitive investors. We assume that a fraction χ of investors are oversensitive, and the remaining $1 - \chi$ rational. For now, we model oversensitive investors in a reduced form, with their demand for asset s , D_s^{os} , given by:

$$D_s^{os}(n, p) = \frac{1}{A} \left(\underbrace{(1 + \Phi_n^{os}(s))\tau_\varepsilon \cdot n}_{\text{News overreaction}} + \underbrace{\Phi_p^{os}(s) \cdot p}_{\text{Price overreaction}} - (\tau_V + \tau_\varepsilon) \cdot p \right). \quad (5)$$

$\Phi_n^{os}(s) \geq 0$ and $\Phi_p^{os}(s) \geq 0$ are exogenous parameters that vary with s which reflect the salience of the news in stock s , or its tendency to attract excess demand from oversensitive investors. For example, salience may depend on the degree of media attention (Tetlock, 2007) and analyst coverage (Hong et al., 2000). It may also be driven by how the news is associated with extreme realization of fundamentals (Kwon and Tang, 2020), fast fundamental growth (Lakonishok et al., 1994; La Porta, 1996; La Porta et al., 1997), or industry-wide developments (Shiller, 2015).¹⁰

Our specification of oversensitive investor demand nests many models of investor behavior. For example, mechanical momentum traders or price extrapolators (Hong and Stein, 1999; Barberis et al., 2018) correspond to $\Phi_p^{os} > 0$ and $\Phi_n^{os} = 0$. Conversely, investors may not learn from prices but over-rely on their private signals (Eyster and Rabin, 2010). Overconfident investors (Scheinkman and Xiong, 2003) who overestimate the precision of their information behave equivalently to an investor with $\Phi_n^{os} > 0$. In the full model, we show that the framework also nests the diagnostic expectations model (Bordalo et al., 2018, 2021).

¹⁰While we assume that $\Phi_n^{os}(s)$ and $\Phi_p^{os}(s)$ are characteristics of asset s in our static model, in the data, we view $\Phi_n^{os}(s)$ and $\Phi_p^{os}(s)$ as properties of each price increase episode, which takes place across time and assets.

Solving the model The equilibrium with oversensitive investors can be solved analogously to the rational benchmark. The rational asset demand is still given by

$$D^{rat}(n, p) = \frac{1}{A} (\tau_\varepsilon \cdot n - (\tau_V + \tau_\varepsilon)p). \quad (6)$$

Combining Equations (5) and (6), market-clearing implies the following price equation:

$$\chi D_s^{os}(n, p) + (1 - \chi) D^{rat}(n, p) = L \cdot p \implies p = \frac{(1 + \chi \Phi_n^{os}(s)) \tau_\varepsilon}{\tau_V + \tau_\varepsilon + AL - \chi \Phi_p^{os}(s)} \cdot n \equiv \psi_s \cdot n. \quad (7)$$

ψ_s measures how asset prices p respond to news n . It is increasing in the salience parameters, $\Phi_n^{os}(s)$ and $\Phi_p^{os}(s)$, and can be seen as the total salience of the news event driving the price increase. Lemma 1 describes the relationship between ψ_s and expected returns.

Lemma 1 (Overreaction and returns). *The expected return following a price increase $p > 0$ is:*

$$E[V - p|p, s] = \left(\frac{\psi^{REE}}{\psi_s} - 1 \right) \cdot p, \quad (8)$$

where $\psi^{REE} \equiv \frac{\tau_\varepsilon}{\tau_V + \tau_\varepsilon}$. *Reversals are expected if and only if $\psi_s > \psi^{REE}$. As one increases $\Phi_n^{os}(s)$ or $\Phi_p^{os}(s)$, asset prices move from momentum to reversals.*

Lemma 1 formalizes the intuitive result that the variation in the future returns of a price increase is driven by the underlying salience of the news event. The benchmark $\psi^{REE} = \frac{\tau_\varepsilon}{\tau_V + \tau_\varepsilon}$ corresponds to the rational benchmark without any inattentive supply: $L = 0$. With $L > 0$, prices exhibit unconditional momentum, $\psi_s < \psi^{REE}$. As ψ_s increases, prices eventually overshoot fundamentals and revert. For the econometrician, however, ψ_s is challenging to directly measure for a specific price increase episode. The literature has pointed to multiple determinants of ψ_s , ranging from fundamentals (Lakonishok et al., 1994), expectations (La Porta, 1996; Greenwood and Shleifer, 2014; Bordalo et al., 2019; De La O and Myers, 2021), media sentiment and investor attention (Tetlock, 2007; DellaVigna and Pollet, 2009), or other intangible information (Daniel and Titman, 2006). Closer to our approach, one can also indirectly infer overreaction from the actions of the supply-side of the asset, as revealed by equity issuance (Greenwood et al., 2019; Lamont and Stein, 2006; Pontiff and Woodgate, 2008; Ma, 2019).

Inferring ψ_s from the holdings gap Relative to the above approaches, we measure overreaction by extracting information from the demand-side of the asset. While ψ_s may be challenging to directly measure, one can identify the set of oversensitive investors by averaging how an investor's asset demand responds to public news across all assets.¹¹ This implies that for each asset s , one can measure D_s^{os} and D_s^{rat} , the average oversensitive and rational holdings of s . We define the overreactive holdings gap of asset s , Gap_s , as the gap between the shares held by oversensitive investors and rational investors, normalized by total shares outstanding.

$$Gap_s = \frac{\chi D_s^{os} - (1 - \chi) D_s^{rat}}{\chi D_s^{os} + (1 - \chi) D_s^{rat}}. \quad (9)$$

Lemma 2 connects our investor holdings measure Gap_s to the strength of total overreaction ψ_s .

Lemma 2 (Overreaction and holdings gap). *The holdings gap for asset s , Gap_s , is given by a monotonically increasing function of ψ_s .*

$$Gap_s = 1 - \frac{2(1 - \chi)(\tau_V + \tau_\varepsilon)}{AL} \left(\frac{\psi^{REE}}{\psi_s} - 1 \right). \quad (10)$$

If $\psi_s < \psi^{REE}$, one has momentum with $Gap_s < 1$: rational investors hold the asset to earn momentum profits. If $\psi_s > \psi^{REE}$, $Gap_s > 1$: rational investors short the asset to profit from expected reversals. In the absence of price and news overreaction ($\Phi_n^{os} = \Phi_p^{os}$), Gap_s is equal to $Gap^0 \equiv 2\chi - 1$. Combining Lemmas 1 and 2, we obtain Proposition 1.

Proposition 1 (Holdings gap and expected returns). *With only rational investors, prices exhibit momentum:*

$$E[V - p|p] = \frac{AL}{\tau_V + \tau_\varepsilon} \cdot p. \quad (11)$$

With oversensitive investors, prices exhibit momentum or reversals, with expected returns:

$$\underbrace{E[V - p|p, s]}_{\text{Future returns}} = \left[1 - \underbrace{\frac{Gap_s - Gap^0}{2(1 - \chi)}}_{\text{Overreaction in holdings}} \right] \cdot \underbrace{\frac{AL}{\tau_V + \tau_\varepsilon}}_{\text{Baseline momentum}} \cdot \underbrace{p}_{\text{Current returns}}. \quad (12)$$

¹¹In the empirical section, we measure an investor's news sensitivity at a given quarter based on how her demand has responded to news across all assets in the past.

Equation (12) captures the core asset pricing prediction of our model. It relates the expected future returns of asset s conditional on a price increase p , $E[V - p|p, s]$, to the interaction between the current price increase p and the contemporaneous overreactive holdings gap Gap_s . For low levels of Gap_s , the model predicts momentum, or positive autocorrelation in returns, with a higher holdings gap implying a lower return autocorrelation. When Gap_s is sufficiently high, expected future returns are negative and proportional to the current return p . In other words, the overreactive holdings gap is a *conditional* predictor of returns: stocks with high overreactive holdings gap will experience negative returns, with the underperformance concentrated in price run-ups. The above prediction separates our model from models in which some investors are better informed than others (Grossman and Stiglitz, 1980). In such models, investor holdings predict returns – stocks held by uninformed investors tend to underperform – but the predictability is not stronger following large price run ups. Thus, we can test whether the holdings gap uniquely isolates overreaction by comparing the performance of high holdings gap stocks that have experienced large price run ups with those that have not.

Proposition 1 summarizes our holdings-based approach to measure overreaction. In contrast to directly estimating ψ_s , which can be difficult, one can infer ψ_s by the relative holdings of over-sensitive investors. The critical prerequisite for our approach is to be able to identify oversensitive investors, and in particular to establish that an investor’s sensitivity to news is a persistent characteristic of that investor. In Section 3, we measure the news sensitivity of an investor based on past trading, and confirm that it is indeed highly persistent across time. Using investor-level news sensitivity, we empirically construct the holdings gap and test Proposition 1 in Section 4.

2.2 Full model: dynamics and psychological microfoundation

Our full model extends the simple model in three ways. First, we introduce dynamics, with our investors learning continuously from news and prices. Second, we make prices partially revealing following Diamond and Verrecchia (1981), Grossman and Stiglitz (1980), and Hellwig (1980), which allows for joint learning from news and prices. Finally, we unify price and news overreaction under a single psychological microfoundation of representativeness, formalized by diagnostic expectations (Bordalo et al., 2018). The full model generates testable predictions regarding the time-varying behavior of oversensitive investors, and the joint dynamics of expected returns and

the holdings gap.

Dynamics Investors learn about fundamentals $V \sim N(0, \tau_V^{-1})$ from two sources of information – news and prices – until prices settle to V at an exogenous horizon $t = T$. For simplicity, we assume that all investors behave like buy-and-hold investors with identical CARA preferences. This implies that investors’ holdings are proportional to the gap between their subjective expected value of fundamentals and prices, normalized by the variance. The assumption means that we abstract away from speculative motives.¹²

News and prices Each investor learns from news by continuously processing $dn_{i,t}$:

$$dn_{i,t} = V dt + \tau_\epsilon^{-1/2} dZ_{i,t}. \quad (13)$$

$\tau_\epsilon^{-1/2}$ represents the rate at which the investor learns from news. $dZ_{i,t}$ reflects idiosyncratic noise in how investor i reacts to the news. By time t , the rational investor extracts from $n_{i,t}$ a cumulative signal of $n_{i,t}/t$ of precision $\tau_\epsilon t$, which averages to V across all investors. Investors also now learn about V from the price path $p_{0:t}$, which is partially revealing due to stochastic supply:

$$S_t = \underbrace{L \cdot p_t}_{\text{inattentive investors}} + \underbrace{q_t}_{\text{random noise trading}} \quad (14)$$

where $L \cdot p_t$ is as before the supply yielded by inattentive agents. q_t is the noisy component of supply, given by a Brownian motion with variance $\text{Var}(dq_t) = \tau_q^{-1} \cdot dt$ (Diamond and Verrecchia, 1981; Hellwig, 1980).¹³ Rational investors learn from prices by positing a linear pricing rule:

$$p_t = a_t F(p_{0:t}) + b_t V - c_t q_t, \quad (15)$$

¹²Adding speculative motives generates a stronger early inflow of rational investors who anticipate selling to oversensitive investors in the future. As t increases, oversensitive investors, who also have speculative motives, will buy even more aggressively than rational investors, which is theoretically explored in Bordalo et al. (2021). Adding speculative motives thus will amplify the qualitative conclusions of our dynamic model.

¹³Our model can easily accommodate supply shocks that follow general Ornstein-Uhlenbeck processes.

where the coefficients a_t , b_t , and c_t are time-varying and known to rational investors. $F(p_{0:t})$ is the signal regarding V , of precision τ_t^P , that rational investors extract from the price path,¹⁴ where F and τ_t^P are equilibrium-consistent.

Diagnostic expectations We microfound how oversensitive investors learn from news and prices by applying diagnostic expectations (Bordalo et al., 2018, 2021), a model of investor psychology based on Kahneman and Tversky’s representativeness heuristic (Tversky and Kahneman, 1973). The model formalizes the notion that price increases and positive news are both representative of high fundamentals: both imply that high values of V have become disproportionately likelier. By overstating the probability of representative outcomes, investors overreact to both news and prices.

Formally, we assume that oversensitive investors, which consist of χ of the total investors, form diagnostic expectations (Bordalo et al., 2018) of fundamentals given prices and private signals:

$$E_t^\theta [V | \mathcal{I}_{t,i}] = E[V | \mathcal{I}_{t,i}] + \theta_s (E[V | \mathcal{I}_{t,i}] - E[V | \mathcal{I}_{0,i}]), \quad (16)$$

where $\mathcal{I}_{t,i} = \{n_{i,t}, p_{0:t}\}$ is the information set of i , which consists of the public price path and the idiosyncratic news process $n_{i,t}$ by time t . The beliefs of oversensitive investors are given by the rational benchmark and an additional term that reflects the overreaction to the total accumulated surprise since $t = 0$. θ_s , the diagnosticity parameter, is assumed to vary across assets (Bordalo et al., 2020b) and is the analogue of ψ_s in the static model.

Following Bordalo et al. (2021), we endogenize how oversensitive investors respond to prices by assuming that they postulate the pricing rule

$$p_t = a_t^\theta F^\theta(p_{0:t}) + b_t^\theta V - c_t^\theta q_t, \quad (17)$$

which is consistent with the equilibrium in which every investor is oversensitive. In other words, oversensitive investors believe that they are rational: they assume that every other investor reacts in the same way as they would to news and prices, and infer V from $p_{0:t}$ in a manner consistent with

¹⁴Given that the average signal across all investors is fixed at V , and the supply shocks dq_t and the idiosyncratic signal process $dZ_{i,t}$ are independent, the public signal and the private signal are two independent signals regarding the fundamental value V .

that assumption.¹⁵ Importantly, this means that oversensitive investors are unaware of the presence of inattentive investors. On the other hand, rational investors are aware of both the existence of inattentive investors and oversensitive investors, which influences how they learn from prices.

Model solution: oversensitive investors as late-stage price extrapolators Appendix A derives the following expressions for the asset demand of oversensitive and rational investors.

Proposition 2 (Investor demand). *Oversensitive and rational asset demand is given by:*

$$\begin{aligned} D_t^\theta(n_{i,t}, p_{0:t}) &= \frac{1}{A} \left(\Phi_{n,t}^\theta \cdot (n_{i,t}/t) + \Phi_p^\theta \cdot p_t - (\tau_V + \tau_\varepsilon t) p_t \right) \\ D_t^{rat}(n_{i,t}, p_{0:t}) &= \frac{1}{A} \left(\Phi_{n,t}^{rat} \cdot (n_{i,t}/t) + \Phi_{p,t}^{rat} \cdot p_t - (\tau_V + \tau_\varepsilon t) p_t \right), \end{aligned} \quad (18)$$

where the news and price sensitivities of each investor are:

$$\begin{aligned} \Phi_{n,t}^\theta &= (1 + \theta_s) \cdot \tau_\varepsilon t, & \Phi_p^\theta &= \frac{(1 + \theta_s)^2 \tau_\varepsilon \tau_q}{A^2 + (1 + \theta_s)^2 \tau_\varepsilon \tau_q} \tau_V \\ \Phi_{n,t}^{rat} &= \tau_\varepsilon t, & \Phi_{p,t}^{rat} &= \frac{(1 + \chi \theta_s) \tau_\varepsilon \tau_q}{A^2 + (1 - \chi)(1 + \chi \theta_s) \tau_\varepsilon \tau_q} \cdot [AL + \tau_V - \chi \Phi_p^\theta - \chi \theta_s \tau_\varepsilon t]. \end{aligned} \quad (19)$$

By endogenizing the price overreaction of oversensitive investors in a dynamic setting, our full model generates a novel prediction that oversensitive investors behave as *late-stage* price extrapolators. Denote:

$$\text{Extrap}_t^\theta = \frac{\text{Cov}(D_t^\theta, p_t)}{\text{Var}[p_t]}, \quad \text{Extrap}_t^{rat} = \frac{\text{Cov}(D_t^{rat}, p_t)}{\text{Var}[p_t]}. \quad (20)$$

Extrap_t^θ can be interpreted as how an oversensitive investor's holdings of asset s comoves with its price increase by time t .¹⁶ Extrap_t^{rat} is the analogous quantity for rational investors. We make the following assumption regarding L :

Assumption 1. *We assume that L is sufficiently high: $L > \frac{A\tau_V}{(1 + \chi \theta_s) \tau_\varepsilon \tau_q}$.*

Assumption 1 is necessary to ensure that there is sufficient initial momentum in asset prices for small t . Proposition 2 then implies the following Corollary.

¹⁵They then combine the signal obtained from news and prices, and overreact to it as specified in Equation (16).

¹⁶ Extrap_t^θ consists of the demand loadings on prices, Φ_p^θ , as well as the correlation induced from learning from the signal $n_{i,t}$, which is also correlated with the price increase.

Corollary 1 (Oversensitive investors as late-stage price extrapolators). *For t small, $Extrap_t^\theta < 0 < Extrap_t^{rat}$: rational investors initially respond to prices more than oversensitive investors. This relationship is reversed as $t \mapsto \infty$: $Extrap_t^\theta > 0 > Extrap_t^{rat}$. Assuming the horizon T is sufficiently large, oversensitive investors on average respond more to prices than rational investors:*

$$Extrap^\theta \equiv \int_0^T Extrap_t^\theta dt > Extrap^{rat} \equiv \int_0^T Extrap_t^{rat} dt.$$

Lastly, oversensitive investors earn lower risk-adjusted returns relative to rational investors:

$$\Pi_{adj}^\theta \equiv \int_{t=0}^T E_{adj} \left[D_t^\theta(n_{i,t}, p_{0:t})(V - p_t) \right] dt < \Pi_{adj}^{rat} \equiv \int_{t=0}^T E_{adj} \left[D_t^{rat}(n_{i,t}, p_{0:t})(V - p_t) \right] dt,$$

where $E_{adj}[X] = E[X] - \frac{\lambda}{2} \text{Var}[X]$ is the risk-adjusted expectation of the random variable X .

In contrast to models (Hong and Stein, 1999; Barberis et al., 2018) that specify a fixed investor response to a price increase, e.g. $D(p) = \phi \cdot p$ for a constant ϕ , Corollary 1 shows that the tendency of oversensitive investors to buy stocks in response to a price increase varies with time. Early on, rational investors behave as sophisticated momentum investors, correctly inferring a positive innovation in fundamentals from recent price increases. In contrast, oversensitive investors are unaware of inattentive investors and do not buy the asset. As time passes and information about V is gradually revealed, rational investors begin to invest in a contrarian manner and sell shares to oversensitive investors who increasingly extrapolate from price increases and earn lower returns. Consequently, in the course of the price run-up, the holdings gap $\text{Gap}_{s,t} \equiv \frac{\chi D_t^\theta - (1-\chi) D_t^{rat}}{\chi D_t^\theta + (1-\chi) D_t^{rat}}$ initially tilts towards rational investors before tilting towards oversensitive investors.

Intuition: overreaction to public information To provide an intuition behind the dynamic behavior of oversensitive investors, note that the difference in how rational investors and oversensitive investors react to prices is proportional to the total amount of information contained in the price: the greater the public information, the greater the overreaction. For small t , when prices are not very informative of fundamentals, oversensitive investors react little to prices.¹⁷ It is only when prices gradually become highly informative of fundamentals that oversensitive investors overreact

¹⁷In particular, given that they assume prices are consistent with the equilibrium in which all investors are oversensitive, they underreact to price increases initially.

to price increases and chase past returns. We test these properties of oversensitive investors directly in Section 4.3.¹⁸

Combining the dynamics of the holdings gap and asset prices, Proposition 3 gives the key predictions of our full model. First, cross-sectionally, we replicate the predictions of Lemmas 1 and 2: asset price increases with higher θ_s have greater overreactive holdings gap and are likelier to revert. Second, we complement our results with new results on dynamics of asset prices and the holdings gap.

- Proposition 3.** *1. Dynamics: asset prices exhibit short-run momentum and long-run reversals. The average holdings gap $E[\text{Gap}_{s,t}|p, \theta_s]$ (in response to good news) tilts towards rational and then diagnostic investors: $E[\text{Gap}_{s,t}|p, \theta_s] < 2\chi - 1$ for t sufficiently small, and $E[\text{Gap}_{s,t}|p, \theta_s] > 2\chi - 1$ for t sufficiently large.*
- 2. Cross-section: holding t fixed, randomizing across V and θ_s , prices exhibit long-run reversals and greater tilt towards oversensitive investors if θ_s is sufficiently high, and long-run momentum with greater tilt towards rational investors if θ_s is sufficiently low.*

To summarize, Figure 1 illustrates the core predictions of our model. All of the price curves in the first panel display the same initial price increase, yet display significant heterogeneity in final returns, shown in the dotted lines. Such variation is generated by differences in the degree of overreaction, captured by θ_s . As the second panel illustrates, by moving beyond the price path and incorporating the variation in investor holdings, one can predict future returns of price increases. Our model yields the following set of core predictions.

Prediction 1 (Holdings gap, momentum, and reversals). *Stocks with high holdings gap experience less short-run momentum and greater long-run reversals.*

Prediction 2 (Holdings gap as a conditional predictor). *Stocks with high holdings gap experience negative returns, conditional on a large price increase.*

¹⁸Our model's prediction that investor biases increase in time thus distinguishes our model from models of private information, where investor biases gradually dissipate as more investors learn about fundamentals. Section 4.2 also documents further empirical evidence that our holdings-based measure captures overreaction instead of investor private information.

Prediction 3 (Dynamics of investor holdings). *Oversensitive investors are late-stage price extrapolators, buying stocks that have gone up in prices, earning lower returns, and lagging rational inflow. In particular, in a price run-up, there is initial entry of rational investors, which is followed by the entry of oversensitive investors.*

In the remaining sections, we implement the holdings-gap methodology and take these predictions to the data. In Section 3, we construct our measure of investor news sensitivity, and show that it is a persistent investor characteristic. In Section 4, we build the overreactive holdings gap at the asset level and test the main predictions of the model. Finally, in Section 5, we analyze the dynamics of our measure around public announcements, and highlight the ability of our measure to aggregate different underlying sources of overreaction.

3 The measurement and persistence of investor news sensitivity

The holdings gap approach predicts that if some investors systematically respond more to news than others, the gap in holdings between oversensitive and rational investors is an equilibrium measure of overreaction. Our model thus suggests the following two-step methodology to predict which short-run price increases revert in the long run. First, at the investor level, we need to construct a measure of news sensitivity, and confirm that it is persistent over time. Second, we need to aggregate the investor-level news sensitivity up to the stock level to construct the holdings gap, our key measure of overreaction in asset prices.

In this Section, we implement this methodology using quarterly institutional holdings data. We measure an investor’s news sensitivity (NS) based on how her holdings have responded to past public news, and show that it is highly persistent over time, especially among active investors. We then define the overreactive holdings gap in a stock by aggregating the news sensitivity of its investors weighed by their holdings. Given the holdings-based measure of overreaction, Section 4 then tests the core asset pricing predictions of our model.

3.1 Defining our main measure

Data Our sample of investors consists of all institutions that report their stock holdings through 13F filings from 1980 to 2020. The SEC requires that all large institutional investors report their

complete holdings of equities as of the last day in the quarter.¹⁹ Our data thus includes a wide variety of institutional investors, including hedge funds, pension funds, insurance companies, mutual funds, and banks. There are two important facets of this data. First, the data is aggregated to the firm level (e.g., Fidelity) rather than the individual fund level (e.g. the different funds operated by Fidelity). Second, although the direct holdings of retail investors (such as through brokerage accounts) are not included in the data, stocks that retail investors indirectly hold through institutions (such as through mutual funds or investment advisors) are. Because the institution must have discretion over the reported portfolio, we can interpret the allocation of a given institution’s funds across stocks as reflecting the institutional managers’ choices.²⁰ Our panel of institutional holdings allows us to observe how a given institution reacts to public information over a wide variety of stocks over a long time period. In particular, despite the exclusion of direct retail holdings, we find that even within the cross-section of large institutional investors, we detect sufficient differences in investor news sensitivity which allows us to measure overreaction at the asset level.

News sensitivity Using 13F filings, we uncover significant heterogeneity in the way that investor holdings react to public information. Given that we only have access to quarterly snapshots of holdings, we focus on how investors respond to relatively slow-moving innovations to fundamentals over horizons corresponding to roughly a year, the same horizon commonly used to form momentum portfolios (Carhart, 1997).

We measure an investor’s news sensitivity by the average news experienced by stocks bought by the investor. In quarter t for investor i , we define:

$$NS_{raw,i,t} = \frac{\sum N_{s,t} \cdot W_{i,s,t}}{\sum W_{i,s,t}}, \quad (21)$$

where $N_{s,t}$ is the aggregate “news” experienced by stock s from quarters $t - 3$ to t , and $W_{i,s,t} \geq 0$ is the amount purchased, in dollars, of stock s by investor i in quarter t . We measure $N_{s,t}$, the news experienced by stock s , from either announcement returns (La Porta et al., 1997) or normalized

¹⁹Large institutional investors are precisely defined as those that “exercise investment discretion” over at least \$100 million in 13F-eligible securities.

²⁰For investment firms whose aggregate portfolio partially or entirely consists of stocks held in index funds, the allocation of funds across stocks does not entirely reflect the managers’ active choices. We partially address this by standardizing our measure across investors, which control for common purchases and sales of stocks across institutional investors.

earnings surprises (Bouchaud et al., 2019).

$$N_{s,t}^{ret} = \sum_{h=0}^3 \text{Announcement Returns}_{s,t-h}, \quad N_{s,t}^{eps} = \frac{1}{P_{s,t-4}} \sum_{h=0}^3 \text{EPS Surprise}_{s,t-h}. \quad (22)$$

Announcement Returns $_{s,t}$ is the returns of stock s over a 10-day window around its earnings announcement in quarter t .²¹ EPS Surprise $_{s,t}$ is the difference between the announced earnings-per-share of stock s and the mean IBES analyst forecast of the same quantity. We normalize the total EPS Surprise measure by $P_{s,t-4}$, the price of stock s in quarter $t-4$, which translates the cumulative surprise into earnings yield space, as in Bouchaud et al. (2019).

For our main specification, we use $N_{s,t}^{ret}$, announcement returns, as our measure of past news. $NS_{raw,i,t}$ then measures the average (dollars-weighted) announcement date returns of stocks purchased by investor i in quarter t . For example, consider two investors, i and j . Suppose in 2020Q4 i buys 100 dollars worth each of stocks A and B , while j only buys 200 dollars worth of stock A . Suppose that in 2020, stock A rose by 10% on each of its 4 announcements, with net announcement returns of 40%, while B did not experience any surprise, with net announcement returns of 0. Then, j 's news sensitivity is 40%, which is higher than that of i , which is $40 \times 0.5 + 0 \times 0.5 = 20\%$.

Processing the raw measure Given the raw investor-by-quarter measure, $NS_{raw,i,t}$, we construct our main news sensitivity in three steps. First, given that we want to measure a persistent characteristic of an investor, we take an 8-quarter moving average of the raw NS measure. Second, to ensure that an investor's news sensitivity is not influenced by recent returns, we lag our measure by 8 quarters.²² Our results are robust to the exact choice of the window of the moving average and the lag. Finally, given that our model identifies oversensitive investors based on their *relative* news sensitivity, we standardize our measure across investors each quarter by taking the percentile

²¹Precisely, we use returns on days $d-4$ to $d+5$ where d is the announcement date. For stocks that only make annual announcements, we only include the annual announcement that takes place from quarters $t-3$ to t . One can also use announcement returns benchmarked to market returns, which makes no qualitative difference.

²²One particular concern that can arise if one allows for recent returns to influence our measure is that our return predictability measure can be confounded by unconditional autocorrelation of industry or factor returns. For example, suppose that the IT sector has outperformed (and experienced positive news) for the past year. Then, if we do not exclude recent data, tech investors would be categorized as having high news sensitivity, even if they may not be necessarily oversensitive to news. On the other hand, if there are industry-level reversals, this means that tech stocks are expected to underperform in the future. Then, the negative return predictability of NS may reflect the momentum and reversals of tech stocks, with no additional predictability of investor composition. There may be similar issues also for factor returns (Barberis and Shleifer, 2003; Ehsani and Linnainmaa, 2022).

rank. Equation (23) summarizes our procedure:

$$NS_{i,t} = \text{Percentile Rank} \left[\frac{1}{8} \sum_{a=0}^7 NS_{raw,i,t-8-a} \right]. \quad (23)$$

For example, the news sensitivity of an investor in 2020Q4 depends on her buys in quarters 2017Q1 to 2018Q4. For each quarter t , we compute the news-sensitivity of investors only for investors with non-missing holdings over the past 16 quarters. An investor who has bought stock with the highest announcement returns will have the highest news sensitivity, $NS_{i,t} = 1$.

3.2 Interpreting news sensitivity

Does high news sensitivity imply oversensitivity? Our measure of investor news sensitivity need not ex ante be a measure of *oversensitivity*. In particular, if there is sufficient underreaction in general following earnings announcements, the tendency to buy following positive earnings may indicate rational behavior, as given by the literature on post-earnings announcement drift (PEAD) (Bernard and Thomas, 1989). However, our measure does not align with sophisticated trading around PEAD: our news sensitivity is based on an investor’s purchase by the end of the quarter of the announcement, which tends to capture purchases occurring at a later timing than the horizon at which PEAD is profitable, which tends to be around 30 days. Furthermore, as documented in Appendix G, the tendency to buy based on announcements in the current quarter is highly correlated with the tendency to buy based on past announcements, again consistent with oversensitive investors overreacting to both current and lagged news. We provide further evidence of oversensitivity in Section 4.3, where we show that investors with high news sensitivity tend to underperform, with the stocks that they purchase earning lower abnormal returns going forward.

Timing of news sensitivity Equation (22) implies that our measure combines reaction to contemporaneous news (same quarter announcement returns) and lagged response to past news. Our combination of the two is motivated by Corollary 1, which shows that oversensitive investors both overreact to current news, especially at the later stage of the price increase. In Appendix G, we explore separating our news sensitivity to response to contemporaneous news and lagged news. Consistent with our model’s predictions, we find a high degree of correlation between these two

alternative measures, with our core predictions robust to either measures.

Focus on earnings announcements We construct our measure based on response to earnings announcements for three reasons. First, earnings announcements are the most standard and systematic way in which a company releases public information about its fundamentals. Second, relative to other events, one can also consider the earnings surprise as an alternative quantitative measure of the news without relying on prices. Third, earnings announcements, especially those accompanied by large surprises, are highly salient and attract investor attention. Consequently, earnings announcements form a natural setting to measure an investor’s tendency to overreact to salient public news. Despite only using earnings announcement returns, we critically show that our measure predicts an investor’s response to information beyond earnings announcement returns. Section 4 demonstrates this in general price increases and Section 5 directly shows that our measure captures reaction to important non-price information.

Focus on buys Our focus on an investor’s buys rather than the entire portfolio reflects our emphasis on measuring the active news-driven decision-making of the investor. First, by focusing on trading and changes in the portfolio, we ensure that an investor’s news sensitivity is not impacted by her passive holdings, which can reflect benchmarking unrelated to any active investment decisions (Antón et al., 2021).²³ This also implies that the persistence we find in news sensitivity is not driven by an investor holding a stock for multiple quarters. Second, our focus on buys rather than sells is motivated by recent work, such as Akepanidtaworn et al. (2021), which finds evidence that buys predominantly reflect the fundamental analysis and active stock-picking of investors, whereas sells may reflect liquidity requirements and other orthogonal factors.²⁴

3.3 Persistence of news sensitivity

Persistence We now confirm that news sensitivity is a persistent feature of an investor. Given that the news sensitivity of an investor i for quarter t aggregates the announcement returns of stocks

²³Our measure may also be influenced by the mechanical rebalancing of passive investors tracking an index. To address this, we standardize our measure each quarter to partial out common flows in the market. In Appendix C, we also directly control for investment style differences.

²⁴See Barber and Odean (2013) and Barber and Odean (2008) for analogous evidence for retail investors.

bought by i from quarters $t - 15$ to $t - 8$, we compute the persistence of NS over a horizon of 8 quarters, as indicated in Equation (24).

$$\text{NS}_{i,t+8} = \alpha + \rho \cdot \text{NS}_{i,t} + \varepsilon_{i,t}, \quad (24)$$

Our choice of the horizon of 8 quarters is to ensure that our persistence is not mechanically generated by overlapping windows.

Table 2 shows the estimates of Equation (24). Observations are at the investor-quarter level, with standard errors two-way clustered at the investor and quarter level. The first column shows an 8-quarter autoregression coefficient of 0.38, which implies that an investor in the top quintile of news sensitivity is on average in the top third ($0.32 + 0.38 \times 0.9 = 0.66$) after 2 years, indicating that our measure is highly persistent over time. Columns (2) through (4) of Table 2 repeat the analysis within each investor type: banks, investment firms, and pensions/insurers. Comparing across each institution type, we find that persistence of news sensitivity is notably stronger within mutual funds and hedge funds, consistent with the intuition that our measure reflects a persistent characteristic of active news-driven investment decisions. Finally, Table A1 compares the level of NS across each investor type: we find that investment firms (mutual funds and hedge funds) on average have the highest NS relative to other institutions, but the differences across institution types are small relative to the heterogeneity within each type.²⁵

Sources of persistence While our paper measures and documents the persistence of investor news sensitivity at the 13F institution level, further work is needed to investigate the sources of the persistence. First, at the individual investor level, there is growing evidence of systematic individual-level heterogeneity in overconfidence and tendency to overreact (Stango and Zinman, 2022), which can also apply to institutional investors (Greenwood and Nagel, 2009) even when aggregated up to the institution level. Another driver is investment style: Figure A1 shows that growth and momentum investors are more news-sensitive, which may be due to passive benchmarking or overreaction to factor returns (Barberis and Shleifer, 2003). Appendix C decomposes the persistence in news sensitivity of an investor into those attributable to an investment style and

²⁵For example, the NS of investment firms are on average 7% higher than that of banks.

the residual, and find that while style explains roughly a quarter of the persistence, the remaining three quarters is driven by investor decisions not attributable to an investment style. Another important driver is clientele, such as retail inflow (Frazzini and Lamont, 2008). For example, Guercio and Reuter (2014) find that mutual funds that are actively managed and predominantly broker-sold underperform. If 13F institutions differ in how much they cater to retail investors who chase positive news and returns, this may also generate systematic differences in news sensitivity.

3.4 Holdings gap measure of overreaction

Defining the holdings gap measure Equipped with a persistent investor-level measure of news sensitivity, we carry out the final step of our empirical strategy: we aggregate our investor measure to the asset level. We define our overreactive holdings gap of stock s in quarter t , $\text{Gap}_{s,t}^{emp}$, as the average holdings-weighted news sensitivity of investors of asset s in quarter t , as given by Equation (25).

$$\text{Gap}_{s,t}^{emp} = \frac{\sum_i H_{i,s,t} \cdot \text{NS}_{i,t}}{\sum_i H_{i,s,t}} \quad (25)$$

$H_{i,s,t}$ is the holdings of asset s by investor i in quarter t . $\text{Gap}_{s,t}^{emp}$ is the empirical counterpart to the theoretical holdings gap and is thus a holding-based measure of overreaction in asset s in quarter t . $\text{Gap}_{s,t}^{emp}$ ranges from 0 (if the stock is held entirely by the least news-sensitive investor) to 1 (if s is held entirely by the most news-sensitive investor), with an increase in $\text{Gap}_{s,t}^{emp}$ reflecting an influx of oversensitive investors.

Alternative measures We also consider alternative measures of the holdings gap, where we focus on the characteristics of recent investors. This allows us to correct for stale investors who are passively holding the asset without responding to any news. We do so in two ways: first, we define $\text{Gap}_{s,t}^{buy}$ as the NS of investors buying the asset in quarter t :

$$\text{Gap}_{s,t}^{buy} = \frac{\sum_i B_{i,s,t} \cdot \text{NS}_{i,t}}{\sum_i B_{i,s,t}}, \quad (26)$$

where $B_{i,s,t} \geq 0$ is the shares, in dollars, of s purchased by investor i in quarter t . We also define $\Delta \text{Gap}_{s,t}^{emp} = \text{Gap}_{s,t} - \text{Gap}_{s,t-4}$ as the change in the holdings gap over the past year. In practice, both measures are highly correlated and yield similar return predictability results to our main measure.

4 Holdings gap and return predictability

In this Section, we test the ability of the overreactive holdings gap to predict the returns following a price increase. Our model generates three core predictions. First, in Section 4.1, we test Prediction 1, which states that a higher holdings gap implies less short-run momentum and greater long-run reversals. We find that while stocks unconditionally exhibit short-run momentum, the degree of momentum is significantly reduced for stocks with a high holdings gap. In the long-run, we also find that stocks with high holdings gap display strong reversals: our measure is thus able to predict which short-run price run-ups revert in the long run.

In Section 4.2, we test Prediction 2: we find that the high holdings gap stocks most strongly underperform following large price increases. By contrast, an alternative investor holdings measure which proxies for private information predicts returns more unconditionally. Together, we find evidence consistent with our holdings gap measure being able to isolate overreaction from private information. Finally, we test Prediction 3 in Section 4.3, which states that oversensitive investors are late-stage price extrapolators, buying stocks that have gone up later than other institutional investors and thus earning lower risk-adjusted returns. This implies that in the course of a run-up episode, the holdings gap initially tilts towards rational investors before tilting towards oversensitive investors. We test both implications in the data and find significant support.

4.1 Prediction 1: momentum, reversals, and the holdings gap

We start with the regression specification in Equation (27), the exact empirical counterpart to Equation (12), which relates the return autocorrelation of a stock to its holdings gap.

$$r_{s,t+1,t+4} = \mu_t + \beta \cdot r_{s,past,t} + \gamma \cdot \text{Gap}_{s,t}^{std} \times r_{s,past,t} + \vartheta \cdot \text{Gap}_{s,t}^{std} + \delta \cdot X_{s,t} + \varepsilon_{s,t} \quad (27)$$

Observations are at the stock by quarter level. $r_{s,past,t}$ is the annualized 11-month net return from of stock s excluding the most recent month,²⁶ and $r_{s,t+1,t+4}$ is its subsequent 12-month returns. $\text{Gap}_{s,t}^{std}$ is the holdings gap standardized across all stocks in quarter t : holding fixed quarter t ,

²⁶This is following Jegadeesh and Titman (1993) and Carhart (1997), which form momentum portfolios excluding the most recent 1 month of data. We calculate the average monthly returns $r_{s,past,t}$ for stocks with at least 6 months of non-missing data leading up to quarter t .

$\text{Gap}_{s,t}^{std} = 1$ for a stock with the highest $\text{Gap}_{s,t}^{emp}$ and $\text{Gap}_{s,t+4}^{std} = 0$ for the stock with the lowest $\text{Gap}_{s,t}^{emp}$. $X_{s,t}$ are stock-level controls, and μ_t are quarter-fixed effects. The momentum/reversal coefficient conditional on the holdings gap is given by $\beta + \gamma \cdot \text{Gap}_{s,t}^{std}$, with the stock exhibiting momentum if $\beta + \gamma \cdot \text{Gap}_{s,t}^{std} > 0$, and reversals otherwise. Equation (12) implies that β is positive and γ is negative: while there is momentum for stocks with the lowest holdings gap, there should be less momentum, if not reversals, for stocks with higher holdings gap.

Table 3 reports the results corresponding to Equation (27). Column (1) computes the baseline return autocorrelation, and finds that stocks unconditionally exhibit momentum over a 4-quarter horizon: roughly 4% of a stock’s average returns over the past year are expected to further accrue over the next 12 months. Column (2) shows the main specification in Equation (27). Consistent with Prediction 1, we find that $\gamma < 0$: momentum is significantly weakened for stocks associated with higher overreactive holdings gap. The momentum coefficient for stocks with the lowest to the highest holdings gap ranges from 0.07 to $0.07 - 0.05 = 0.02$: the most overreactive stock only has a third of the momentum of the least overreactive stock. Column (3) shows that our results are further strengthened when one includes past 3-year returns in the stock-level controls $X_{s,t}$, which may contain additional information about the overreaction in the asset (De Bondt and Thaler, 1985).²⁷ Lastly, Column (4) also shows that our results are stronger when we refine our holdings gap measure to reflect the news sensitivity of recent investors in the asset, given by the average news sensitivity of investor inflow over the past year: $\frac{1}{4} \sum_{h=0}^3 \text{Gap}_{s,t-h}^{buy}$.²⁸

Overreactive and non-overreactive momentum To assess whether our measure delivers excess returns beyond standard factors, we construct two momentum portfolios based on the holdings gap and compute their abnormal returns. For each month, we sort stocks into deciles past returns, $r_{s,past,t}^{dec}$,²⁹ and quintiles of the holdings gap measure in the most recent quarter: $\text{Gap}_{s,t}^{quint}$. We then construct two momentum portfolios, “non-overreactive” and “overreactive” momentum, in the following manner. For “overreactive” momentum, one would expect a high level of holdings

²⁷Using 5 year past returns instead of 3 year past returns yields an analogous result.

²⁸One potential reason for doing so is to control for the component of the holdings gap driven by stale and inactive holdings of investors. One can also similarly replace the holdings gap with the one year change in the measure, which produces analogous results.

²⁹Average past returns is defined to be the average 12 month returns excluding the recent 1 month return, following (Carhart, 1997).

gap for winners (oversensitive investors buy more), and conversely a lower level of holdings gap for losers (oversensitive investors sell more). Conversely, one would expect the opposite for “non-overreactive” momentum, with a low holdings gap for winners and high holdings gap for losers. Thus, we construct the overreactive momentum portfolio by buying stocks in the 10th decile of returns and 5th quintile of the holdings gap and shorting stocks in the 1st decile of returns and 1st quintile of the holdings gap. Similarly, we construct the non-overreactive momentum portfolio by buying stocks in the 10th decile of returns and 1st quintile of the holdings gap and shorting stocks in the 1st decile of returns and 5th quintile of the holdings gap.

Table 4 compares the excess returns of our non-overreactive and overreactive momentum portfolios over the subsequent 12 months. Columns (1) and (4) report the raw returns, and Columns (2), (3), (5), and (6) control for the market, SMB, and HML factors. In all specifications, we find that the outperformance of momentum strategies is concentrated in the non-overreactive momentum portfolio with an annualized alpha of 10 – 15%, while the overreactive momentum portfolio has no significant outperformance. Consistent with the intuition that non-overreactive momentum occurs for relatively undervalued stocks, the non-overreactive momentum portfolio has a positive loading on the value factor, while overreactive momentum has a negative loading.

Short-run momentum and long-run reversals: a reconnect By applying the same methodology as in Table 4 across multiple horizons, Figure 2 traces out the long-run cumulative excess returns of non-overreactive and overreactive momentum portfolios. Comparing the long-run performance of the two portfolios reveals a striking dichotomy. While both momentum portfolios earn excess returns in the short-run, overreactive momentum experiences sharp long-run reversals, with negative excess returns of up to -20% accruing over three years. By contrast, the cumulative returns of non-overreactive momentum are relatively more sustained. Thus, consistent with Prediction 1, our holdings gap measure of overreaction predicts which short-run price increase will revert in the long-run. While the connection between short-run momentum and long-run reversals is unconditionally weak (Fama and French, 2010; Greenwood et al., 2019), by focusing on stocks with a high degree of overreaction, we are able to strengthen this connection.

4.2 Prediction 2: holdings gap as a conditional predictor of returns

Another consequence of our measure predicting the degree of long-run reversals is that the holdings gap of a stock should be a *conditional* predictor of returns: Prediction 2 states that the underperformance of stocks with high holdings gap should be concentrated in large price run-ups. Intuitively, given that the holdings gap measures overreaction based on heterogeneous response to public information, the predictability of our measure should be greatest when there have recently been widespread public news and price increases regarding the asset. In contrast, a measure of private information, which captures informed investors trading against uninformed investors, will be more unconditionally predictive: the purchases of informed investors predict positive future returns, regardless of current price increases.³⁰ To test Prediction 2 and focus on the difference between overreaction and private information, we show that stocks with high holdings gap strongly underperform in run-up episodes, while much less so otherwise. We then contrast our overreaction measure with a measure of informed investors, which predicts returns much more unconditionally.

Stock-level run-ups We systematically collect episodes in which a stock has experienced at least 100% returns over the past 4 quarters.³¹ Overall, we identify 13,000 episodes in our sample, and restrict to 10,000 episodes with non-missing values of the holdings gap measure.³² Figure A2 shows the fraction of stocks experiencing a run-up episode each quarter over our sample period from 1980 to 2020: while the number tends to increase over periods of overall high returns, the long-run average remains stable over our sample period. Stocks in our episodes are generally smaller (\$1.4 bn market cap) and have similar level of institutional holdings as the CRSP sample (41%).

High holdings gap stocks underperform following run-ups Figure 3 plots the cumulative log returns of stocks in our run-up episodes, where a run-up episode e begins in quarter $t^e + 1$. Quarters $t^e + 1$ to $t^e + 4$ correspond to the run-up period, where the stock experiences greater than 100%

³⁰Examples of such an approach include Campbell et al. (2009), which shows that institutional trades are predictive of earnings surprises. Similarly, Rapach et al. (2016) finds that short interest is predictive of future cash-flows.

³¹Greenwood et al. (2019) defines an industry-wide bubble as those that have experienced $\geq 100\%$ returns over a 2 year period. Given that we focus on episodes at an individual stock level where such run-ups are relatively more common, we take a shorter time horizon.

³²To prevent overlap in our episodes, we enforce a minimum of 8 quarters between the episodes of a given stock.

returns. Following the run-up, the black curve tracks the mean 3 year cumulative returns of our episodes, from quarters $t^e + 5$ to $t^e + 16$. On average, our episodes lead to a total reversal of 7.8% over 3 years (2.6% annualized), with significant dispersion, ranging from continuation to reversals. To visualize the predictive power of our holdings gap measure, we sort each episode based on quintiles of the level of the holdings gap at the end of the run-up period, $\text{Gap}_{s,t^e+4}^{emp}$. The red curves correspond to episodes with high levels of the holdings gap and the blue curves correspond to episodes with low levels of the holdings gap.³³ Our measure has significant power to predict the long-run returns of our episodes: run-ups associated with the top quartile of overreactive holdings are associated with a reversal of roughly 17% over 3 years, double the magnitude of the unconditional average.

We regress r_{s,t^e+5,t^e+16} , the three-year future return of the stock (cumulative returns of quarters $t^e + 5$ to $t^e + 16$ in the graph) on Gap_{s,t^e+4} , as given by Equation (28).

$$r_{s,t^e+5,t^e+16} = \alpha + \beta \cdot \text{Gap}_{s,t^e+4}^{std_epi} + \mu_{t^e} + \gamma \cdot X_{s,t^e+4} + \varepsilon_{s,t^e}. \quad (28)$$

Observations are at the run-up episode level. $\text{Gap}_{s,t^e+4}^{std_epi}$ is $\text{Gap}_{s,t^e+4}^{emp}$ standardized at the episode level.³⁴ β measures the average difference in 3-year returns of run-ups with the highest holdings gap to the lowest holdings gap in a given quarter. To control for common time-varying returns, we include time fixed-effects μ_t . Table 5 shows the estimates of Equation (28). Standard errors are two-way clustered at the stock and quarter level. Column (1) shows the results of the baseline regression. Even after controlling for time-varying market returns, run-ups with the greatest overreactive holdings on average revert by 12% over the subsequent 3 years more than those with the lowest holdings gap.³⁵ Column (2) adds in the run-up period returns r_{s,t^e+1,t^e+4} as controls.³⁶ Column (3) translates our return predictability results into crash-risk: both run-up period returns and

³³Figure A3 shows how the 3-year returns and crash probabilities, defined as the likelihood of 3-year returns being lower than -40% , vary with $\text{Gap}_{s,t^e+4}^{emp}$.

³⁴In other words, for stock run-ups beginning in quarter t , $\text{Gap}_{s,t^e+4}^{std_epi} = 1$ for the run-up with the highest $\text{Gap}_{s,t^e+4}^{emp}$ and $\text{Gap}_{s,t^e+4}^{std_epi} = 0$ for run-ups with the lowest $\text{Gap}_{s,t^e+4}^{emp}$.

³⁵By conditioning on previous returns being 100%, the linear coefficient on the holdings gap in episodes is effectively the discrete version of the interaction term in Equations (27) and (12).

³⁶The attenuation of the coefficient on $\text{Gap}_{s,t^e+4}^{std}$ is consistent with the fact that episodes with higher run-up period returns are associated with higher overreactive holdings and lower future returns. Regardless, the predictability of our measure on 3-year returns remains statistically and economically significant at 9%.

our overreactive holdings gap significantly increases the probability of extreme negative returns.³⁷ To summarize, our holdings-based measure of overreaction can explain a substantial fraction of the variation in the returns of stock-level run-ups, with greater overreactive holdings associated with greater long-run reversals and heightened crash risk.³⁸

Overreaction vs private information To highlight the difference between our measure of overreaction and measures of private information, we proxy the tendency of an investor to be privately informed of fundamentals by her average benchmarked returns, denoted as $P\&L_{i,t}$. We then use $P\&L_{i,t}$ to analogously construct a P&L-based holdings gap $\text{Gap}_{s,t}^{P\&L}$ for stock s in quarter t by taking the holdings-weighted average of $P\&L_{i,t}$ instead of $NS_{i,t}$. In Appendix E, we give details on how we orthogonalize our two measures to construct a portfolio that tracks overreaction, which we denote as the “purged NS” portfolio, and a portfolio that tracks private information, which we denote as the “purged P&L” portfolio. We then compare the alphas of these two portfolios, both restricted to stocks in run-up episodes and otherwise (non-episodes).³⁹

Consistent with Prediction 2, Table 6 shows that whereas the predictive power of the purged P&L portfolio is roughly the same across run-up episodes and non-episodes, the purged NS portfolio has virtually no predictive power in non-episodes, in contrast to its strong predictive power for run-up episodes. This is consistent with the notion that $\text{Gap}_{s,t}^{NS}$ measures overreaction based on differences in how investors react to public news. This analysis therefore demonstrates the unique ability of our investor composition measure to capture overreaction, which drives future reversals.

³⁷Appendix D confirms that our return predictability results hold even after controlling for the standard priced risk factors. An equal-weighted long-short portfolio trading stocks in run-up episodes based on their overreactive holdings gap has negative alpha, with stocks in the highest quintile of the holdings gap underperforming those in the lowest quintile at an annualized rate of roughly 8%.

³⁸Appendix H analyzes the symmetric case of extreme negative returns. Unlike the case of positive run-ups, our measure has limited predictive power conditional on stocks crashing. One can rationalize this asymmetry by accounting for short-sales constraints: intuitively, in response to good news, a greater degree of overreaction translates to higher levels of holdings gap, as oversensitive investors continue to buy the asset. On the other hand, in response to bad news, increasing overreaction may not change the holdings gap as oversensitive investors no longer hold any of the asset. Appendix H gives a sketch of the model extension and additional supportive evidence.

³⁹Precisely, we define a stock s in quarter t as an “episode” sample if the stock has had a run-up over the quarters $[t - 4, t - 1]$ (greater than 100% returns), and a “non-episode” sample otherwise.

4.3 Prediction 3: oversensitive investors are late-stage price extrapolators

Lastly, the dynamic version of our model predicts that oversensitive investors are late-stage price extrapolators: they buy stocks that have gone up in value later in the price increase when they are relatively more overvalued. This generates a systematic dynamic in investor composition in a run-up episode, where rational investors increase their holdings in the early stages before yielding the shares to oversensitive investors. We test these properties in two steps: first, we correlate the investor-level news sensitivity with three additional investor-level measures: the tendency to buy stocks that have gone up, the performance on these purchases, and the tendency to lag past institutional inflows. Second, we measure the average dynamics of the holdings gap in our run-up episodes, and confirm that oversensitive investors systematically enter later in the episodes.

Price extrapolation and buy performance We measure a investor’s price extrapolativeness and buy performance by measuring the average past 12-month returns of stocks bought by the investor and their subsequent 12-month benchmarked returns.⁴⁰

$$\text{Price Extrapolation}_{raw,i,t} = \frac{\sum M_{s,t} \cdot W_{i,s,t}}{\sum W_{i,s,t}}, \quad \text{Buy Performance}_{raw,i,t+4} = \frac{\sum_s R_{s,t,t+4}^{adj} \cdot W_{i,s,t}}{\sum_s W_{i,s,t}} \quad (29)$$

$W_{i,s,t}$ is the total purchase, in dollars, of stock s by investor i in quarter t . $M_{s,t}$ is the past 12-month return of stock s in quarter t , and $R_{s,t,t+4}^{adj}$ is the cumulative annual return of stock s from quarter $t + 1$ to $t + 4$, benchmarked to the average returns of stocks in the same book-to-market, size, and momentum quintile as s (Daniel et al., 1997).⁴¹ From the raw measures, we construct the final measures in the same way as news sensitivity, by taking the percentile rank of an 8-quarter moving average lagged by 8 quarters.

Late Finally, we measure the tendency of an investor to purchase stocks already bought by other institutional investors. We run a rolling 1-year regression for each investor of her trading in quarter

⁴⁰We follow Carhart (1997) in excluding the recent 1 month returns, although doing so makes no difference in our investor measure.

⁴¹We also test returns over alternative horizons of 2 and 5 years, as well as returns benchmarked instead to market returns, with similar results.

t of asset s by its lagged institutional inflows, as given in Equation 30.

$$\frac{\Delta H_{i,s,t'}}{\text{AUM}_{i,t'}} = \alpha_i + f_{raw,i,t} \cdot \Delta \text{Institutional Share}_{s,t'-1} + \varepsilon_{i,s,t'}, \quad t' \in \{t-3, t-2, t-1, t\} \quad (30)$$

$\Delta H_{i,s,t}$ is the change in holdings, in dollars, of stock s by investor i in quarter t , and $\text{AUM}_{i,t} = \sum_s H_{i,s,t}$ is the total value of investor i 's stock holdings in quarter t . $\Delta \text{Institutional Share}_{s,t-1}$ is the change in the fraction of institutional holdings of stock s from quarter $t-5$ to $t-1$. The higher the regression coefficient, $f_{raw,i,t}$, the stronger the tendency of investors to lag other institutional investors. We obtain our final measure by lagging $f_{raw,i,t}$ by 8 quarters and taking a percentile rank.

Oversensitive investors are late-stage price extrapolators We show that an investor's news sensitivity, which measures how investors buy stocks based on past news, is tightly correlated with our three measures in a way consistent with Prediction 3. Oversensitive investors are late-stage price extrapolators: while oversensitive investors tend to buy stocks that have gone up, they do so after other investors when the asset is already overvalued. Table 7 shows the relationship of news sensitivity to the three alternative investor measures. Observations are at the investor-by-quarter level for all investors in the 13F, with two-way clustered standard errors. Columns (1), (3), and (5) show the relationship of the three measures with our main news sensitivity, while columns (2), (4), and (6) show the same relationship using an alternative measure of news sensitivity computed based on earnings surprises, as defined in Equation (22). Columns (1) and (2) show that oversensitive investors have higher price extrapolativeness: investors who buy stocks that have experienced positive news also buy stocks that have risen in prices. Columns (3) and (4) show a negative relationship between an investor's news sensitivity and her buy performance, consistent with oversensitive investors buying stocks when they are relatively overvalued. Columns (5) and (6) further show a positive relationship between news sensitivity and the tendency to buy stocks already purchased by other investors: oversensitive investors are late investors.⁴²

⁴²The connection between investor overreactiveness and institutional flows relate to the literature on institutional herding (Sias (2004), Dasgupta et al. (2011)), which documents positive autocorrelation in institutional flows. In particular, Dasgupta et al. (2011) find that consecutive institutional inflows predict future reversals, which is consistent with the intuition that sustained institutional inflows are associated with overreaction. We contribute to that literature by further documenting the existence of institutions that systematically lag the inflows of other institutions, with oversensitive investors trailing other institutions in episodes of institutional inflows.

Dynamics of holdings gap The fact that oversensitive investors are late-stage price extrapolators generates non-monotonicity in the dynamics of the overreactive holdings gap: rational investors enter early, then leave as oversensitive investors enter late. Figure 4 plots the average dynamics of the holdings gap for 8 quarters leading up to the final quarter of the run-up episode (quarters -3 to 4). Consistent with our model’s predictions, the holdings gap decreases slightly prior to the run-up, and increases strongly during the run-up. Given that investor news sensitivity is based on asset prices at least lagged by 8 quarters, our results are not mechanically driven by contemporaneous asset returns during the run-up. Lastly, consistent with our return predictability results, run-ups that are associated with negative 3-year returns are associated with a persistently higher holdings gap through the run-up period than run-ups associated with positive 3-year returns.

News vs prices: news sensitivity beyond price extrapolation While our theory predicts that news sensitive investors endogenously extrapolate from price increases, it is important to distinguish oversensitive investors from mechanical price extrapolators or momentum traders. First, as we have demonstrated, oversensitive investors respond to prices more so at a later stage when it is unprofitable to do so. Second, and more importantly, the inflows of oversensitive investors respond to *news*, and not just price increases. The return predictability results in Sections 4.1 and 4.2 already demonstrate this indirectly: if the inflows of oversensitive investors are determined just by past returns, then stocks with the same run-ups will have no variation in the overreactive holdings gap. The fact that variations in our measure holding fixed the price path is able to predict future returns highlights its capacity to reflect overreaction driven by factors beyond the actual price increase. In Section 5, we show this more directly by studying how the holdings gap responds to news, and show in particular that our measure also aggregates various non-price information associated with overreaction.

5 Mechanism: holdings gap measures reaction to information

The return predictability analysis of Section 4 shows results that are consistent with our hypothesis that our holding-based measure sorts stocks based on the degree of overreaction. In this Section, we provide more direct evidence by analyzing how our measure responds to public news. Analogously to our findings in Section 4.3, we find that the public release of good news is preceded by rational

inflow and followed by oversensitive inflow. Holding fixed past returns, the degree of oversensitive inflow depends on non-price information commonly associated with overreaction, such as industry top performers (Shiller, 2015) and fast growth of fundamentals (Lakonishok et al., 1994). We conclude by showing that relative to any of these individual drivers of overreaction, our holdings gap has greater explanatory power over predicting post-announcement returns, thus aggregating many different channels of overreaction.

5.1 Holdings gap response to news

We analyze how our holdings gap responds to information by measuring its multi-quarter dynamics around positive earnings announcements.⁴³ Through the lens of our model, positive announcements correspond to a discrete jump in the amount of information processed by investors. Prediction 3 implies that positive announcements should on average be preceded by a decrease in the holdings gap, reflecting inflow from rational investors,⁴⁴ and then followed by an increase, or inflow from oversensitive investors. In the cross-section, the post-announcement increase in the holdings gap should reflect the degree of overreaction to the announcement.

Unconditional dynamics Starting 4 quarters prior to the announcement, we trace out the change in the holdings gap according to the following specification:

$$\Delta\text{Gap}_{s,t-5,t-5+h}^{emp} = \beta_h \cdot \text{positive}_{s,t} + \delta \cdot X_{s,t} + \phi_{t+h} + \epsilon_{s,t+h}, \quad (31)$$

for $1 \leq h \leq 9$. Observations are at the stock by quarter level. $\text{positive}_{s,t}$ is an indicator for stock s having a positive annual earnings announcement in quarter t , and $\Delta\text{Gap}_{s,t-4,t-4+h}^{emp}$ is the change in the holdings gap from quarter $t-5$ to quarter $t-4+h$. Each coefficient β_h then measures the average change in the holdings gap relative to its level in quarter $t-5$, both before ($h < 4$) and after ($h \geq 4$) the announcement. We control for unconditional autocorrelation in earnings announcement returns and the holdings gap by including lagged indicators $\text{positive}_{s,t-k}$ for $k \leq 4$ and $\Delta\text{Gap}_{t-8,t-4}$ as controls, along with quarter fixed effects. Figure 5 plots the response of β_h . Consistent with our

⁴³Appendix H shows the results for negative announcements.

⁴⁴In the model, this is driven by rational investors learning from prices more aggressively early on. In practice, this effect can also be driven by rational investors having greater access to information, or other related forces.

predictions, β_h is negative for $h < 4$ – the holdings gap weakly decreases before the announcement – and $\beta_h \geq 0$ for $h \geq 4$: there is inflow from oversensitive investors following the announcement that persists for multiple quarters. Table A4 presents a predictive version of our results, showing that a decrease in the holdings gap predicts positive news, which in turn predicts future changes in the holdings gap.

Holdings gap response to price and non-price information We now examine how the increase in the holdings gap following positive announcements depends on various drivers of investor overreaction. One concern is that holdings gap response may be entirely driven by the price increase around the announcement, which triggers inflows from oversensitive investors that extrapolate past returns. By controlling for the recent price path, we now show that the holdings gap also responds to important non-price drivers of overreaction.

We analyze two primary sources of non-price information associated with overreaction. First, we consider industry developments associated with the stock, in particular the past performance of winners within the same industry. Overextrapolation of industry developments is a salient feature of many narrative accounts of overreaction, such as the British railway mania of the 1840s, the conglomerate boom of the 1960s, the real estate investment trust boom of the 1970s (Soros, 2015), the tech bubble of the 1990s (Shiller, 2015), and the cryptocurrency boom of 2020s, where the success of early companies and assets in a given industry spurred future speculation in stocks in the same industry. To capture industry effects, we set $Z_{s,t,ind}$ as the excess returns of firms in the top decile of stock s 's industry, as measured by its three-digit SIC code, over the past 4 quarters.

Second, we also consider features of company fundamentals, and in particular the growth in its earnings and sales. For example, Bordalo et al. (2019) document evidence from analyst expectations that investors over-extrapolate fundamentals from companies whose earnings have grown rapidly in the past: overreaction to past fundamentals can thus provide a foundation for the underperformance of growth stocks (Lakonishok et al., 1994; La Porta, 1996; La Porta et al., 1997). Similarly, Bordalo et al. (2022c) and De La O and Myers (2021) show that overreactive expectations of stock fundamentals can explain a substantial fraction of asset price movements. To capture overreaction to fundamentals, we set $Z_{s,t,sales}$, $Z_{s,t,eps}$ as the growth in sales and earnings-per-share of company s over the past 4 quarters.

Context versus news In all of our examples, the non-price information we consider are developments that have occurred *prior* to the quarter t earnings announcement. In other words, the non-price information that we consider is not information released simultaneously with the announcement, but can be viewed as non-price features that shape the *context* in which investors interpret the positive earnings. Our choice of these variables reflects the fact that investor overreaction is shaped not just by immediately released information, but also how it interacts with the recruitment of past relevant information.⁴⁵

Holdings gap response to non-price information We examine whether our overreactive holdings gap measure reflects reaction to price and non-price information by analyzing how the over-sensitive inflow following the positive announcement varies with price and non-price variables. Holding fixed horizon h , we run the following regression:

$$\Delta \text{Gap}_{s,t-1,t+h}^{emp} = \beta_h \cdot \text{positive}_{s,t} + \alpha_h \cdot Z_{s,t} + \gamma_h \cdot Z_{s,t} \times \text{positive}_{s,t} + \delta \cdot X_{s,t} + \phi_{k,t+h} + \varepsilon_{s,t+h}. \quad (32)$$

Observations are at the stock by quarter level. $\Delta \text{Gap}_{s,t-1,t+h}^{emp}$ is the h -quarter change in the holdings gap after the announcement, and $Z_{s,t} \in \{Z_{s,t,past}, Z_{s,t,contemp}, Z_{s,t,ind}, Z_{s,t,sales}, Z_{s,t,eps}\}$ is either a price variable, which consists of the past 4 quarter returns ($Z_{s,t,past}$) and the announcement quarter returns ($Z_{s,t,contemp}$), or a non-price variable, which consists of industry winners ($Z_{s,t,ind}$), sales growth ($Z_{s,t,sales}$), and earnings growth ($Z_{s,t,eps}$). The interaction term γ_h measures how the post-announcement change in the holdings gap varies each of these variables. For non-price variables, we residualize out price dynamics and common industry movements by controlling for the past four quarter log returns of stock s and including industry-by-quarter fixed effects.

Table 8 shows the estimates of Equation (32) for $h = 1$: we analyze the the 2-quarter response (the difference in the holdings gap in quarters $t + 1$ relative to $t - 1$) in the holdings gap in response to the announcement in quarter t . Columns (1) and (2) show that the holdings gap increases in the past and contemporaneous returns, consistent with oversensitive investors extrapolating from price increases. Furthermore, Columns (3), (4), and (5) show that for all three measures of non-

⁴⁵For an example in which investor reaction varies with immediate announcement characteristics, [Kwon and Tang \(2020\)](#) explore how investor overreaction to an event is predicted by with its association with extreme realizations of fundamentals.

price information, the interaction term γ_h is significant and positive, even after controlling for the price path.⁴⁶ A one standard deviation increase in industry winner returns, sales growth, and EPS growth are associated with, respectively, 20%, 20%, and 10% higher increase in $\Delta\text{Gap}_{s,t-1,t+1}^{emp}$ relative to its average increase. Figure A4 plots γ_h for $h \geq 1$: each non-price information has a persistent multi-quarter effect on future oversensitive inflow following the announcement. To summarize, all of our measures of non-price information significantly predict greater oversensitive inflow, even after controlling for past returns. This indicates that the our holdings gap measure of overreaction can also account for the variation in observable non-price information that drives investor overreaction to news.

5.2 Explanatory power: holdings gap as an aggregator of overreaction

While our measure captures observable drivers of non-price overreaction, it should also reflect the total degree of overreaction, and thereby act as an aggregator of both observable and unobservable drivers of overreaction. In particular, this implies that the holdings gap measure should provide substantial incremental explanatory power in predicting returns beyond using just the price path, with its explanatory power larger than any individual non-price variable.

Aggregate explanatory power To test this, we measure how much our variables can account for the variation in future returns. Our sample continues to be the set of stocks that had a positive annual earnings announcement in quarter t . For each of these stocks, we predict $r_{s,t,t+2}$, the two-quarter returns of asset s following the quarter t announcement, given by Equation (33),

$$r_{s,t,t+2} = \gamma \cdot Z_{s,t} + \phi_{k,t} + \varepsilon_{s,t}, \quad (33)$$

where $\phi_{k,t}$ are industry-by-quarter fixed effects. We start by setting $Z_{s,t}$ as the recent price path, consisting of returns over quarters $t - 3$ to t , and then add in the Fama-French 4 factor loadings as of quarter t . From this baseline, we then measure the increase in the R^2 of Equation (33) when one adds the holdings gap over quarters $t - 3$ to t , and compare it to the increase when one separately

⁴⁶In our main specification, we do not control for announcement quarter returns, as it is endogenous in the degree of oversensitive inflow. Table A5 shows the results controlling for announcement quarter returns, which shows similar results.

adds each of our three non-price information variables. Lastly, we compute the R^2 of Equation (33) when we combine all three non-price information measures, as well as when one also adds in the holdings gap.⁴⁷

The results are summarized in Figure 6. Each regression includes the path of past returns in $Z_{s,t}$. The gray bar establishes the baseline R^2 when $Z_{s,t}$ only includes the path of past returns, as well as adding in the Fama-French loadings. The blue and purple bars reflect the additional explanatory power, beyond prices, of each of our non-price variables, with our holdings gap in blue and the other non-price characteristics in purple. The first orange bar shows the R^2 when $Z_{s,t}$ includes all three non-price characteristics, and the second orange bar shows the R^2 with all of our variables added in. While the explanatory power of each regression is low in absolute terms – stock-level returns within a quarter-industry category are extremely noisy and hard to predict – Figure 6 shows that adding in the path of the holdings gap leading up to the announcement increases by 20% the explanatory power just of using the price paths and the Fama French characteristics, highlighting the importance of non-price information as a determinant of overreaction. Furthermore, the explanatory power of the holdings gap is substantially higher than any one of the three non-price characteristics, as well as the regression that includes the three non-price measures simultaneously. In Appendix F, we also show that the outperformance of our holdings gap measure relative to the other individual non-price information is even stronger in out-sample predictions (Campbell and Thompson, 2008). In Appendix I, we compare the explanatory power of our variables in predicting long-run reversals of short-run momentum portfolios, and find additional predictive power of our holdings gap variable. Overall, we find evidence consistent with the ability of our holdings gap to aggregate overreaction to non-price information, both observed and unobserved.⁴⁸

6 Conclusion

A large price increase is not necessarily a symptom of overreaction: to predict whether a short-run rise in prices will revert, one needs a separate measure of overreaction. In this paper, we contribute

⁴⁷To ensure that these comparisons are apples-to-apples, we restrict to stocks that have non-missing values for each of the different stock-level characteristics and transform each characteristic to be the (normalized) quarterly rank within the set of included stocks.

⁴⁸The regression with all three non-price characteristics and our holdings gap has a higher explanatory power than just the regression with the holdings gap, reflecting that our measure is a noisy aggregator of non-price information.

by developing a methodology to measure overreaction based on investor holdings. We develop a model in which investors have heterogeneous sensitivity to news. In equilibrium, the relative holdings between rational and oversensitive investors then reveal the degree of overreaction and expected future returns of a price increase. Turning to the data, we validate the fundamental assumption behind our approach: there is persistence in an investor's news sensitivity, measured by her portfolio response to past news. We then aggregate an investor's news sensitivity to construct the overreactive holdings gap at the asset level, our holdings-based measure of overreaction. Consistent with our theory, our measure is able to predict which price increases revert in the long-run. By analyzing how the holdings gap responds to news, we further validate its ability to measure overreaction to observed and non-observed information regarding an asset.

Our work suggests that understanding the heterogeneity in investor information sensitivity – how different investors have responded to different information in the past – is a promising way to better understand the current drivers of market prices and trading activity. While we have focused on a single dimension of news sensitivity, the tendency to overreact to all public news, our methodology can extend to settings where investor sensitivity to information may vary in a more granular manner: for example, some investors may be more attentive to certain types of information similar to what they have experienced in the past. Such heterogeneity can be driven by deeper psychological forces, such as different investor experiences ([Malmendier, 2021](#)) and memory ([Bordalo et al., 2020a, 2022a](#); [Wachter and Kahana, 2019](#)). By better understanding investor sensitivity to information, one can thus gain a deeper understanding of the granular drivers of asset prices and investor beliefs.

References

- Afrouzi, Hassan, Spencer Yongwook Kwon, Augustin Landier, Yueran Ma, and David Thesmar**, “Overreaction in expectations: Evidence and theory,” *Available at SSRN*, 2020.
- Akepanidaworn, Klakow, Rick Di Mascio, Alex Imas, and Lawrence Schmidt**, “Selling Fast and Buying Slow: Heuristics and Trading Performance of Institutional Investors,” Technical Report, National Bureau of Economic Research 2021.
- Antón, Miguel, Randolph B. Cohen, and Christopher Polk**, “Best Ideas,” *Available at SSRN 1364827*, 2021.
- Asness, Clifford S., Tobias J. Moskowitz, and Lasse Heje Pedersen**, “Value and Momentum Everywhere,” *The Journal of Finance*, 2013, 68 (3), 929–985.
- Barber, Brad M. and Terrance Odean**, “All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors,” *The Review of Financial Studies*, 2008, 21 (2), 785–818.
- and —, “The Behavior of Individual Investors,” *Handbook of the Economics of Finance*, 2013, 2, 1533–1570.
- Barberis, Nicholas and Andrei Shleifer**, “Style Investing,” *Journal of Financial Economics*, 2003, 68 (2), 161–199.
- , —, and **Robert Vishny**, “A Model of Investor Sentiment,” *Journal of Financial Economics*, 1998, 49 (3), 307–343.
- , **Robin Greenwood, Lawrence Jin, and Andrei Shleifer**, “Extrapolation and Bubbles,” *Journal of Financial Economics*, 2018, 129 (2), 203–227.
- Bernard, Victor L and Jacob K Thomas**, “Post-earnings-announcement drift: delayed price response or risk premium?,” *Journal of Accounting research*, 1989, 27, 1–36.
- Bollen, Nicolas P. B. and Jeffrey A. Busse**, “Short-Term Persistence in Mutual Fund Performance,” *The Review of Financial Studies*, 08 2004, 18 (2), 569–597.
- Bondt, Werner F. M. De and Richard Thaler**, “Does the Stock Market Overreact?,” *The Journal of Finance*, 1985, 40 (3), 793–805.
- Bordalo, Pedro, John J. Conlon, Nicola Gennaioli, Spencer Yongwook Kwon, and Andrei Shleifer**, “Memory and Probability,” *The Quarterly Journal of Economics*, 08 2022.
- , **Nicola Gennaioli, and Andrei Shleifer**, “Diagnostic Expectations and Credit Cycles,” *The Journal of Finance*, 2018, 73 (1), 199–227.
- , —, and —, “Memory, Attention, and Choice,” *The Quarterly Journal of Economics*, 2020, 135 (3), 1399–1442.
- , —, and —, “Salience,” *Annual Review of Economics*, 2022, 14, 521–544.

- , – , **Rafael La Porta, and Andrei Shleifer**, “Diagnostic Expectations and Stock Returns,” *The Journal of Finance*, 2019, 74 (6), 2839–2874.
- , – , **Rafael LaPorta, and Andrei Shleifer**, “Belief Overreaction and Stock Market Puzzles,” Technical Report, Working paper 2022.
- , – , **Spencer Yongwook Kwon, and Andrei Shleifer**, “Diagnostic Bubbles,” *Journal of Financial Economics*, 2021, 141 (3), 1060–1077.
- , – , **Yueran Ma, and Andrei Shleifer**, “Overreaction in Macroeconomic Expectations,” *American Economic Review*, 2020, 110 (9), 2748–82.
- Bose, Devdepta, Henning Cordes, Sven Nolte, Judith Christiane Schneider, and Colin Farrell Camerer**, “Decision weights for experimental asset prices based on visual salience,” *The Review of Financial Studies*, 2022, 35 (11), 5094–5126.
- Bouchaud, Jean-Philippe, Philipp Krueger, Augustin Landier, and David Thesmar**, “Sticky Expectations and the Profitability Anomaly,” *The Journal of Finance*, 2019, 74 (2), 639–674.
- Brooks, John**, *The Go-Go Years: The Drama and Crashing Finale of Wall Street’s Bullish 60s*, John Wiley & Sons, 1999.
- Campbell, John Y and Samuel B Thompson**, “Predicting excess stock returns out of sample: Can anything beat the historical average?,” *The Review of Financial Studies*, 2008, 21 (4), 1509–1531.
- Campbell, John Y., Tarun Ramadorai, and Allie Schwartz**, “Caught on Tape: Institutional Trading, Stock Returns, and Earnings Announcements,” *Journal of Financial Economics*, 2009, 92 (1), 66–91.
- Carhart, Mark M.**, “On Persistence in Mutual Fund Performance,” *The Journal of Finance*, 1997, 52 (1), 57–82.
- Cassella, Stefano and Huseyin Gulen**, “Extrapolation Bias and the Predictability of Stock Returns by Price-Scaled Variables,” *The Review of Financial Studies*, 02 2018, 31 (11), 4345–4397.
- Cella, Cristina, Andrew Ellul, and Mariassunta Giannetti**, “Investors’ horizons and the amplification of market shocks,” *The Review of Financial Studies*, 1993, 26 (7).
- Chernenko, Sergey, Samuel G. Hanson, and Adi Sunderam**, “Who Neglects Risk? Investor Experience and the Credit Boom,” *Journal of Financial Economics*, 2016, 122 (2), 248–269.
- Coppola, Antonio**, “In Safe hands: The financial and real impact of investor composition over the credit cycle,” *Working paper*, 2022.
- Cutler, David M, James M Poterba, and Lawrence H Summers**, “What moves stock prices?,” 1988.
- Daniel, Kent, Alexander Klos, and Simon Rottke**, “The Dynamics of Disagreement,” *The Review of Financial Studies*, 10 2022.

- **and Sheridan Titman**, “Market Reactions to Tangible and Intangible Information,” *The Journal of Finance*, 2006, 61 (4), 1605–1643.
- **and Tobias J. Moskowitz**, “Momentum Crashes,” *Journal of Financial Economics*, 2016, 122 (2), 221–247.
- **, David Hirshleifer, and Avanidhar Subrahmanyam**, “Investor Psychology and Security Market Under- and Overreactions,” *The Journal of Finance*, 1998, 53 (6), 1839–1885.
- **, Mark Grinblatt, Sheridan Titman, and Russ Wermers**, “Measuring Mutual Fund Performance with Characteristic-Based Benchmarks,” *The Journal of Finance*, 1997, 52 (3), 1035–1058.
- Dasgupta, Amil, Andrea Prat, and Michela Verardo**, “Institutional Trade Persistence and Long-Term Equity Returns,” *The Journal of Finance*, 2011, 66 (2), 635–653.
- Della Vigna, Stefano and Joshua M. Pollet**, “Investor Inattention and Friday Earnings Announcements,” *The Journal of Finance*, 2009, 64 (2), 709–749.
- Diamond, Douglas W. and Robert E. Verrecchia**, “Information Aggregation in a Noisy Rational Expectations Economy,” *Journal of Financial Economics*, 1981, 9 (3), 221–235.
- Ehsani, Sina and Juhani T. Linnainmaa**, “Factor Momentum and the Momentum Factor,” *The Journal of Finance*, 2022, 77 (3), 1877–1919.
- Eyster, Erik and Matthew Rabin**, “Naive Herding in Rich-Information Settings,” *American Economic Journal: Microeconomics*, 2010, 2 (4), 221–43.
- Fama, Eugene F.**, “Two Pillars of Asset Pricing,” *American Economic Review*, 2014, 104 (6), 1467–85.
- **and Kenneth R. French**, “Common Risk Factors in the Returns on Stocks and Bonds,” *Journal of Financial Economics*, 1993, 33 (1), 3–56.
- **and –**, “Luck versus Skill in the Cross-Section of Mutual Fund Returns,” *The Journal of Finance*, 2010, 65 (5), 1915–1947.
- Frazzini, Andrea and Owen A Lamont**, “Dumb Money: Mutual Fund Flows and the Cross-Section of Stock Returns,” *Journal of Financial Economics*, 2008, 88 (2), 299–322.
- Fung, William, David A. Hsieh, Narayan Y. Naik, and Tarun Ramadorai**, “Hedge Funds: Performance, Risk, and Capital Formation,” *The Journal of Finance*, 2008, 63 (4), 1777–1803.
- Gagnon-Bartsch, Tristan and Benjamin Bushong**, “Learning with misattribution of reference dependence,” *Journal of Economic Theory*, 2022, 203, 105473.
- Gerakos, Joseph, Juhani Linnainmaa, and Adair Morse**, “Asset Managers: Institutional Performance and Factor Exposures,” *The Journal of Finance*, 2021, 76 (4), 2035–2075.

- Greenwood, Robin and Andrei Shleifer**, “Expectations of returns and expected returns,” *The Review of Financial Studies*, 2014, 27 (3), 714–746.
- **and Stefan Nagel**, “Inexperienced Investors and Bubbles,” *Journal of Financial Economics*, 2009, 93 (2), 239–258.
- , **Andrei Shleifer, and Yang You**, “Bubbles for Fama,” *Journal of Financial Economics*, 2019, 131 (1), 20–43.
- Griffin, John M. and Jin Xu**, “How Smart Are the Smart Guys? A Unique View from Hedge Fund Stock Holdings,” *The Review of Financial Studies*, 05 2009, 22 (7), 2531–2570.
- Grinblatt, Mark and Sheridan Titman**, “The Persistence of Mutual Fund Performance,” *The Journal of Finance*, 1992, 47 (5), 1977–1984.
- Grossman, Sanford J. and Joseph E. Stiglitz**, “On the Impossibility of Informationally Efficient Markets,” *The American Economic Review*, 1980, 70 (3), 393–408.
- Guercio, Diane Del and Jonathan Reuter**, “Mutual fund performance and the incentive to generate alpha,” *The Journal of Finance*, 2014, 69 (4), 1673–1704.
- Hellwig, Martin F.**, “On the Aggregation of Information in Competitive Markets,” *Journal of Economic Theory*, 1980, 22 (3), 477–498.
- Hong, Harrison and Jeremy C. Stein**, “A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets,” *The Journal of Finance*, 1999, 54 (6), 2143–2184.
- **and** – , “Disagreement and the Stock Market,” *Journal of Economic Perspectives*, 2007, 21 (2), 109–128.
- , **Terence Lim, and Jeremy C. Stein**, “Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies,” *The Journal of Finance*, 2000, 55 (1), 265–295.
- Jagannathan, Ravi, Alexey Malakhov, and Dmitry Novikov**, “Do Hot Hands Exist among Hedge Fund Managers? An Empirical Evaluation,” *The Journal of Finance*, 2010, 65 (1), 217–255.
- Jegadeesh, Narasimhan and Sheridan Titman**, “Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency,” *The Journal of Finance*, 1993, 48 (1), 65–91.
- Kindleberger, Charles P.**, *Manias, Panics, and Crashes: A History of Financial Crises* 1978.
- Koijen, Ralph SJ and Motohiro Yogo**, “A demand system approach to asset pricing,” *Journal of Political Economy*, 2019, 127 (4), 1475–1515.
- Kosowski, Robert, Allan Timmermann, Russ Wermers, and Hal White**, “Can Mutual Fund ”Stars” Really Pick Stocks? New Evidence from a Bootstrap Analysis,” *The Journal of Finance*, 2006, 61 (6), 2551–2595.

- Kwon, Spencer Yongwook and Johnny Tang**, “Extreme Events and Overreaction to News,” Available at SSRN 3724420, 2020.
- Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny**, “Contrarian Investment, Extrapolation, and Risk,” *The Journal of Finance*, 1994, 49 (5), 1541–1578.
- Lamont, Owen A. and Jeremy C. Stein**, “Investor Sentiment and Corporate Finance: Micro and Macro,” *American Economic Review*, 2006, 96 (2), 147–151.
- Lee, Charles M. C. and Bhaskaran Swaminathan**, “Price Momentum and Trading Volume,” *The Journal of Finance*, 2000, 55 (5), 2017–2069.
- Lewellen, Jonathan**, “Institutional Investors and the Limits of Arbitrage,” *Journal of Financial Economics*, 2011, 102 (1), 62–80.
- Li, Xiaomin and Colin F Camerer**, “Predictable effects of visual salience in experimental decisions and games,” *The Quarterly Journal of Economics*, 2022, 137 (3), 1849–1900.
- Lou, Dong**, “A Flow-Based Explanation for Return Predictability,” *The Review of Financial Studies*, 12 2012, 25 (12), 3457–3489.
- **and Christopher Polk**, “Comomentum: Inferring arbitrage activity from return correlations,” *The Review of Financial Studies*, 2022, 35 (7), 3272–3302.
- Ma, Yueran**, “Nonfinancial Firms as Cross-Market Arbitrageurs,” *The Journal of Finance*, 2019, 74 (6), 3041–3087.
- Mackay, Charles**, *Extraordinary Popular Delusions and the Madness of Crowds* 1841.
- Malmendier, Ulrike**, “Experience Effects in Finance: Foundations, Applications, and Future Directions,” *Review of Finance*, 2021, 25 (5), 1339–1363.
- Nagel, Stefan**, “Short sales, institutional investors and the cross-section of stock returns,” *Journal of Financial Economics*, 2005, 78 (2), 277–309.
- O, Ricardo De La and Sean Myers**, “Subjective Cash Flow and Discount Rate Expectations,” *The Journal of Finance*, 2021, 76 (3), 1339–1387.
- Pontiff, Jeffrey and Artemiza Woodgate**, “Share issuance and cross-sectional returns,” *The Journal of Finance*, 2008, 63 (2), 921–945.
- Porta, Rafael La**, “Expectations and the Cross-Section of Stock Returns,” *The Journal of Finance*, 1996, 51 (5), 1715–1742.
- , **Josef Lakonishok, Andrei Shleifer, and Robert Vishny**, “Good News for Value Stocks: Further Evidence on Market Efficiency,” *The Journal of Finance*, 1997, 52 (2), 859–874.
- Rabin, Matthew and Dimitri Vayanos**, “The Gambler’s and Hot-Hand Fallacies: Theory and Applications,” *The Review of Economic Studies*, 2010, 77 (2), 730–778.

- Rapach, David E., Matthew C. Ringgenberg, and Guofu Zhou**, “Short Interest and Aggregate Stock Returns,” *Journal of Financial Economics*, 2016, 121 (1), 46–65.
- Scheinkman, José A. and Wei Xiong**, “Overconfidence and Speculative Bubbles,” *Journal of Political Economy*, 2003, 111 (6), 1183–1220.
- Shiller, Robert J.**, *Irrational Exuberance*, Princeton University Press, 2015.
- , “Narrative Economics,” *American Economic Review*, 2017, 107 (4), 967–1004.
- Sias, Richard W.**, “Institutional Herding,” *The Review of Financial Studies*, 2004, 17 (1), 165–206.
- Soros, George**, *The Alchemy of Finance*, John Wiley & Sons, 2015.
- Stango, Victor and Jonathan Zinman**, “We Are All Behavioural, More, or Less: A Taxonomy of Consumer Decision-Making,” *The Review of Economic Studies*, 2022.
- Tetlock, Paul C.**, “Giving Content to Investor Sentiment: The Role of Media in the Stock Market,” *The Journal of Finance*, 2007, 62 (3), 1139–1168.
- Tversky, Amos and Daniel Kahneman**, “Availability: A Heuristic for Judging Frequency and Probability,” *Cognitive Psychology*, 1973, 5 (2), 207–232.
- Wachter, Jessica A and Michael Jacob Kahana**, “A retrieved-context theory of financial decisions,” Technical Report, National Bureau of Economic Research 2019.

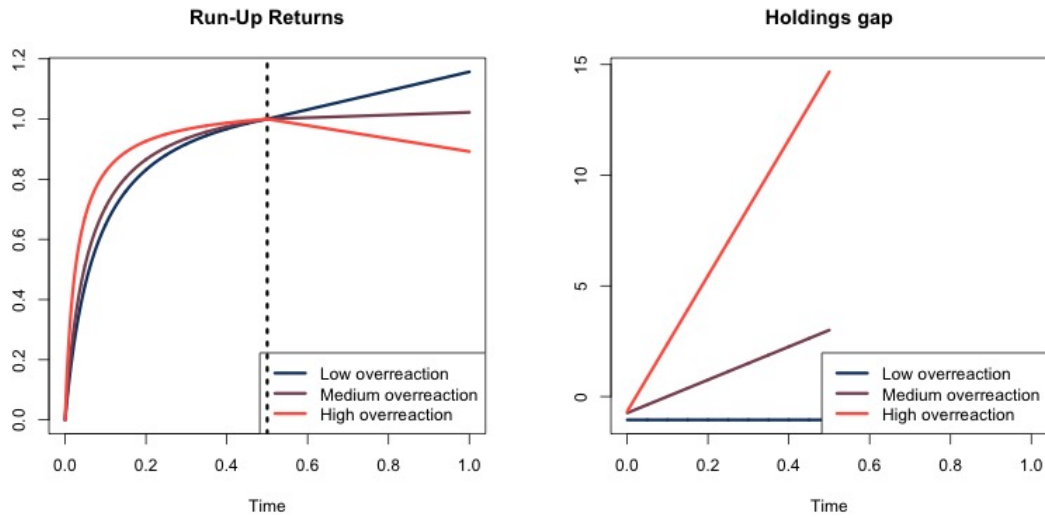


Figure 1: Final model prediction: diagnostic equilibrium

Notes: The left subplot is the average price path for $\theta_s = 0, 0.5, 2$, corresponding to low, medium, and high overreaction. The fundamentals V corresponding to each three scenarios are normalized to have the three average price paths evaluate to the same value at $t = 0.5$. The right subplot is the average holdings gap associated with the three price paths. Put together, the contemporaneous level of the holdings gap forecasts future momentum or reversals.

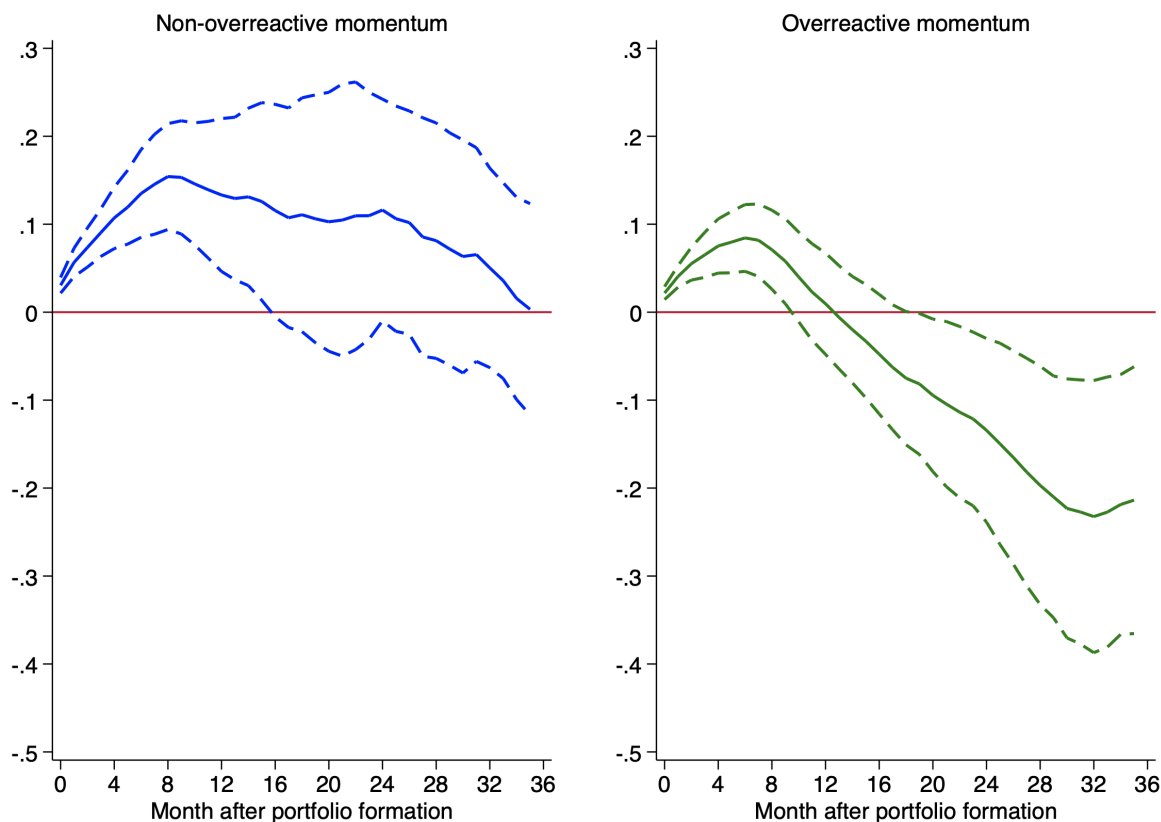


Figure 2: Cumulative momentum returns, sorted by overreactive holdings gap

Notes: Figure 2 plots the cumulative abnormal returns of two momentum portfolios, which are formed by double sorting stocks in each month t into (a) the decile of cumulative log returns over the past 12 months excluding the most recent month $t - 1$ (Carhart, 1997) and (b) the quintile of the holdings gap measure in the quarter before month t 's quarter, which is defined following Equation (25). Observations are at the monthly level. In the left panel, we plot the cumulative abnormal returns of the non-overreactive momentum portfolio, which goes long stocks in the 10th decile of returns and 1st quintile of the holdings gap and shorts stocks in the 1st decile of returns and 5th quintile of the holdings gap. In the right panel, we plot the cumulative abnormal returns of the overreactive momentum portfolio, which goes long stocks in the 10th decile of returns and 5th quintile of the holdings gap and shorts stocks in the 1st decile of returns and 1st quintile of the holdings gap. We compute abnormal returns by regressing, for each horizon, the unweighted mean returns of each sorted portfolio on the returns of the market portfolio (the CRSP value-weighted portfolio), SMB (size), and HML (value) (Fama and French, 1993) over the same horizon. The dotted lines show the 95% confidence intervals, where the standard errors are estimated using the Newey-West method with a lag of twelve.

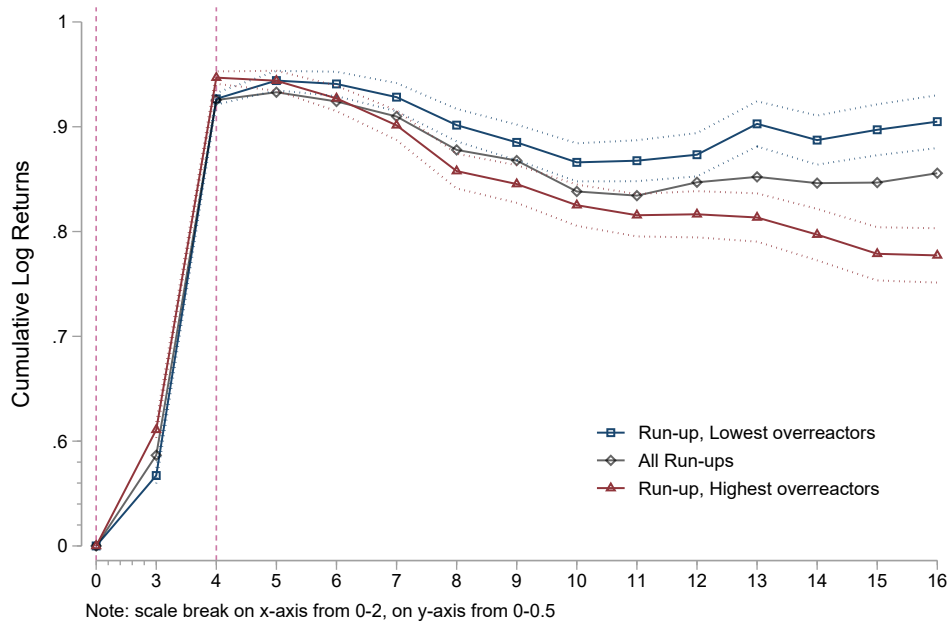


Figure 3: Return predictability: stock-level run-ups

Notes: Figure 3 plots the cumulative log returns of the stock-level run-ups, which is defined by a stock experiencing at least 100% returns over 4 quarters, which correspond to quarters 1 through 4 in the figure. Quarters 5 through 16 correspond to the subsequent 3-year returns. We winsorize our episodes by their holdings gap in quarter 4 at the 10% level. The black curve plots the average across all episodes, while the blue and red curves correspond to the bottom and top quintiles of the holdings gap at the end of the selection period (quarter 4). The dotted lines represent one standard-error intervals around the cumulative returns of high and low holdings gap episodes.

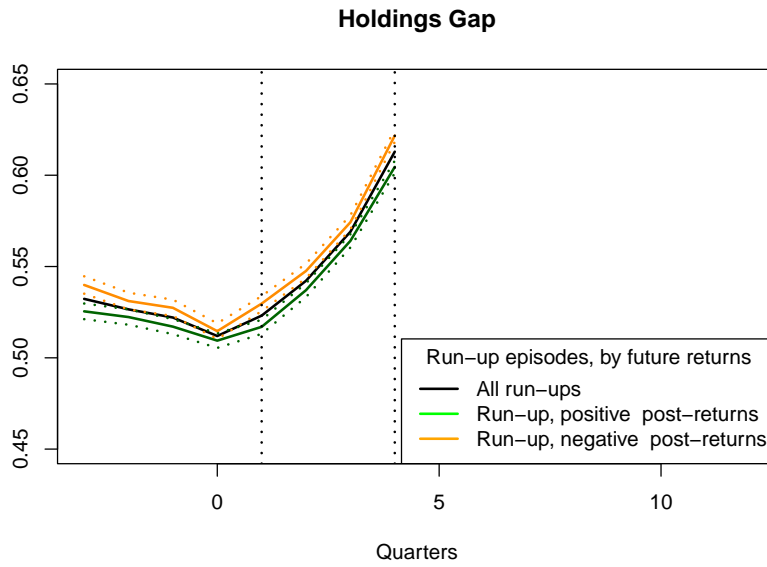


Figure 4: Dynamics of holdings gap: stock-level run-ups

Notes: Figure 4 shows the standardized holdings gap of a stock in the run-up episodes, which is defined to be a stock experiencing at least 100% returns over 4 quarters. The standardized holdings gap is given by the percentile rank across each quarter of the holdings-gap measure of the stock, as defined in Equation (25). Quarters -3 to 0 correspond to 4 quarters prior to the run-up, and quarters 1 through 4 correspond to the run-up period. The black curve plots the average level of the measure across all run-up episodes, while the orange and green curves plot the measure across run-up episodes with negative and positive future 3-year returns following the run-up. The dotted lines represent one standard-error intervals around the mean.

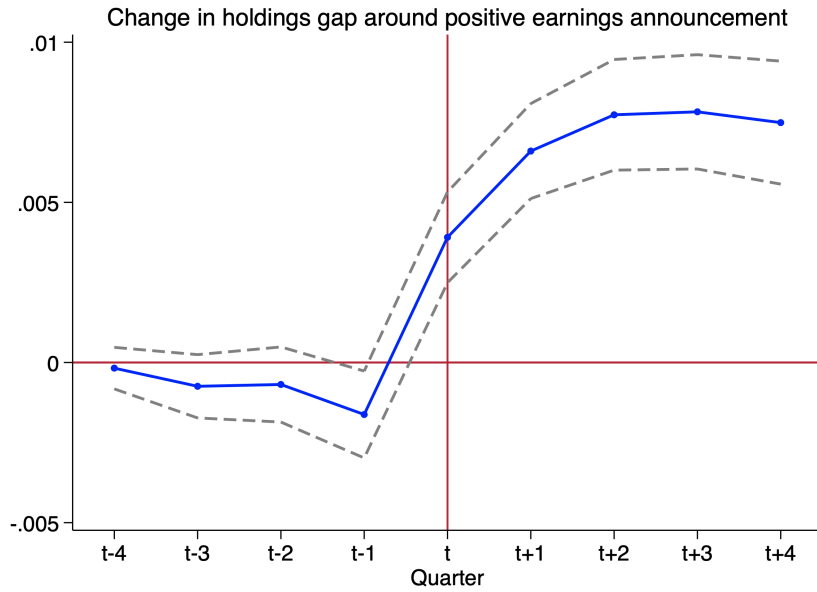


Figure 5: Change in holdings gap around positive earnings announcement

Notes: Figure 5 plots the regression coefficient β_{it} of the regression specification in Equation (31), which traces out the change in the holdings gap measure of stocks with a positive earnings announcement in quarter t , controlling for quarter fixed effects and unconditional autocorrelation in announcements (indicators for having a positive earnings announcement in quarters $t - 4$ to $t - 1$). Observations are at the stock by quarter level. The dotted lines represent the 95% confidence interval around each estimate of β_{it} .

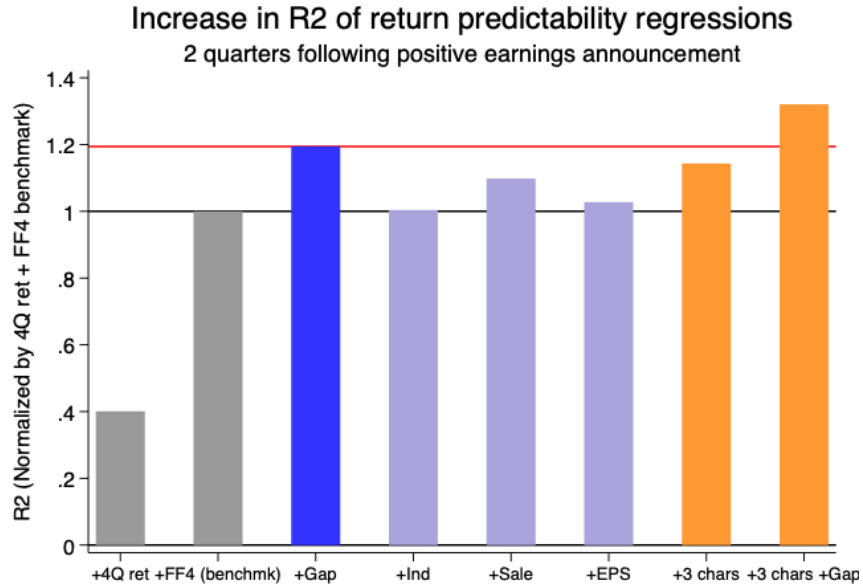


Figure 6: Increase in return prediction R^2 , by non-price measure

Notes: Figure 6 shows the explanatory power of our holdings gap measure in predicting the 2-quarter post-announcement returns following positive announcements, as specified in Equation (33). The first column shows (in basis points) the adjusted R^2 of using just the past 4 quarter returns including the announcement quarter return. The second column shows the adjusted R^2 of adding the stock's four FF4 loadings in addition to the returns. The third column shows the R^2 after adding the 4-quarter path of the holdings gap measure in addition to the returns and FF4 loadings. The fourth, fifth, and sixth columns show the the R^2 after separately adding, respectively, industry winners, sales, and EPS growth in addition to returns and FF4 loadings. The seventh column shows the R^2 of adding all three of these variables in addition to returns and FF4 loadings. The final column shows the total explanatory power of adding in all three non-price variables and the holdings gap measure in addition to returns and FF4 loadings.

Variable	N	Mean	SD	P25	P50	P75
Quarterly Returns	617259	0.034	0.304	-0.099	0.016	0.133
Market Capitalization [Millions]	617259	2695.348	15429.058	55.167	217.521	962.821
Holdings Gap	617259	0.521	0.134	0.445	0.538	0.606
Investor Equity Holdings [Millions of dollars]	346357	3231.709	29441.910	92.633	240.121	889.751
Active Years [Years]	346357	8.732	8.416	2.249	6.005	13.008

Table 1: Summary Statistics

Notes: Table 1 presents the summary statistics for the main stock and investor variables used in our analysis using data from 1980 to 2020. The holdings gap measure of a stock is given by Equation (25), which is the average news sensitivity of investors in the stock weighted by their holdings. Investor equity holdings at a given quarter is the total dollar value of the CRSP stock holdings reported in the 13F filing. The active years of an investor in a given quarter is given by the number of years from the first filings of the investor to the given quarter.

	Future NS			
	(1)	(2)	(3)	(4)
NS	0.38*** (0.02)	0.11*** (0.04)	0.36*** (0.02)	0.19*** (0.04)
Constant	0.32*** (0.01)	0.42*** (0.02)	0.34*** (0.01)	0.41*** (0.02)
Sample	All	Banks	Investment Firms	Pensions/Insurers
Observations	71,910	9,065	44,634	7,729
R ²	0.15	0.01	0.13	0.04

Table 2: Persistence of NS

Notes: Table 2 reports the estimates corresponding to a linear regression specification of the 8-quarter average standardized investor news sensitivity (NS) against the same variable 2 years in the future. NS is defined following Equation (23). Column (1) uses all 13F institutions, and Columns (2), (3), and (4) run the same analysis within banks, investment firms, and pensions/insurers. Observations are at the investor-quarter level. Standard errors are two-way clustered at the investor and quarter level and are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

	Next Year Returns			
	(1)	(2)	(3)	(4)
Past Returns	0.04 (0.03)	0.07** (0.03)	0.05* (0.03)	0.07** (0.04)
Gap		-0.05** (0.02)	-0.03** (0.01)	
Past Returns X Gap		-0.05** (0.03)	-0.06** (0.03)	
Past 3 Year Returns			0.04 (0.04)	
Gap (Buys)				-0.03 (0.01)
Past Returns X Gap (Buys)				-0.06** (0.02)
Observations	471,160	471,160	436,020	358,173
R ²	0.14	0.14	0.14	0.16

Table 3: Momentum conditional on holdings gap

Notes: Table 3 reports the momentum of a stock conditional on the level of the holdings gap measure, given by Equation (27). Observations are at the stock by quarter level. Past returns of the stock is given by the annualized average monthly returns of the stock for the past 12 months, excluding the recent month (Carhart, 1997). Gap is the percentile rank of the holdings gap measure $\text{Gap}_{s,t}^{emp}$ across all stocks in quarter t , with $\text{Gap}_{s,t}^{emp}$ defined following Equation (25). Next year returns is the log returns of stock s from quarters $t + 1$ to $t + 4$ inclusive. The observation is at the stock by quarter level, consisting of all stock-quarter pairs with non-missing values of Gap with at least 6 months of returns as of quarter t . Column (1) shows the baseline return predictability without conditioning on the holdings gap. Column (2) shows our main specification, which indicates a lower momentum coefficient for stocks with higher holdings gap. Columns (3) controls for the 3 year returns of the stock. Column (4) replaces the holdings gap measure by the average news sensitivity of investors buying the stock between quarters $t - 3$ to t , standardized to its percentile rank across all stocks in quarter t . Driscoll-Kraay standard errors with four lags are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

	Non-overreactive momentum			Overreactive momentum		
	(1)	(2)	(3)	(4)	(5)	(6)
Alpha (%)	0.0906** (0.0415)	0.152*** (0.0454)	0.139*** (0.0378)	-0.0207 (0.0362)	-0.0128 (0.0463)	0.0229 (0.0279)
Market return		-0.853*** (0.282)	-0.790*** (0.249)		-0.109 (0.320)	-0.233 (0.189)
FF SMB			-0.190 (0.336)			-0.529* (0.307)
FF HML			0.508** (0.256)			-1.044*** (0.303)
Observations	337	337	337	337	337	337

Table 4: Momentum portfolio returns, double sorted by investor overreaction score

Notes: Table 4 examines the 12 month returns of two double-sorted momentum strategies which condition on both past returns and the level of holdings gap. Observations are at the monthly level. The portfolios are formed by double sorting stocks in each month t into (a) the decile of cumulative log returns over the past 12 months excluding the most recent month $t - 1$ (Carhart, 1997) and (b) the quintile of the holdings gap measure in the quarter before month t 's quarter, which is defined following Equation (25). The non-overreactive momentum portfolio (columns (1) to (3)) goes long stocks in the 10th decile of returns and 1st quintile of the holdings gap and shorts stocks in the 1st decile of returns and 5th quintile of the holdings gap. The overreactive momentum portfolio (columns (4) to (6)) goes long stocks in the 10th decile of returns and 5th quintile of the holdings gap and shorts stocks in the 1st decile of returns and 1st quintile of the holdings gap. We compute abnormal returns by regressing the unweighted mean four-quarter returns of each sorted portfolio on the four-quarter returns of the market portfolio (the CRSP value-weighted portfolio), SMB (size), and HML (value) (Fama and French, 1993). The regressions are run at the monthly level. Columns (1) and (4) present the raw returns of both momentum portfolios. Columns (2) and (5) present the excess returns controlling for the market portfolio. Columns (3) and (6) present the alphas controlling for the size and value portfolios. Standard errors are estimated using the Newey-West method with a lag of twelve and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

	Returns		1(Returns < -40%)
	(1)	(2)	(3)
Gap	-0.12*** (0.04)	-0.09** (0.04)	0.04** (0.02)
Run-up Period Returns		-0.50*** (0.07)	0.23*** (0.03)
Quarter FEs	Yes	Yes	Yes
Observations	9,776	9,776	9,776
R ²	0.17	0.19	0.15

Table 5: Return Predictability: Price Run-Ups

Notes: Table 5 reports the estimates corresponding to Equation (28). Observations are at the run-up episode level. Returns is the three-year future return of the stock following quarter $t^e + 4$. $1(\text{Returns} < -40\%)$ is an indicator variable for whether the three-year future return of the stock is less than -40%. $\text{Gap}_{s,t^e+4}^{emp}$ is the standardized holdings gap measure for the stock that quarter. Run-Up Period Returns is the four-quarter return of the stock up to the most recent quarter. Standard errors are two-way clustered at the stock by quarter level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

	Long NS		Long P&L	
	(1) Run-ups	(2) Others	(3) Run-ups	(4) Others
Alpha (%)	-0.109** (0.0454)	-0.00761 (0.0189)	0.0552* (0.0319)	0.0448** (0.0191)
Market	0.441** (0.182)	0.00740 (0.0629)	-0.197 (0.122)	-0.0855 (0.0729)
Size	-0.254* (0.132)	-0.0402 (0.0565)	0.108 (0.115)	0.0374 (0.0523)
Value	-0.149 (0.270)	-0.0952 (0.0886)	0.307 (0.214)	0.375*** (0.0832)
Momentum	0.0366 (0.286)	-0.172 (0.108)	0.144 (0.184)	0.159* (0.0856)
Observations	115	121	115	121

Table 6: Long-short portfolio returns double-sorted by NS and P&L measure

Notes: Table 6 compares the 4-quarter abnormal returns of a long-short portfolio of a purged NS portfolio to that of a purged P&L portfolio. Observations are at the quarterly level. The returns of the purged NS and P&L portfolios are given by Equation (71). Columns (1) and (3) show the performance of the two portfolios in run-ups, where the sorted portfolios in Equation (71) are constructed using stock-quarter pairs s, t where s has experienced a run-up (greater than 100% returns) from quarters $t - 4$ to $t - 1$. Columns (2) and (4) show the unconditional performance of the portfolios, constructed using the remaining stock-quarter pairs. We compute abnormal returns by regressing the 4-quarter returns of the two portfolios on the 4-quarter returns of market, which is the CRSP value-weighted portfolio, size, value, and momentum, which are constructed based off of stock-level characteristics (Fama and French, 1993; Carhart, 1997) as of the given quarter and held fixed for the next four quarters. The regressions are run at the quarterly level. Standard errors are estimated using the Newey-West method with a lag of four and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

	(1)	(2)	(3)	(4)	(5)	(6)
	NS	NS ^{eps}	NS	NS ^{eps}	NS	NS ^{eps}
Price Extrapolation	0.598*** (0.00829)	0.354*** (0.0181)				
Buy Performance			-0.149*** (0.0105)	-0.235*** (0.00956)		
Late					0.0328*** (0.00405)	0.0935*** (0.00607)
Quarter FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	178941	177647	159322	158202	127804	126867

Table 7: Correlates of NS with other measures of overreactiveness

Notes: Table 7 correlates investors' news sensitivity with other investor characteristics. The price extrapolation and buy performance of an investor are defined by Equation (29), which measures the average past year returns of stocks bought by the investor and their benchmarked future 4-quarter returns. Lastly, the late measure is defined by Equation (30), which reflects the tendency of an investor's trading to lag past institutional flow. See the main text for more details. Observations are at the investor-quarter level. Columns (1), (3), and (5) show the correlation between the three measures with the main NS measure, defined in Equation (23), where we add in quarter fixed effects. Columns (2), (4), and (6) show the correlation between the three measures with an alternative NS measure, computed from EPS surprises, as defined in Equation (22). Standard errors are two-way clustered at the investor and quarter level and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

	Past returns (1)	Curr. returns (2)	Industry (3)	Sales growth (4)	EPS growth (5)
Positive news	0.831*** (0.0499)	0.281*** (0.0453)	0.481*** (0.0907)	0.841*** (0.0506)	0.782*** (0.0514)
Price var.	-0.0547 (0.0611)	0.959*** (0.0922)			
Positive news X price var.	0.184*** (0.0590)	0.243*** (0.0561)			
Nonprice var.			0.584 (0.374)	-0.0184 (0.0436)	-0.0542 (0.0427)
Positive news X nonprice var.			0.193*** (0.0396)	0.140*** (0.0531)	0.0930** (0.0466)
Lagged return and score controls	No	Yes	Yes	Yes	Yes
Quarter-industry FEs	Yes	Yes	Yes	Yes	Yes
Observations	92567	92567	91855	89104	67165

Table 8: Holdings gap response (BPs) to positive earnings announcement (t-1 to t+1)

Notes: Table 8 shows the estimates of Equation (32), which predicts the 2-quarter response in the holdings gap (the change in the level from quarter $t - 1$ to $t + 1$) to the price and non-price characteristics of the announcement in quarter t , including quarter-by-industry fixed effects and (other than in Column (1)) controlling for $t - 1$ level of the holdings gap and the past 4 quarter log returns. Observations are at the stock by quarter level. The independent variables include $positive_{s,t}$, the indicator for whether stock s has a positive announcement in quarter t , $price_{s,t}$, the price variable of stock s as of quarter t and their interaction $positive_{s,t} \times price_{s,t}$, as well as $Z_{s,t}$, the non-price variable of stock s in quarter t and the interaction $positive_{s,t} \times Z_{s,t}$. In Column (1), $price_{s,t}$ is cumulative log returns over quarters $t - 4$ to $t - 1$, while in Column (2), $price_{s,t}$ is the log return in the announcement quarter t . In Column (3), $Z_{s,t}$ is the performance of industry winners, defined as the average top decile 4-quarter returns of stocks in the same 3-digit SIC industry as s . In Column (4), $Z_{s,t}$ is the past 4 quarter sales growth of s . In Column (5), $Z_{s,t}$ is the past 4 quarter EPS growth of s . Standard errors are clustered at the quarter-by-industry level, where we cluster at the level of 2-digit SIC code. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Online Appendix

Michael Blank, Spencer Kwon and Johnny Tang

January 14, 2023

A Proofs

A.1 Derivation for the static model

To simplify the derivation, we prove the following lemma that derives the correspondence between the price-news relationship and the momentum-reversal coefficient.

Lemma 3. *Let the fundamentals V be distributed according to the prior distribution centered around 0 with variance τ_V^{-1} . Given $n = V + \varepsilon$ with $\text{Var}[\varepsilon] = \tau_\varepsilon^{-1}$, and prices given by the linear equilibrium: $p = \psi^0 \cdot n$, the expected returns is given by:*

$$E[V - p|p] = \left(\frac{\tau_\varepsilon}{\tau_V + \tau_\varepsilon} (\psi^0)^{-1} - 1 \right) \cdot p \quad (34)$$

Proof. By properties of normal random variables, it suffices to compute the covariance of V and p , as well as the variance of p . By assumption, the variance of p is given by $(\psi^0)^2 \cdot \text{Var}[n]$, where $\text{Var}[n] = \tau_V^{-1} + \tau_\varepsilon^{-1}$. Second, the covariance of V and p are given by:

$$\text{Cov}(V, p) = \psi^0 \cdot \text{Cov}(V, V + \varepsilon) = \psi^0 \cdot \tau_V^{-1} \quad (35)$$

Thus, one can evaluate the expected value of V conditional on p :

$$\begin{aligned} E[V - p|p] &= \left(\frac{\text{Cov}(V, p)}{\text{Var}[p]} - 1 \right) \cdot p \\ &= \left(\frac{\psi^0 \cdot \tau_V^{-1}}{(\psi^0)^2 (\tau_V^{-1} + \tau_\varepsilon^{-1})} - 1 \right) \cdot p \\ &= \left(\frac{\tau_\varepsilon}{\tau_V + \tau_\varepsilon} (\psi^0)^{-1} - 1 \right) \cdot p, \end{aligned} \quad (36)$$

as desired. □

Thus, to prove Lemma 1, it suffices to derive the expression for ψ^0 , the news-price coefficient. This follows from the market-clearing condition:

$$\begin{aligned} L \cdot p &= (1 - \chi) \frac{1}{A} (\tau_\varepsilon \cdot n - (\tau_V + \tau_\varepsilon)p) + \chi \frac{1}{A} ((1 + \Phi_n^{os}(s)) \tau_\varepsilon \cdot n + \Phi_p^{os}(s) \cdot p - (\tau_V + \tau_\varepsilon)p) \\ &= \frac{1}{A} ((1 + \chi \Phi_n^{os}(s)) \tau_\varepsilon n + \chi \Phi_p^{os}(s) \cdot p - (\tau_V + \tau_\varepsilon)p). \end{aligned} \quad (37)$$

Rearranging, one obtains:

$$p = \frac{(1 + \chi \Phi_n^{os}(s)) \tau_\varepsilon}{\tau_V + \tau_\varepsilon + AL - \chi \Phi_p^{os}(s)} \cdot n. \quad (38)$$

Thus, Lemma 3 immediately implies Lemma 1.

To derive Lemma 2 and Proposition 1, one can simply impose market-clearing and plug in the expression for rational demand to obtain:

$$\begin{aligned} \text{Gap}_s &= \frac{\chi \bar{D}_s^{os} - (1 - \chi) \bar{D}_s^{rat}}{\chi \bar{D}_s^{os} + (1 - \chi) \bar{D}_s^{rat}} = 1 - 2(1 - \chi) \frac{\bar{D}_s^{rat}}{L \cdot p} = 1 - 2(1 - \chi) \frac{1}{A} \frac{\tau_\varepsilon n - (\tau_V + \tau_\varepsilon)p}{L \cdot p} \\ &= 1 - 2(1 - \chi) \frac{1}{AL} (\tau_\varepsilon \psi_s^{-1} - (\tau_V + \tau_\varepsilon)), \end{aligned} \quad (39)$$

which immediately implies Lemma 2. One can obtain Proposition 1 by inverting the above expression to have ψ_s as a function of Gap_s and applying it to the expression in Lemma 1.

A.2 Derivation for the dynamic model and proof of Proposition 2

Diagnostic demand Consider the equilibrium in which all investors are diagnostic, with no price adjustment frictions (in other words, diagnostic investors do not think there are rational or sluggish investors as well). Suppose that diagnostic investors extract a signal of value $F^\theta(p_{0:t})$ of precision $\tau_{i,p}^\theta$ from public prices. Diagnostic expectations implies that they overreact to both the learning from prices and signals, which implies that the diagnostic demand function is given by:

$$D^\theta(n_{i,t}, p_t) = \overbrace{(1 + \theta_s) \left(\tau_{\varepsilon t} \cdot (n_{i,t}/t) + \tau_i^{p,\theta} \cdot F^\theta(p_{0:t}) \right)}^{\text{Diagnostic expectations of fundamentals}} - (\tau_V + \tau_\varepsilon + \tau_i^{p,\theta}) p_t. \quad (40)$$

In other words, $F^\theta(p_{0:t})$ is the signal that the diagnostic investors would infer from the prices prior to overreacting to it.

Aggregating across diagnostic investors, market-clearing implies the following price equation:

$$p_t = \frac{(1 + \theta_s) \tau_i^{p,\theta}}{\tau_V + \tau_\varepsilon + \tau_i^{p,\theta}} F^\theta(p_{0:t}) + \frac{(1 + \theta_s) \tau_{\varepsilon t}}{\tau_V + \tau_\varepsilon + \tau_i^{p,\theta}} V - \frac{A}{\tau_V + \tau_\varepsilon + \tau_i^{p,\theta}} q_t. \quad (41)$$

From the above, one can solve using continuous-time Gaussian updating to obtain:

$$\tau_i^{p,\theta} = \frac{(1 + \theta_s)^2 \tau_\varepsilon^2 \tau_q t}{A^2}, \quad (42)$$

and

$$(1 + \theta) F^\theta(p_{0:t}) = \frac{\tau_V + \tau_{\varepsilon t} \left(1 + \frac{(1 + \theta_s)^2 \tau_\varepsilon \tau_q}{A^2} \right)}{\tau_{\varepsilon t} \left(1 + \frac{(1 + \theta_s)^2 \tau_\varepsilon \tau_q}{A^2} \right)} p_t, \quad (43)$$

which immediately implies the expression for oversensitive demand, as desired.

Rational demand The diagnostic asset demand function is thus given by:

$$D^\theta(n_{i,t}, p_t) = \frac{1}{A} \left((1 + \theta_s) \tau_{\varepsilon t} (n_{i,t}/t) + \Phi_p^\theta \cdot p_t - (\tau_V + \tau_{\varepsilon t}) p_t \right). \quad (44)$$

The rational agent obtains two independent signals involving the fundamentals V : one from the public price path, which is of value $F(p_{0:t})$ with precision τ_i^p , and the other from the idiosyncratic news signal, which is centered at V with precision $\tau_\varepsilon \cdot t$. Thus, Bayesian normal updating and CARA investor demand implies:

$$D^{rat}(n_{i,t}, p_t) = \frac{1}{A} \left(\tau_{\varepsilon t} \cdot (n_{i,t}/t) + \tau_i^p \cdot F(p_{0:t}) - (\tau_V + \tau_{\varepsilon t} + \tau_i^p) p_t \right) \quad (45)$$

Market clearing then implies the following equilibrium price function:

$$p_t = \frac{(1 - \chi)\tau_t^p}{\bar{\tau}_t} F(p_{0:t}) + \frac{(1 + \chi\theta_s)\tau_\varepsilon t}{\bar{\tau}_t} V - \frac{A}{\bar{\tau}_t} q_t = a_t F(p_{0:t}) + b_t V - c_t q_t, \quad (46)$$

where $\bar{\tau}_t = AL + \tau_V + \tau_\varepsilon t + (1 - \chi)\tau_t^p - \chi\Phi_p^\theta$.

Denote $\xi_t = \frac{b_t}{c_t} = \frac{(1 + \chi\theta_s)\tau_\varepsilon t}{A}$. Rearranging the above into a differential form, one obtains:

$$d\left(\frac{p_t}{c_t} - \frac{a_t}{c_t} F(p_{0:t})\right) = [d\xi_t \cdot V - dq_t] = \frac{(1 + \chi\theta_s)\tau_\varepsilon}{A} dt \cdot \mathbf{v}_t, \quad (47)$$

where $\mathbf{v}_t = V - \frac{A \cdot dq_t}{(1 + \chi\theta_s)\tau_\varepsilon \cdot dt}$ is the marginal independent signal of V obtained from prices at time t , of precision $\text{Var}\left[\frac{A \cdot dq_t}{(1 + \chi\theta_s)\tau_\varepsilon \cdot dt}\right]^{-1} = (1 + \chi\theta_s)^2 \tau_\varepsilon^2 \frac{\tau_q}{A^2} dt$.

Gaussian updating implies that $F(p_{0:t})$ evolves according to the following differential equation:

$$d(\tau_t^p \cdot F(p_{0:t})) = d\tau_t^p \cdot \mathbf{v}_t. \quad (48)$$

This can be viewed as the continuous time version of standard Gaussian updating of independent normal signals, where the posterior is given by the weighted average of each signal by its precision. Given that the marginal signal is of precision $(1 + \chi\theta_s)^2 \tau_\varepsilon^2 \frac{\tau_q}{A^2} dt$, integrating across time immediately implies:

$$\tau_t^p = (1 + \chi\theta_s)^2 \tau_\varepsilon^2 \frac{\tau_q}{A^2} t. \quad (49)$$

Rearranging Equation (48) yields

$$\mathbf{v}_t = \frac{A}{(1 + \chi\theta_s)\tau_\varepsilon dt} \cdot d\left(\frac{p_t}{c_t} - \frac{a_t}{c_t} F(p_{0:t})\right). \quad (50)$$

We then plug the above expression into Equation (48), and obtain:

$$d(\tau_t^p \cdot F(p_{0:t})) = \frac{(1 + \chi\theta_s)\tau_\varepsilon \tau_q}{A} \cdot d\left(\frac{p_t}{c_t} - \frac{a_t}{c_t} F(p_{0:t})\right). \quad (51)$$

Integrating both sides and simplifying yields the following expression for $F(p_{0:t})$, the value of

the public signal obtained from the price path $p_{0:t}$:

$$\begin{aligned}
\frac{(1 + \chi\theta_s)\tau_\varepsilon t}{A} \cdot F(p_{0:t}) &= \left(\frac{p_t}{c_t} - \frac{a_t}{c_t} F(p_{0:t}) \right) \\
\iff F(p_{0:t}) &= \left(a_t + \frac{(1 + \chi\theta_s)\tau_\varepsilon t}{A} c_t \right)^{-1} p_t \\
&= \frac{AL + \tau_V + \tau_\varepsilon t + (1 - \chi)\tau_t^p - \chi\Phi_p^\theta}{(1 - \chi)\tau_t^p + (1 + \chi\theta_s)\tau_\varepsilon t} p_t \\
&= \left(1 + \frac{AL + \tau_V - \chi\Phi_p^\theta - \chi\theta_s\tau_\varepsilon t}{(1 + \chi\theta_s)\tau_\varepsilon t \left(1 + (1 - \chi)\frac{(1 + \chi\theta_s)\tau_\varepsilon \tau_q}{A^2} \right)} \right) p_t.
\end{aligned} \tag{52}$$

Thus, we obtain that the signal of fundamentals given the entire price path $p_{0:t}$ only depends on the current price p_t .

Furthermore, setting $\Phi_{p,t}^{rat} = (F(p_{0:t})/p_t - 1) \cdot \tau_t^p$, one obtains:

$$\Phi_{p,t}^{rat} = \frac{\frac{(1 + \chi\theta_s)\tau_\varepsilon \tau_q}{A^2}}{1 + (1 - \chi)\frac{(1 + \chi\theta_s)\tau_\varepsilon \tau_q}{A^2}} \cdot [AL + \tau_V - \chi\Phi_p^\theta - \chi\theta_s\tau_\varepsilon t]. \tag{53}$$

Plugging the above expression into Equation (45) yields Proposition 2.

A.3 Momentum-reversal coefficient, holdings-gap, and the proof of Proposition 3

Recall that we have the following pricing equation:

$$p_t = \frac{(1-\chi)\tau_t^p}{\bar{\tau}_t} F(p_{0:t}) + \frac{(1+\chi\theta_s)\tau_\varepsilon t}{\bar{\tau}_t} V - \frac{A}{\bar{\tau}_t} q_t = a_t F(p_{0:t}) + b_t V - c_t q_t. \quad (54)$$

Inverting the pricing equation, one obtains:

$$p_t = \left(1 + (1-\chi) \frac{(1+\chi\theta_s)\tau_\varepsilon \tau_q}{A^2} \right) (b_t V - c_t q). \quad (55)$$

One can then compute the return predictability coefficient $\beta_t = \frac{\text{Cov}(V-p_t, p_t)}{\text{Var}(p_t)}$ using the above pricing equation. We denote $J = \frac{(1+\chi\theta_s)\tau_\varepsilon \tau_q}{A^2}$, $\bar{\tau}_\varepsilon = (1+\chi\theta_s)\tau_\varepsilon$, and $\bar{\tau}_t = AL + \tau_V + \tau_\varepsilon t + (1-\chi)\tau_t^p - \chi\Phi_p^\theta$. Then, the momentum-reversal coefficient β_t is given by:

$$\beta_t = \frac{\text{Cov}(V-p_t, p_t)}{\text{Var}(p_t)} = \frac{J(\tau_V + AL - \chi\theta_s \tau_\varepsilon t - \chi\Phi_p^\theta) - (1 + (1-\chi)J)\tau_V}{(1 + (1-\chi)J)[J\bar{\tau}_\varepsilon t + \tau_V]}. \quad (56)$$

Suppose we have sufficiently high investor sluggishness, as given by:

$$L > \frac{A\tau_V}{(1+\chi\theta_s)\tau_\varepsilon \tau_q}. \quad (57)$$

Lemma 4. *Assume L is sufficiently high, as given by Equation (57). Then, $\lim_{t \rightarrow 0} \beta_t > 0$: there is initial momentum. As t increases, there exists a t^* such that $\beta_t < 0$ if and only if $t > t^*$. Finally, β_t is bounded as $t \mapsto \infty$. In particular, $\lim_{t \rightarrow \infty} \beta_t = -\frac{\chi\theta_s}{(1+(1-\chi)J)(1+\chi\theta_s)} \equiv \beta_\infty$. In particular, $-1 < \beta_\infty < 0$.*

Proof. To investigate the dynamics of β_t , note that rearranging the numerator above, we find that $\beta_t < 0$ if and only if:

$$\frac{AL + \left(\frac{\chi A^2}{A^2 + (1+\theta_s)^2 \tau_\varepsilon \tau_q} - J^{-1} \right) \tau_V}{\chi\theta_s \tau_\varepsilon} < t. \quad (58)$$

The above inequality is binding as long as L is sufficiently large, and it is trivial to see that the assumption in Equation (57) is sufficient to ensure $\beta_t > 0$ for $t = 0$ and $\beta_t < 0$ for $t > t^* \equiv \frac{AL + \left(\frac{\chi A^2}{A^2 + (1+\theta_s)^2 \tau_\varepsilon \tau_q} - J^{-1} \right) \tau_V}{\chi\theta_s \tau_\varepsilon}$, as desired.

To obtain the long-run dynamics of β_t , note that Equation (56) immediately implies:

$$\lim_{t \rightarrow \infty} \beta_t = \frac{-\chi\theta_s \tau_\varepsilon}{(1 + (1-\chi)J)\bar{\tau}_\varepsilon} = -\frac{\chi\theta_s}{(1 + (1-\chi)J)(1 + \chi\theta_s)}, \quad (59)$$

as desired. \square

Next, to see whether the holdings gap is above or below $\text{Gap}^0 = 2\chi - 1$, we compare the total rational versus diagnostic demand. Note:

$$E[\bar{D}^{rat} - \bar{D}^\theta | p_t, \theta_s] = \frac{1}{A} \left((\Phi_{p,t}^{rat} - \Phi_p^\theta) - \theta \tau_\varepsilon t \eta_t \right) p_t, \quad (60)$$

where $\eta_t = 1 + \beta_t = E[V | p_t, \theta_s] / p_t$. In other words, whether $\text{Gap}_t = \frac{\chi \bar{D}^\theta - (1-\chi) \bar{D}^{rat}}{\chi \bar{D}^\theta + (1-\chi) \bar{D}^{rat}}$ is greater than $2\chi - 1$ if and only if

$$(\Phi_{p,t}^{rat} - \Phi_p^\theta) < \theta \tau_\varepsilon t \eta_t. \quad (61)$$

Lemma 5. *Continue to assume that L is sufficiently high, as specified by Equation (57). Then, as t goes from 0 to ∞ , $\Phi_{p,t}^{rat} - \Phi_p^\theta$ is a decreasing affine function in t which goes from positive to negative.*

Proof. First, we compute the asymptotics of $\Phi_{p,t}^{rat} - \Phi_p^\theta$ as $t \mapsto 0$.

$$\lim_{t \rightarrow 0} \Phi_{p,t}^{rat} - \Phi_p^\theta = \frac{(1 + \chi \theta_s) \tau_\varepsilon \tau_q \cdot [AL + \tau_V] - (A^2 + (1 + \chi \theta_s) \tau_\varepsilon \tau_q) \Phi_p^\theta}{A^2 + (1 - \chi)(1 + \chi \theta_s) \tau_\varepsilon \tau_q}. \quad (62)$$

The above quantity is positive if and only if:

$$\begin{aligned} (1 + \chi \theta_s) \tau_\varepsilon \tau_q \cdot [AL + \tau_V] &> (A^2 + (1 + \chi \theta_s) \tau_\varepsilon \tau_q) \frac{(1 + \theta_s)^2 \tau_\varepsilon \tau_q}{A^2 + (1 + \theta_s)^2 \tau_\varepsilon \tau_q} \tau_V \\ \iff L &> \left(\frac{((1 + \theta_s)^2 - (1 + \chi \theta_s)) \tau_\varepsilon \tau_q}{A^2 + (1 + \theta_s)^2 \tau_\varepsilon \tau_q} \right) \frac{A \tau_V}{(1 + \chi \theta_s) \tau_\varepsilon \tau_q}. \end{aligned} \quad (63)$$

The last inequality clearly follows from our assumption of the lower bound of L in Equation (57).

To conclude, note that Φ_p^θ is constant in t , and $\Phi_{p,t}^{rat}$ is a monotonically decreasing function in t , and is in fact an affine function in t . Consequently, the term $\Phi_{p,t}^{rat} - \Phi_p^\theta$ decreases linearly in t and eventually becomes negative, as desired. \square

To summarize, we have the following necessary and sufficient condition for there to be reversals, as well as a greater tilt towards oversensitive investors:

$$\begin{aligned} \chi \theta_s \tau_\varepsilon t &> AL + \left(\frac{\chi A^2}{A^2 + (1 + \theta_s)^2 \tau_\varepsilon \tau_q} - \frac{A^2}{(1 + \chi \theta_s) \tau_\varepsilon \tau_q} \right) \tau_V \\ \theta \tau_\varepsilon t \eta_t &> \frac{(1 + \chi \theta_s) \tau_\varepsilon \tau_q \cdot [AL + \tau_V - \chi \theta_s \tau_\varepsilon t] - (A^2 + (1 + \chi \theta_s) \tau_\varepsilon \tau_q) \Phi_p^\theta}{A^2 + (1 - \chi)(1 + \chi \theta_s) \tau_\varepsilon \tau_q} \end{aligned} \quad (64)$$

Cross-sectionally, when one varies θ_s , the first condition is clearly satisfied as θ_s becomes sufficiently large. For the second condition, note that Lemma 4 implies that the RHS of the second condition converges to a negative linear function in θ_s , while the RHS is positive, given that $\eta_t = 1 + \beta_t \geq 0$. In terms of dynamics, combining Equation (61) and Lemma 4 and 5 immediately implies short-run momentum/long-run reversals as well as the dynamics of the holdings gap, as desired.

A.4 Time-varying price extrapolativeness and proof of Corollary 1

We compute the difference in the momentum tendency of oversensitive investors $\text{Extrap}_t^\theta = \frac{\text{Cov}(D_t^\theta(n_{i,t}, p_{0:t}), p_t)}{\text{Var}[p_t]}$ and rational investors $\text{Extrap}_t^{\text{rat}} = \frac{\text{Cov}(D_t^{\text{rat}}(n_{i,t}, p_{0:t}), p_t)}{\text{Var}[p_t]}$ as a function of time.

Note:

$$\text{Extrap}_t^\theta = \frac{1}{A} \left(\Phi_p^\theta - (\tau_V + \tau_\varepsilon t) + \Phi_{n,t}^\theta \cdot \eta_t \right), \quad (65)$$

where $\eta_t = \frac{\text{Cov}(V, p_t)}{\text{Var}[p_t]}$. In other words, the price extrapolativeness of an investor depends on Φ_p^θ , the direct learning from prices, $\tau_V + \tau_\varepsilon t$, the standard price effect on demand, and $\Phi_{n,t}^\theta \cdot \eta_t$, the endogenous momentum, which results from investors inferring value from their private signals, which is correlated with the contemporaneous price increase. Analogously, the momentum tendency for rational investors is given by:

$$\text{Extrap}_t^{\text{rat}} = \frac{1}{A} \left(\Phi_{p,t}^{\text{rat}} - (\tau_V + \tau_\varepsilon t) + \Phi_{n,t}^{\text{rat}} \cdot \eta_t \right). \quad (66)$$

First, consider t to be sufficiently close to 0. In that case, the momentum tendency of oversensitive investors converges to:

$$\lim_{t \rightarrow 0} \text{Extrap}_t^\theta = -\frac{1}{A} \left(\frac{A^2}{A^2 + (1 + \theta_s)^2 \tau_\varepsilon \tau_q} \tau_V \right) < 0, \quad (67)$$

which is negative. Intuitively, by assumption oversensitive investors are not aware of the presence of sluggish investors. Consequently, early on in the momentum process, where there is no information either in prices or news to overreact to, they do not display momentum behavior.

In contrast, the price extrapolativeness of rational investors converges to:

$$\begin{aligned} \lim_{t \rightarrow 0} \text{Extrap}_t^{\text{rat}} &= \frac{1}{A} \left(\frac{(1 + \chi \theta_s) \tau_\varepsilon \tau_q}{A^2 + (1 - \chi)(1 + \chi \theta_s) \tau_\varepsilon \tau_q} \cdot [AL + \left(1 - \chi \frac{(1 + \theta_s)^2 \tau_\varepsilon \tau_q}{A^2 + (1 + \theta_s)^2 \tau_\varepsilon \tau_q} \right) \tau_V] - \tau_V \right) \\ &\propto L - A \left[\frac{1}{(1 + \chi \theta_s) \tau_\varepsilon \tau_q} - \frac{\chi}{A^2 + (1 + \theta_s)^2 \tau_\varepsilon \tau_q} \right] \tau_V. \end{aligned} \quad (68)$$

In other words, if $L > \frac{A}{(1 + \chi \theta_s) \tau_\varepsilon \tau_q} \tau_V$, rational investors engage in early-stage momentum, profiting off of the sluggishness of investors that do not react to any information, as proxied by L .

Next, we analyze the asymptotics of Extrap_t^θ and $\text{Extrap}_t^{\text{rat}}$ as t becomes sufficiently large. For oversensitive investors, we have:

$$\text{Extrap}_t^\theta = \tau_\varepsilon t \cdot ((1 + \theta_s) \eta_\infty - 1 + o(t)) = \tau_\varepsilon t \theta_s \cdot \left(1 - \frac{\chi(1 + \theta_s)}{1 + \chi \theta_s} \cdot \frac{1}{1 + (1 - \chi)(1 + \chi \theta_s) \tau_\varepsilon \tau_q / A^2} \right) > 0, \quad (69)$$

where the last inequality follows from $1 + \chi\theta_s > \chi(1 + \theta_s)$.

In contrast, for rational investors, we have:

$$\text{Extrap}_t^{\text{rat}} = \tau_\varepsilon t \cdot \left(\eta_\infty - 1 + o(t) - \frac{\chi\theta_s(1 + \chi\theta_s)\tau_\varepsilon\tau_q}{A^2 + (1 - \chi)(1 + \chi\theta_s)\tau_\varepsilon\tau_q} \right) < 0, \quad (70)$$

where the last inequality follows from Lemma 4 with $\beta_\infty < 0$, as desired.

Lastly, given that Extrap_t^θ and $\text{Extrap}_t^{\text{rat}}$ are asymptotically linear functions of t that are increasing (decreasing) in t , one immediately obtains $\text{Extrap}^\theta = 1/T \int_t \text{Extrap}_t^\theta dt > \text{Extrap}^{\text{rat}} = 1/T \int_t \text{Extrap}_t^{\text{rat}} dt$ if T is sufficiently large.

B Additional Figures and Tables

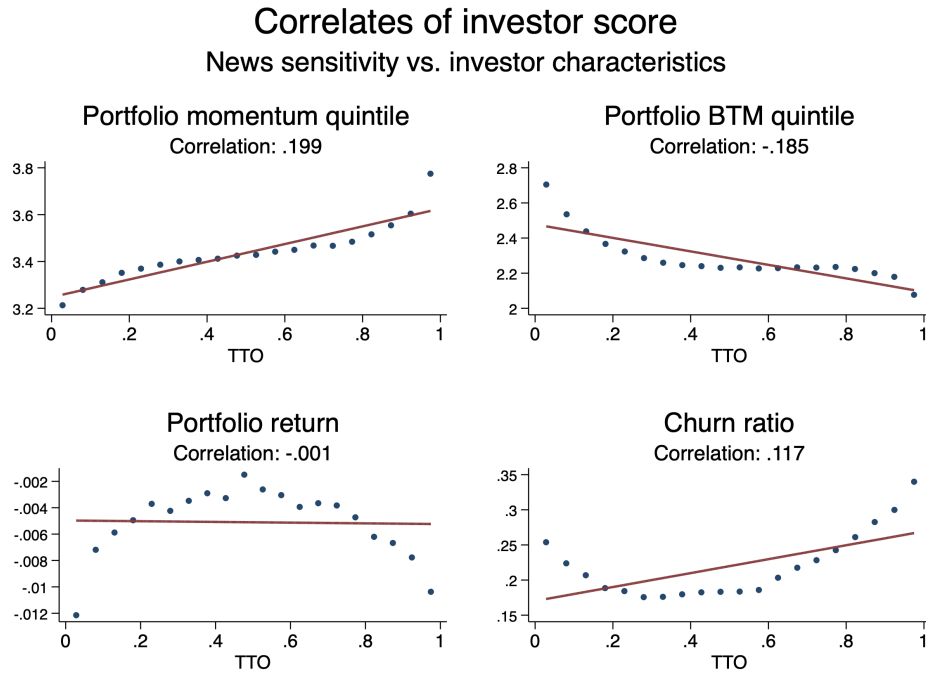


Figure A1: NS vs styles, returns, and volume

Notes: Figure A1 presents a binned scatterplot of the investor-level correlation between an investor's news sensitivity and other alternative measures. The portfolio momentum quintile is the dollars-weighted average momentum quintile of stocks held by the investor. The portfolio BTM quintile is the average book-to-market quintiles of stocks held by the investor. The portfolio return is the average returns (computed from quarterly holdings) of the investor, benchmarked to size, value, and momentum benchmarks. Finally, the churn ratio is the share of total holdings that an investor changes to a different asset each quarter. All variables are first demeaned at the quarterly level and then averaged across all quarters that the given investor is in the 13F. Observations are at the investor-quarter level.

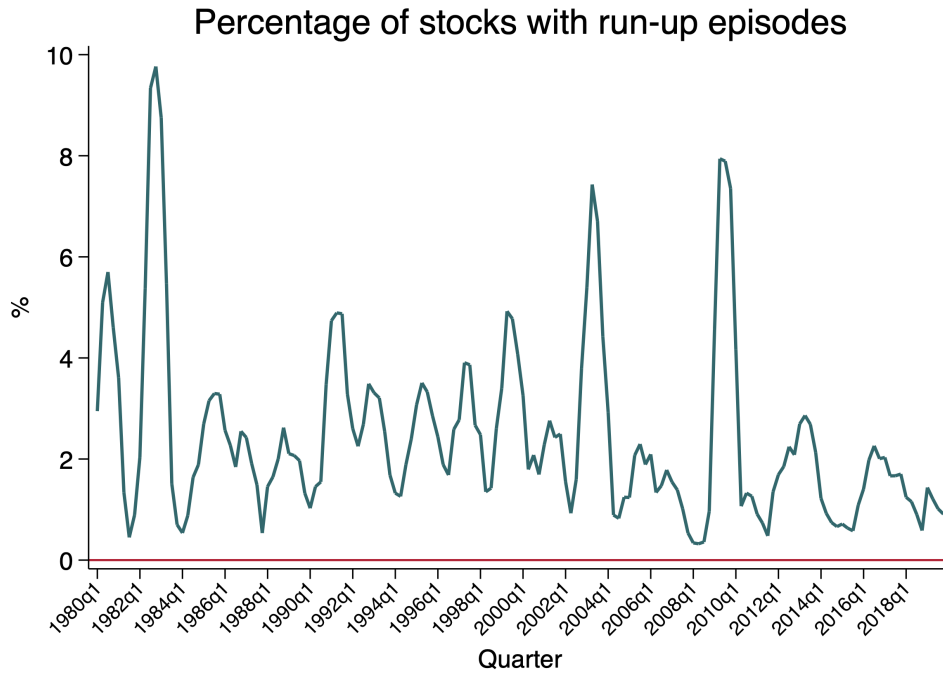
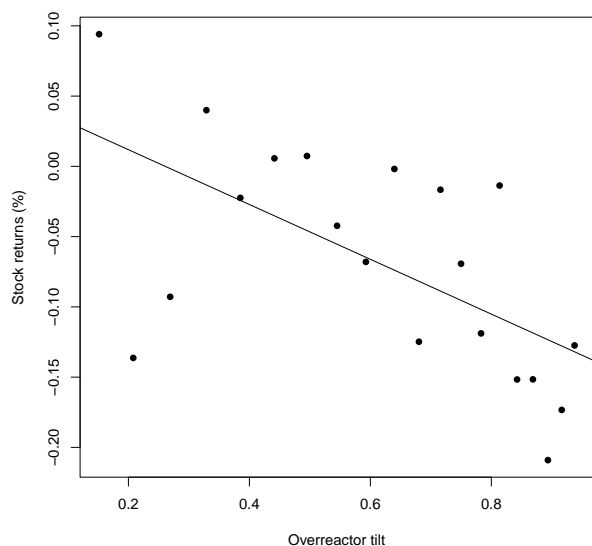


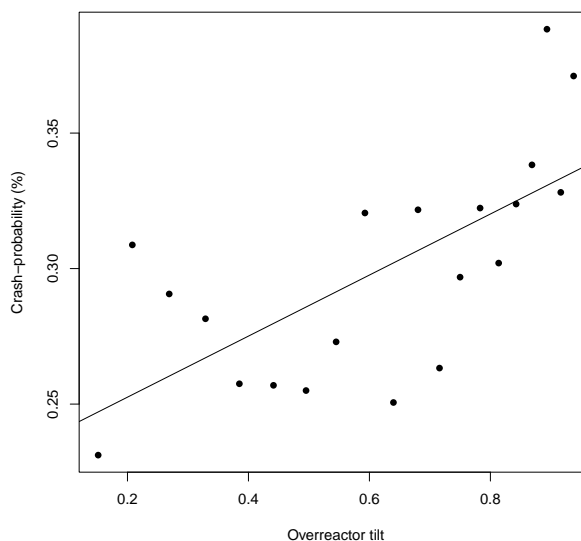
Figure A2: Percentage of stocks with run-up episodes

Notes: Figure A2 displays the proportion of stocks experiencing a run-up episode each quarter from 1980 to 2020, where a run-up episode is defined as a stock experiencing greater than 100% returns in the past 4 quarters.

Figure A3: Return predictability in bubbles



(a) Return Predictability: Momentum Episodes



(b) Return Predictability: Crash Probability

Notes: Figure A3a displays a binned scatterplot of the future 3-year return and the holdings gap for run-up episodes. The holdings gap is defined following Equation (25). Observations are at the stock-quarter level. Figure A3b displays a regression line of the crash probability and the holdings gap. Crash probability is an indicator variable for whether the stock had a three-year future return lower than -40% . Observations are at the stock-quarter level.

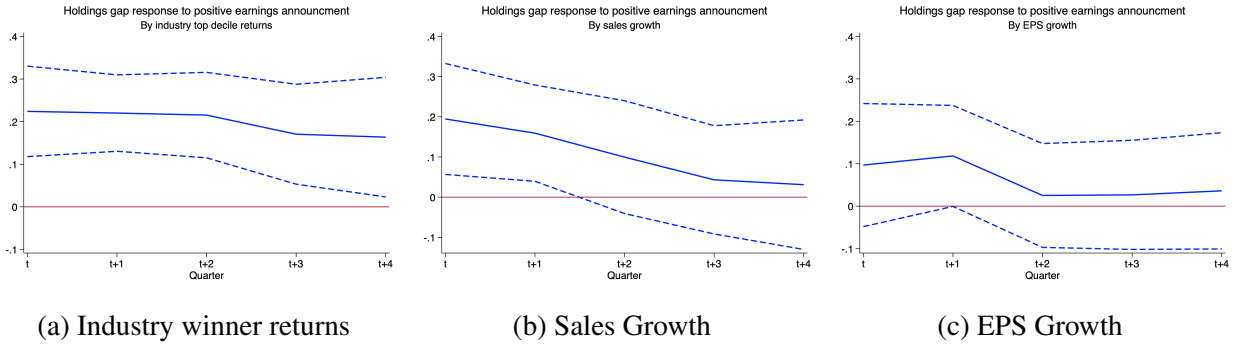


Figure A4: Response of holdings gap to non-price information

Notes: Figure A4 shows the estimates of Equation (32), where we regress the change in the holdings gap from quarter $t - 1$ to quarter $t + h$ from $h = 0$ to $h = 4$ to the non-price characteristics of the announcement in quarter t , controlling for the past 4 quarter log returns and including quarter-by-industry fixed effects. In Figure A4a, the non-price variable is the performance of industry winners, defined as the average top decile 4-quarter returns of stocks in the same 3-digit SIC industry as s . In Figure A4b and Figure A4c, we use the past 4 quarter sales growth and earnings-per-share growth of s . The dotted lines show the standard errors, which are clustered at the quarter-by-industry level, where we cluster at the level of 2-digit SIC code.

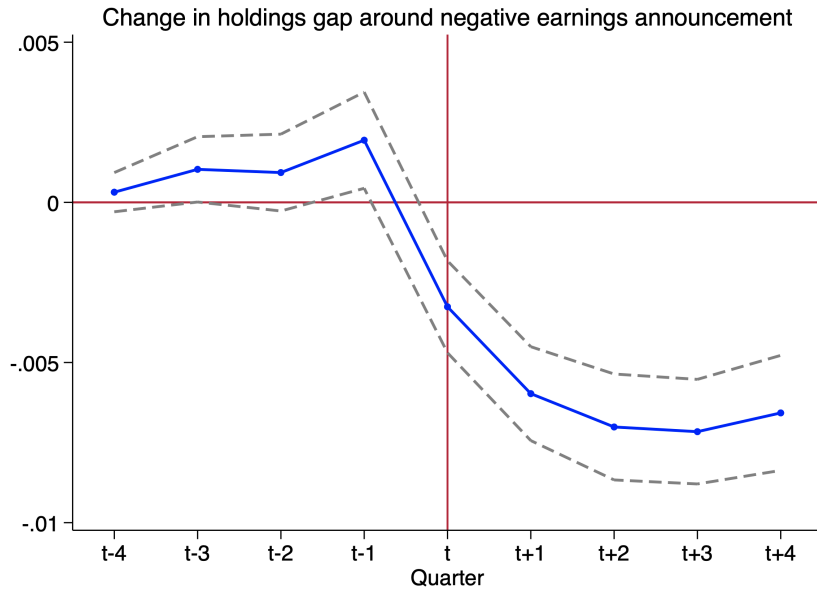


Figure A5: Change in holdings gap around negative earnings announcement

Notes: Figure 5 plots the regression coefficient β_h of the regression specification in Equation (31), which traces out the change in the holdings gap measure of stocks with a negative earnings announcement in quarter t , controlling for quarter fixed effects and unconditional autocorrelation in announcements (indicators for having a negative earnings announcement in quarters $t - 4$ to $t1$). The dotted lines represent the 95% confidence interval around each estimate of β_h .

1(<i>IIA</i>)	0.07*** (0.01)
1(<i>Others</i>)	0.04 (0.02)
1(<i>PIE</i>)	0.04*** (0.01)
ln(<i>AUM</i>)	0.01*** (0.003)
Quarter FEs	X
Observations	52,087
R ²	0.02

Table A1: Other investor characteristics associated with news sensitivity

Notes: Table A1 correlates the investor news sensitivity, as defined in Equation (23), with other investor characteristics. 1(*IIA*) is the indicator for whether an institution is an investment firm, 1(*PIE*) is the indicator for whether an institution is an pension or insurer, 1(*Others*) is the residual indicator. The coefficient compares the average news sensitivity of these institution groups to the final category, which are banks. *AUM* is the asset under management of these institutions. Observations are at the institution by quarter level, and we include quarter fixed effects as controls. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

	Long NS				Long P&L			
	(1) Ep	(2) Ep	(3) Non-ep	(4) Non-ep	(5) Ep	(6) Ep	(7) Non-ep	(8) Non-ep
Alpha (%)	-0.105*** (0.0366)	-0.109** (0.0454)	-0.0224 (0.0166)	-0.00761 (0.0189)	0.0698*** (0.0247)	0.0552* (0.0319)	0.0585*** (0.0195)	0.0448** (0.0191)
Market	0.432** (0.188)	0.441** (0.182)	0.0453 (0.0814)	0.00740 (0.0629)	-0.234** (0.108)	-0.197 (0.122)	-0.121 (0.0822)	-0.0855 (0.0729)
Size	-0.253* (0.129)	-0.254* (0.132)	-0.0371 (0.0686)	-0.0402 (0.0565)	0.111 (0.119)	0.108 (0.115)	0.0345 (0.0622)	0.0374 (0.0523)
Value	-0.176 (0.172)	-0.149 (0.270)	0.0303 (0.0831)	-0.0952 (0.0886)	0.200 (0.150)	0.307 (0.214)	0.375*** (0.0575)	0.492*** (0.0832)
Momentum		0.0366 (0.286)		-0.172 (0.108)		0.144 (0.184)		0.159* (0.0856)
Observations	115	115	121	121	115	115	121	121

Table A2: Long-short portfolio returns double-sorted by NS and P&L

Notes: Table A2 compares the 4-quarter abnormal returns of a long-short portfolio of a purged NS portfolio to that of a purged P&L portfolio. The returns of the purged NS and P&L portfolios are given by Equation (71). Columns (1), (2), (5), and (6) show the performance of the two portfolios in run-ups, where the sorted portfolios in Equation (71) are constructed using stock-quarter pairs s, t where s has experienced a run-up (greater than 100% returns) from quarters $t - 4$ to $t - 1$. Columns (3), (4), (7) and (8) show the unconditional performance of the portfolios, constructed using the remaining stock-quarter pairs. For columns (2), (4), (6), and (8), we compute abnormal returns by regressing the 4-quarter returns of the two portfolios on the 4-quarter returns of market, which is the CRSP value-weighted portfolio, size, value, and momentum, which are constructed based off of stock-level characteristics (Fama and French, 1993; Carhart, 1997) as of the given quarter and held fixed for the next four quarters. For the remaining columns, we do not control for the momentum factor. The regressions are run at the quarterly level. Standard errors are estimated using the Newey-West method with a lag of four and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

	Long NS				Long Buy Performance			
	(1) Ep	(2) Ep	(3) Non-Ep	(4) Non-Ep	(5) Ep	(6) Ep	(7) Non-Ep	(8) Non-Ep
Alpha (%)	-0.110*** (0.0356)	-0.130*** (0.0466)	-0.0212 (0.0158)	-0.00888 (0.0177)	0.115*** (0.0324)	0.130** (0.0509)	0.0533** (0.0206)	0.0389* (0.0204)
Market	0.426** (0.172)	0.480*** (0.177)	0.0291 (0.0743)	-0.00230 (0.0577)	-0.250** (0.104)	-0.290** (0.146)	-0.0926 (0.0812)	-0.0557 (0.0705)
Size	-0.309** (0.144)	-0.310** (0.152)	-0.00982 (0.0657)	-0.0125 (0.0560)	0.335** (0.148)	0.336** (0.155)	0.0531 (0.0654)	0.0562 (0.0558)
Value	-0.243 (0.148)	-0.0833 (0.206)	-0.0158 (0.0711)	-0.120 (0.0800)	0.301** (0.146)	0.186 (0.236)	0.308*** (0.0520)	0.431*** (0.0804)
Momentum		0.210 (0.239)		-0.142 (0.0992)		-0.152 (0.258)		0.167* (0.0924)
Observations	117	117	121	121	117	117	121	121

Table A3: Long-short portfolio returns double-sorted by NS and Buy Performance

Notes: Table A3 compares the 4-quarter abnormal returns of a long-short portfolio of a purged NS portfolio to that of a purged Buy-Performance portfolio. The returns of the purged NS and Buy-Performance portfolios are given by Equation (71), where one replaces P&L with an investor's buy-performance measure, as defined in Equation (?). Columns (1), (2), (5), and (6) show the performance of the two portfolios in run-ups, where the sorted portfolios in Equation (71) are constructed using stock-quarter pairs s, t where s has experienced a run-up (greater than 100% returns) from quarters $t - 4$ to $t - 1$. Columns (3), (4), (7) and (8) show the unconditional performance of the portfolios, constructed using the remaining stock-quarter pairs. For columns (2), (4), (6), and (8), we compute abnormal returns by regressing the 4-quarter returns of the two portfolios on the 4-quarter returns of market, which is the CRSP value-weighted portfolio, size, value, and momentum, which are constructed based off of stock-level characteristics (Fama and French, 1993; Carhart, 1997) as of the given quarter and held fixed for the next four quarters. For the remaining columns, we do not control for the momentum factor. The regressions are run at the quarterly level. Standard errors are estimated using the Newey-West method with a lag of four and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

	Positive announcement		Post-announcement change in holdings gap	
	(1)	(2)	(3)	(4)
	Positive	Positive	Δ gap, t-1 to t	Δ gap, t-1 to t+5
Δ gap, t-9 to t-1	-0.0395** (0.0186)			
Δ gap, t-5 to t-1		-0.0381 (0.0239)		
Positive announcement			0.00482*** (0.000340)	0.0100*** (0.000947)
8Q lagged return control	Yes	Yes	Yes	Yes
Quarter FEs	Yes	Yes	Yes	Yes
Observations	80148	86063	89919	79429

Table A4: Holdings gap around positive earnings announcement

Notes: Table A4 shows the predictive version of Figure 5, which shows that around the release of positive news, the holdings gap decreases (rational investors increase their shares) prior to the announcement, and increases after the announcement. Observations are at all stock-by-quarter pairs (s, t) for stock s having an annual earnings announcement in quarter t . In Columns (1) and (2), we regress $\text{positive}_{s,t}$, the indicator of stock s having an earnings announcement with positive 10-day window returns, on the change in the holdings gap from quarters $t - 9$ to $t - 1$ or $t - 5$ to $t - 1$, with a negative coefficient indicating that a decrease in the holdings gap predicts future positive announcements. In Columns (3) and (4), we regress future changes in the holdings gap, from quarters $t - 1$ to t or $t - 1$ to $t + 5$, on the indicator $\text{positive}_{s,t}$. A positive coefficient implies that a positive earnings announcement predicts a rise in the holdings gap, or an inflow of oversensitive investors. In all regressions, we control for lagged stock returns and quarter fixed effects. Standard errors are clustered at the stock-by-quarter level and reported in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

	Past returns (1)	Industry (2)	Sales growth (3)	EPS growth (4)
Positive announcement	0.228*** (0.0513)	0.0164 (0.105)	0.265*** (0.0483)	0.187*** (0.0442)
Price var	-0.0768 (0.0683)			
Positive X Price var	0.251*** (0.0615)			
Non-Price var.		0.554 (0.376)	0.0202 (0.0464)	-0.0503 (0.0435)
Positive X Non-Price var.		0.146*** (0.0467)	0.114** (0.0522)	0.108** (0.0477)
Lagged return and score controls	No	Yes	Yes	Yes
Quarter-industry FEs	Yes	Yes	Yes	Yes
Quarter t return control	Yes	Yes	Yes	Yes
Observations	92567	91855	89104	67165

Table A5: Holdings gap response (BPs) to positive earnings announcement (t-1 to t+1)

Notes: Table A5 shows the estimates of Equation (32), which predicts the 2-quarter response in the holdings gap (the change in the level from quarter $t - 1$ to $t + 1$) to the price and non-price characteristics of the announcement in quarter t , including quarter-by-industry fixed effects and (other than in Column (1)) controlling for $t - 1$ level of the holdings gap and the past 4 quarter log returns. The quarter t return is also included as a control. Observations are at the stock by quarter level. The independent variables include $positives_{s,t}$, the indicator for whether stock s has a positive announcement in quarter t , $price_{s,t}$, the return of stock s as of quarter t and their interaction $positive_{s,t} \times price_{s,t}$, as well as $Z_{s,t}$, the non-price variable of stock s in quarter t and the interaction $positive_{s,t} \times Z_{s,t}$. In Column (1), $Z_{s,t}$ is the performance of industry winners, defined as the average top decile 4-quarter returns of stocks in the same 3-digit SIC industry as s . In Column (2), $Z_{s,t}$ is the past 4 quarter sales growth of s . In Column (3), $Z_{s,t}$ is the past 4 quarter EPS growth of s . Standard errors are clustered at the quarter-by-industry level, where we cluster at the level of 2-digit SIC code. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

C Accounting for factor investing in news sensitivity

In this section, we decompose the persistence of our measure into one that can be attributed to a particular investment style (momentum, size, and value), and one that is more idiosyncratic to the investor. Holding fixed a given period $[t, t+h]$, we run the following pooled regression, for $0 \leq k < h$, across all investors and quarters:

$$NS_{i,t+k} = \phi_{t+k} + \widehat{NS}_{i,t \rightarrow t+h} + \varepsilon_{j,t},$$

with ϕ_{t+k} the quarter fixed-effect and $\widehat{NS}_{i,t \rightarrow t+h}$ capturing investor i 's NS during $[t, t+h]$. We also measure investor i 's average exposure to value, momentum, and size factors over the same period:

$$\bar{m}_{i,t \rightarrow t+h}^{style} = \frac{1}{h} \sum_{k=0}^{h-1} \frac{\sum_s H_{i,s,t+k} \cdot style_{s,t+k}}{\sum_s H_{i,s,t+k}}, \quad style \in \{Val, Mom, Size\},$$

where $H_{i,s,t}$ is the holdings of asset s by investor i in quarter t and $style_{s,t}$ is the style quintile of stock s in quarter t .

We run a cross-sectional regression of $\widehat{NS}_{i,t \rightarrow t+h}$, investor i 's NS in $[t, t+h]$, on $\bar{m}_{i,t \rightarrow t+h}^{style}$ over all investors in the period:

$$\widehat{NS}_{i,t \rightarrow t+h} = \sum_{s \in \{Val, Mom, Size\}} \beta_s \bar{m}_{j,t \rightarrow t+h}^s + \widehat{NS}_{i,t \rightarrow t+h}^{residual}$$

The fitted value of the regression, $\widehat{NS}^{style} = \sum_{s \in \{Val, Mom, Size\}} \beta_s \bar{m}_{j,t \rightarrow t+h}^s$, can be interpreted as the investor's NS attributable to her loadings on the three factors. The residual of the regression, $\widehat{NS}_{i,t \rightarrow t+h}^{residual}$, is then the residual component of NS. We can then decompose the persistence in NS by those attributable to style (\widehat{NS}^{style}) and the residual ($\widehat{NS}_{i,t \rightarrow t+h}^{residual}$):

$$\begin{aligned} \text{Cov}(\widehat{NS}_{i,t \rightarrow t+h}, \widehat{NS}_{i,t+h \rightarrow t+2h}) &= \underbrace{\text{Cov}(\widehat{NS}_{i,t \rightarrow t+h}^{style}, \widehat{NS}_{i,t+h \rightarrow t+2h}^{style})}_{\text{Persistence driven by style}} + \\ &\quad \underbrace{\text{Cov}(\widehat{NS}_{i,t \rightarrow t+h}^{residual}, \widehat{NS}_{i,t+h \rightarrow t+2h}^{residual})}_{\text{Residual persistence}} + \text{cross terms.} \end{aligned}$$

Style Cov. Fraction	0.165
Residual Cov. Fraction	0.555

Table A6: Persistence of investor score FE

Table A6 shows the results of the decomposition. The factor loadings of the investor account for roughly a quarter of the persistence of our measure, with a substantial component of persistence not attributable to a fixed investment factor.

D Trading strategy in price run-ups

Trading strategy To ensure that our return predictability results have incremental predictive power beyond standard priced risk factors, we complement our baseline regression results with a portfolio-based predictability exercise. For each quarter t , we form an equal-weighted long-short portfolio based on stocks that have run up by more than 100% in the past 4 quarters, where we sort the stocks into quintiles of $\text{Gap}_{s,t}$. We take as the long leg the set of stocks in the highest quintile, and as the short leg the set of stocks in the lowest quintile. We then compute the subsequent 4-quarter returns of the portfolio, and measure its alpha with respect to size, value, and momentum factors, as constructed by [Fama and French \(1993\)](#) and [Carhart \(1997\)](#).

One potential concern regarding this standard approach is that the horizon of return predictability (4 quarters) can be longer than the rebalancing frequency of the factor portfolios.⁴⁹ In particular, this implies that the 4 quarter returns of a momentum portfolio is systematically different from the 4 quarter returns of a high momentum stock, the latter being a more reasonable benchmark to compare to the 4-quarter returns of our episode stocks. Thus, we also construct our own [Fama and French \(1993\)](#)-like portfolios that explicitly track the 4-quarter forecasting horizon. In particular, for each quarter t , we form long-short portfolios based off size, momentum, and value constructed in a similar way as the [Fama and French \(1993\)](#) and [Carhart \(1997\)](#) portfolios.

In particular, we follow the exact same conventions as [Fama and French \(1993\)](#) and [Carhart \(1997\)](#) in performing the characteristic-specific sorts. For value and size, these are based off double-sorted portfolios using the median of size (market capitalization) and the tercile of value (the market-to-book ratio). The long-short size portfolio is, on the long end, the equal-weighted average of the three portfolios that include above-median size stocks, while the short end is the equal-weighted average of the three portfolios that include below-median size stocks. The long-short value portfolio is, on the long end, the equal-weighted average of the three portfolios that include the upper tercile of value stocks, and on the short end, the equal-weighted average of the three portfolios that include the lower tercile of value stocks. Finally, the momentum portfolio is constructed in the same way as the value portfolio, but using portfolios that are double-sorted by the stock's size and previous twelve month returns (going back two months before t).

We then compute the four-quarter return of these portfolios, holding fixed the composition of the portfolios over the entire four quarters. We refer to these factors as 4Q-horizon factors.

Table [A7](#) shows the estimated alpha of our the long-short overreaction portfolio. Column (1) includes no factor portfolios, such that the constant measures the overall return on the portfolio. We see that the unadjusted annual portfolio return is -8.6% and highly statistically significant (based on Newey-West errors computed with a lag of four). Thus, consistent with the above regression-

⁴⁹While size and value are rebalanced every June, momentum is rebalanced every month.

	(1)	(2)	(3)	(4)
	Unadjusted	Beta-adjusted	FF factors	4Q-horizon factors
Alpha (%)	-0.0860** (0.0343)	-0.0791** (0.0340)	-0.0794** (0.0352)	-0.143** (0.0651)
Market return		-0.104 (0.255)	-0.131 (0.204)	0.0209 (0.245)
Size			-0.185 (0.274)	-0.0889 (0.160)
Value			-0.359 (0.258)	-0.0913 (0.348)
Momentum			0.116 (0.172)	0.282 (0.302)
Observations	119	119	119	119

Table A7: Alpha of long-short extrapolative investor portfolios of stock run-up episodes

based predictability results, stocks with higher overreactive holdings experience significantly lower returns over the next four quarters. Columns (2) and (3) show that our results remain similar controlling for market returns and the [Fama and French \(1993\)](#) and [Carhart \(1997\)](#) factors. Finally, Column (4) shows the alpha with respect to the 4Q-horizon factors that match our forecasting horizon, which shows an even stronger underperformance of overreactive stocks. To summarize, our measure can explain a substantial fraction of the variation in the returns of stock-level run-ups, with greater overreactive holdings associated with greater long-run reversals and heightened crash risk.

E Overreaction vs private information: constructing two portfolios

In this section, we describe in detail the construction of the informed investor measure and the associated stock-level measure $\text{Gap}^{P\&L}$.

Investor-level and asset-level measure We denote $\text{P\&L}_{raw,i,t}$ of investor i in quarter t as her quarter t portfolio characteristics-benchmarked returns, defined as:

$$\text{P\&L}_{raw,i,t} = \frac{R_{s,t,t+1}^{\text{char adj}} \cdot H_{i,s,t}}{\sum_{s \in S_{i,t}} H_{i,s,t}},$$

where $R_{s,t,t+1}^{\text{char adj}}$ is the characteristics-adjusted return (Daniel et al., 1997) of stock s in the subsequent quarter and $H_{i,s,t}$ is the holdings of investor i of stock s in quarter t . In exactly the same procedure as investor news sensitivity, we process the raw P&L measure by computing the 8-quarter lagged moving average and taking the percentile rank among all investors. We then aggregate $\text{P\&L}_{i,t}$ in exactly the same manner as Equation (25) to construct at the stock by quarter level the holdings gap measure based on P&L, $\text{Gap}_{s,t}^{P\&L}$.

Orthogonalizing and comparing the two measures Given that our two holdings gap measures are correlated, we need to orthogonalize each measure and construct two portfolios that separate the predictability generated by overreaction, the “purged NS” portfolio, and that generated by private information, given by the alpha of the “purged P&L” portfolio.

For each quarter t , we form a double-sorted portfolio with respect to the quintiles of both measures, $\text{Gap}_{s,t}^{NS}$ and $\text{Gap}_{s,t}^{P\&L}$. We then form two long-short portfolios: one that takes the equal-weighted mean of the five portfolios with $\text{Gap}_{s,t}^{NS,quint} = 5$ subtracted by the mean return of the five portfolios with $\text{Gap}_{s,t}^{NS,quint} = 1$, and the other that does the same but with $\text{Gap}_{s,t}^{P\&L,quint}$. Let $r_t(NS^a, P\&L^b)$ be the 4-quarter returns of the portfolio with stocks in the a -th quintile of $\text{Gap}_{s,t}^{NS}$ and b -th quintile of $\text{Gap}_{s,t}^{P\&L}$. Then, the 4-quarter returns of the two portfolios are given by Equation (71),

$$\begin{aligned} r_t^{NS} &= \frac{1}{5} \left[\sum_{b=1}^5 r_t(NS^5, P\&L^b) - r_t(NS^1, P\&L^b) \right] \\ r_t^{P\&L} &= \frac{1}{5} \left[\sum_{a=1}^5 r_t(NS^a, P\&L^5) - r_t(NS^a, P\&L^1) \right]. \end{aligned} \tag{71}$$

Intuitively, r_t^{NS} (the “purged NS portfolio”) holds fixed the level of $\text{Gap}_{s,t}^{P\&L}$ and looks at the variation in $\text{Gap}_{s,t}^{NS}$, thus isolating the predictability generated by overreaction. Similarly, $r_t^{P\&L}$ (the “purged P&L portfolio”) isolates the predictability generated by information frictions.

Analysis We perform the above portfolio formation for the stock-quarter pairs (s,t) where s is experiencing a stock-level run-up (i.e. s has greater than 100% returns from quarters $t - 3$ to t), as well as the residual stock-quarter pairs (s,t) , which we call “non-episodes.” Table A2 shows the performance of the two portfolios during run-up episodes and out of run-up episodes. Columns (1), (2), (5), and (6) show the performance of the two portfolios in run-ups, while Columns (3), (4), (7) and (8) show the performance of the portfolios out of run-ups. Table A3 reports analogous results where we proxy investor informedness by her Buy-Performance measure, as defined in Equation (27), which tracks the benchmarked returns of stocks bought by the investor. In all of our specifications, while the underperformance of high overreactive holdings gap stocks is concentrated in run-ups, stocks with a high informed-investor holdings gap unconditionally outperform.

F Out-of-sample predictability analysis

In this section, we present an out-of-sample predictability version of the analysis in Section 5. Instead of comparing the adjusted R^2 of the regression in Equation (33), we instead use the out-of-sample R^2 defined in (Campbell and Thompson, 2008).

To see the explanatory power of our variables out of sample, we require a minimum training window of 10-years, starting from 1990 in our sample. Then, for each quarter t , we train a model predicting returns based on our set of variables in Section 5.2. In other words, holding the set of predictive variables Z fixed, we train our model using data up to quarter $t - 1$:

$$r_{s,t'+1,t'+2} = \alpha_{t-1} + \gamma_{t-1} \cdot Z_{s,t'-1}, \quad (72)$$

where $t' \leq t - 3$. Using the trained parameters $(\hat{\alpha}_{t-1}, \hat{\gamma}_{t-1})$, we then predict

$$\hat{r}_{s,t+1,t+2} = \hat{\alpha}_{t-1} + \hat{\gamma}_{t-1} \cdot Z_{s,t}, \quad (73)$$

for all stock s with positive earnings announcement in quarter t . Then, we compute the out-of-sample mean square error of the model associated with Z , $MSE(Z)$, by averaging the squared errors $(r_{s,t+1,t+2} - \hat{r}_{s,t+1,t+2})^2$ across s and t in our sample.

Following Campbell and Thompson (2008), we compare this to the mean square error using past average returns, which we denote as $MSE^{default}$. The out-of-sample R^2 of the model using the set of variables Z is then given by:

$$OOSR^2(Z) = 1 - \frac{MSE(Z)}{MSE^{default}} \quad (74)$$

	Ret	Gap	Industry	Sales	EPS	3 chars	3 chars +Gap
R^2 (BPs)	0.092	0.137	0.099	0.123	0.104	0.146	0.189
OOS R^2 (BPs)	-1.940	0.075	0.017	-0.016	-0.004	-0.002	0.042

Table A8: Effect of non-price characteristics on R^2 of two-quarter return predictability regressions

Table A8 shows the resulting out-of-sample R^2 of each model in Section 5.2, along with the in-sample adjusted R^2 computed for the main text. Consistent with the existence of significant in-sample bias, the out of sample R^2 is significantly lower than the in-sample R^2 of our variables. Regardless, the holdings gap measure has a significant improvement to the out-of-sample predictability, and outperforms the models only using the direct non-price information. Of the non-price information, only industry is able to outperform the default model.

G Late reaction or overreaction

In this section, we consider alternative measures of investor news sensitivity than the one defined in Equation (23). In our main specification, we had defined an investor's news sensitivity based on her tendency to purchase stocks in quarter t that have experienced positive news throughout quarters $t - 3$ to t . This combines two reactions: delayed reaction to positive news in quarters $t - 3$ to $t - 1$, as well as potential (contemporaneous) overreaction to positive news in quarter t .

Theoretical motivation In our full model, Corollary 1 implies a tight link between delayed reaction to past news and overreaction to contemporaneous news. In a setting where a single innovation in fundamentals V is continuously revealed over time (empirically, this corresponds to an innovation that is revealed through repeated positive news), oversensitive investors not only overreact contemporaneously, but do so at a later stage of the price increase process, effectively also reacting in a delayed manner to past information. Given the tight theoretical and intuitive link between contemporaneous overreaction and delayed reaction to news, our main measure combines the two forces.

Two alternative measures In this section, we explore separating out these two characteristics, and investigate which characteristic drives our core empirical results, we define two close alternatives of investor news sensitivity.

$$\begin{aligned}
 N_{s,t}^{lagged} &= \sum_{h=1}^4 \text{Announcement Returns}_{s,t-h} \\
 N_{s,t}^{contemp} &= \text{Announcement Returns}_{s,t}
 \end{aligned} \tag{75}$$

Then, one can construct in an analogous manner to Equation (21)

$$\begin{aligned}
 NS_{raw,i,t}^{lagged} &= \frac{\sum N_{s,t}^{lagged} \cdot W_{i,s,t}}{\sum W_{i,s,t}} \\
 NS_{raw,i,t}^{contemp} &= \frac{\sum N_{s,t}^{contemp} \cdot W_{i,s,t}}{\sum W_{i,s,t}},
 \end{aligned} \tag{76}$$

where $W_{i,s,t} \geq 0$ is the amount purchased, in dollars, of stock s by investor i in quarter t .

$NS_{raw,i,t}^{lagged}$ measures an investor's tendency to purchase stocks that have had good news in the prior 4 quarters before quarter t , while $NS_{raw,i,t}^{contemp}$ captures an investor's tendency to purchase stocks that have had good news in the same quarter. One can then obtain the final news sensitivity measures as well as the asset-level holdings gap measure in exactly the same manner as described in Section 3.

	Next Year Returns			
	(1)	(2)	(3)	(4)
Past 3 Year Returns			0.04 (0.03)	
Past Returns	0.04** (0.02)	0.07*** (0.02)	0.05** (0.02)	0.07*** (0.03)
Gap		-0.05*** (0.01)	-0.04*** (0.01)	
Past Returns X Gap		-0.04** (0.02)	-0.05** (0.02)	
Gap (Buys)				-0.03** (0.01)
Past Returns X Gap (Buys)				-0.05** (0.03)
Observations	468,119	468,119	433,308	356,292
R ²	0.14	0.14	0.14	0.16

Table A9: Momentum predictability using lagged news sensitivity

Notes: Table A9 reports the momentum of a stock conditional on the level of the holdings gap measure, given by Equation (27). Past returns of the stock is given by the annualized average monthly returns of the stock for the past 12 months, excluding the recent month (Carhart, 1997). Gap is the percentile rank of the holdings gap measure using lagged news sensitivity, as defined by Equation (75). Next year returns is the log returns of stock s from quarters $t + 1$ to $t + 4$ inclusive. The observation is at the stock by quarter level, consisting of all stock-quarter pairs with non-missing values of Gap with at least 6 months of returns as of quarter t . Column (1) shows the baseline return predictability without conditioning on the holdings gap. Column (2) shows our main specification, which indicates a lower momentum coefficient for stocks with higher holdings gap. Columns (3) controls for the 3 year returns of the stock. Column (4) replaces the holdings gap measure by the average news sensitivity of investors buying the stock between quarters $t - 3$ to t , standardized to its percentile rank across all stocks in quarter t . Standard errors are two-way clustered at the stock by quarter level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

	Next Year Returns			
	(1)	(2)	(3)	(4)
Past 3 Year Returns			0.04 (0.03)	
Past Returns	0.04** (0.02)	0.08*** (0.02)	0.06*** (0.02)	0.08*** (0.02)
Gap		-0.03*** (0.01)	-0.02** (0.01)	
Past Returns X Gap		-0.06*** (0.02)	-0.06*** (0.02)	
Gap (Buys)				-0.02** (0.01)
Past Returns X Gap (Buys)				-0.08*** (0.02)
Observations	471,146	471,146	436,006	358,148
R ²	0.14	0.14	0.14	0.16

Table A10: Momentum predictability using contemporaneous news sensitivity

Notes: Table A10 reports the momentum of a stock conditional on the level of the holdings gap measure, given by Equation (27). Past returns of the stock is given by the annualized average monthly returns of the stock for the past 12 months, excluding the recent month (Carhart, 1997). Gap is the percentile rank of the holdings gap measure using contemporaneous news sensitivity, as defined by Equation (75). Next year returns is the log returns of stock s from quarters $t + 1$ to $t + 4$ inclusive. The observation is at the stock by quarter level, consisting of all stock-quarter pairs with non-missing values of Gap with at least 6 months of returns as of quarter t . Column (1) shows the baseline return predictability without conditioning on the holdings gap. Column (2) shows our main specification, which indicates a lower momentum coefficient for stocks with higher holdings gap. Columns (3) controls for the 3 year returns of the stock. Column (4) replaces the holdings gap measure by the average news sensitivity of investors buying the stock between quarters $t - 3$ to t , standardized to its percentile rank across all stocks in quarter t . Standard errors are two-way clustered at the stock by quarter level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Tables A9 and A10 show that our main results of momentum predictability remain robust to our choice of the news sensitivity measure. Both versions of the holdings gap measure of overreaction can modulate the degree of momentum, even after controlling for longer horizon past returns.

	Next Year Returns		
	(1)	(2)	(3)
Past Returns	0.07*** (0.02)	0.08*** (0.02)	0.08*** (0.02)
Lagged Gap	-0.05*** (0.01)		-0.05*** (0.01)
Past Returns X Lagged Gap	-0.04** (0.02)		-0.01 (0.02)
Contemp. Gap		-0.03*** (0.01)	0.002 (0.01)
Past Returns X Contemp. Gap		-0.06*** (0.02)	-0.06*** (0.02)
Observations	468,105	468,105	468,105
R ²	0.14	0.14	0.14

Table A11: Horse race between our two news sensitivity measures

Notes: Table A11 reports the momentum of a stock conditional on the level of the holdings gap measure, given by Equation (27). Column (1) uses the lagged news sensitivity and Column (2) uses the contemporaneous news sensitivity to construct the holdings gap. In Column (3), one includes both versions of the holdings gap in the predictive regression.

Table A11 compares the degree of predictability of the two measures. The interaction term, i.e. the gap in momentum coefficients between the least and most overreactive stocks, is larger for the holdings gap measure built based on contemporaneous response to news. Column (3) shows that when one predicts future returns using both measures, most of the predictive power goes to the interaction term using the contemporaneous holdings gap measure. However, given the high correlation between the two gap measures, as well as the low frequency of both our holdings and news measures, our ability to separate the relative contribution of delayed and contemporaneous overreaction remains limited.

Relationship with post-earnings-announcement-drift We document in Section 4.3 that over-sensitive investors, who tend to buy stocks that have had high earnings returns in the same quarter, have lower benchmarked returns on their purchases. On the other hand, the evidence on post earnings announcement drift (Bernard and Thomas, 1989) documents the outperformance of buying stocks that have had positive earnings surprises. To reconcile these findings, note that most of the earnings announcement drift returns accrue in a horizon of 10 days to a month after the

event, while our “contemporaneous” purchases occur over a lower frequency of a quarter. Hence, our oversensitive agents need not be agents that respond immediately following the earnings announcement, but rather enter later on and earn lower returns. The fact that our results hold using lagged news sensitivity is consistent with the finding that investors with high news sensitivity tend to also respond later to older news.

H Investor response to negative news

In this section, we analyze the predictability of our holdings gap measure focused on cases of negative news. We document significant asymmetry in our results compared to the case of positive news, and sketch an extension of the theory that incorporates short-sales constraints that can account for this asymmetry.

Negative stock episodes First, we perform an analogous exercise of episodes of extreme negative returns, where we define a negative episode as a period in which a stock experiences lower than -60% returns over 4 quarters, which corresponds roughly to reverse percentile of stocks experiencing greater than 100% returns.

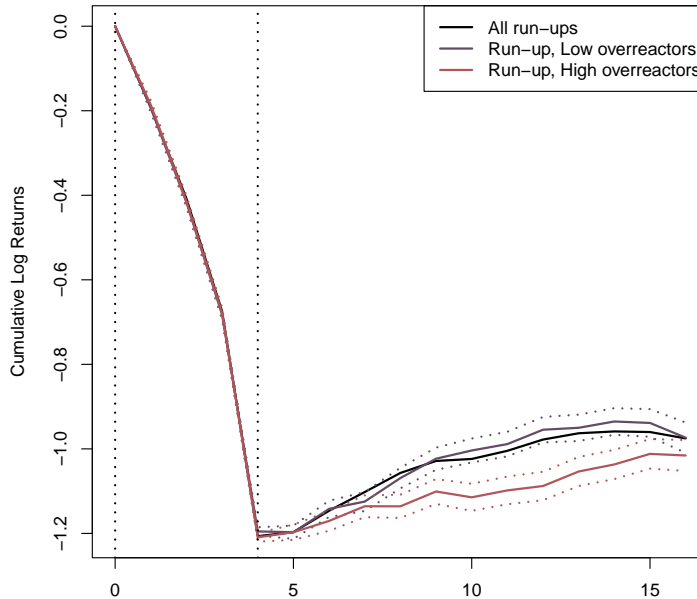


Figure A6: Negative episodes: holdings gap predictability

Notes: Figure A6 plots the cumulative log returns of the stock-level negative episodes, which is defined by a stock experiencing lower than -60% returns over 4 quarters, which correspond to quarters 1 through 4 in the figure. Quarters 5 through 16 correspond to the subsequent 3-year returns. We winsorize our episodes by their holdings gap in quarter 4 at the 10% level. The black curve plots the average across all episodes, while the the gray and red curves correspond to the bottom and top quintiles of the holdings gap at the end of the selection period (quarter 4). The dotted lines represent one standard-error intervals around the cumulative returns of high and low holdings gap episodes.

Figure A6 shows the cumulative returns of our negative episodes, sorted by the level of the holdings gap. Unlike the case for positive episodes, there is no significant difference in the 3-year cumulative returns between episodes sorted by our holdings gap measure, although there are

reversals on average, or positive returns following the formation period. However, one should be cautious in interpreting this reversal, as this can be generated by a survivorship bias, as we condition on stocks continuing to be traded for 12 quarters after the run-down.

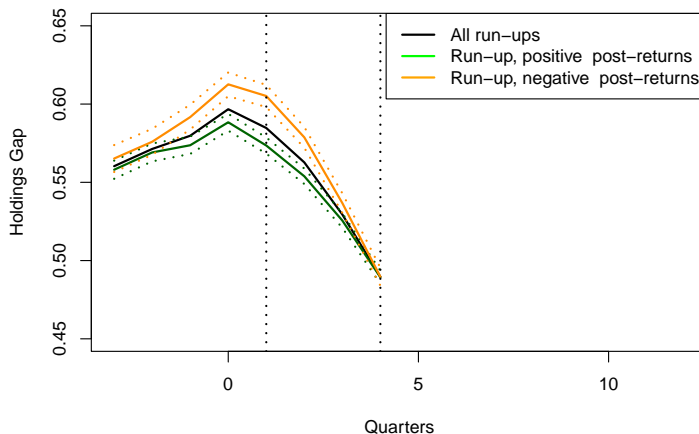


Figure A7: Dynamics of holdings gap, negative episodes

Notes: Figure A7 shows the standardized holdings gap of a stock in negative episodes, which is defined to be a stock experiencing lower than -60% returns over 4 quarters. The standardized holdings gap is given by the percentile rank across each quarter of the holdings-gap measure of the stock, as defined in Equation (25). Quarters -3 to 0 correspond to 4 quarters prior to the episode, and quarters 1 through 4 correspond to the crash period. The black curve plots the average level of the measure across all run-up episodes, while the orange and green curves plot the measure across run-up episodes with negative and positive future 3-year returns following the run-up. The dotted lines represent one standard-error intervals around the mean.

Figure A7 shows that through the course of the negative episode, oversensitive investors reduce their holdings, resulting in a drop in the holdings gap measure across the episodes. On the other hand, consistent with the lack of cross-sectional predictability, we find that the level of the holdings gap converges to a similar value for episodes with both positive and negative future 3 year returns.

Short-sales constraints and model extension To summarize, in the case of negative extreme returns, we also find evidence consistent with overall overreaction and reversals, but we do not find evidence that of the holdings gap measure’s ability to cross-sectionally predict future returns. One can rationalize the asymmetry in our findings by incorporating short-sales constraints into our model. As documented in Nagel (2005), short-sales constraints can introduce asymmetric effects in how rational arbitrageurs can correct mispricing. In our setting, this also plays a role in how oversensitive investors respond to news.

To sketch out this insight, consider a simple modification of our model, where there is a continuum of overreactive agents, which constitute χ of investors, and the remaining $1 - \chi$ consisting of

rational investors. We now assume a uniformly binding short-sales constraint: rational and over-sensitive investors cannot hold short positions. Also, instead of inattentive investors yielding the shares, we assume a fixed supply of assets, $S > 0$. Finally, we assume that the degree of investor oversensitivity is heterogeneous across oversensitive investors:

$$D^{os}(n, p) = \frac{1}{A} ((1 + \Phi(c) \cdot NS_i) \tau_\epsilon n - (\tau_V + \tau_\epsilon) p), \quad (77)$$

where $NS_i \geq 0$ varies across investor i . Then, in response to positive news, as one increases $\Phi(c)$, one can show that the holdings gap measure continuously adjusts upwards, even after rational agents hold none of the asset. In that case, our holdings empirical holdings gap measure then corresponds to $\text{Gap} \equiv \int D^{os}(n, p) \cdot NS_i / S$. Intuitively, this is driven by even more oversensitive agents buying the asset from less oversensitive agents. This implies that our measure is cross-sectionally predictive in the case of positive news despite the short-sales constraint.

On the other hand, for sufficiently negative news, once all oversensitive investors hold none of the asset, the asset is entirely in the hands of the rational investors. Thus, even as one increases $\Phi(c)$, there is no change in the holdings gap measure and thus no cross-sectional predictability, although there is an unconditional reversal, which is given by rational investors being compensated for holding the positive supply of the assets.

Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
Run-up Holdings Gap	10054	0.613	0.243	0.126	0.413	0.829	0.947
Run-down Holdings Gap	4625	0.49	0.231	0.095	0.29	0.689	0.897

Table A12: Distribution of holdings gap

The above mechanism makes a further prediction beyond the asymmetry of predictability: given that the lack of predictability is driven by oversensitive investors hitting the short sales constraint, this implies that the dispersion in the holdings gap should be lower for negative episodes. Table A12 compares the distribution of the holdings gap through our run-up and negative (“run-down”) episodes, and find indeed that not only is the holdings gap lower for the run-down episodes, but also that the standard deviation is lower than that of the run-up episodes.⁵⁰ However, the difference in the dispersion is small relative to the total dispersion, which can be driven by the general noise in measuring investor news sensitivity.

Negative announcements We also perform the same analysis as in Section 5, where we analyze the dynamics of the holdings gap around negative earnings announcements. Figure A5 shows the

⁵⁰There is a substantial difference in the sample size as we restrict episodes for stocks that continue to be traded for 12 quarters following the run-down.

exact reverse pattern as shown in Figure 5. Consistent with our theory, on average, oversensitive investors sell the asset in response to negative news. However, also consistent with our findings that our measure is limited in its ability to predict cross-sectional variation in response to negative news, Figure A8 shows that our measure does not increase the explanatory power of predicting post-announcement returns conditioned on negative news.

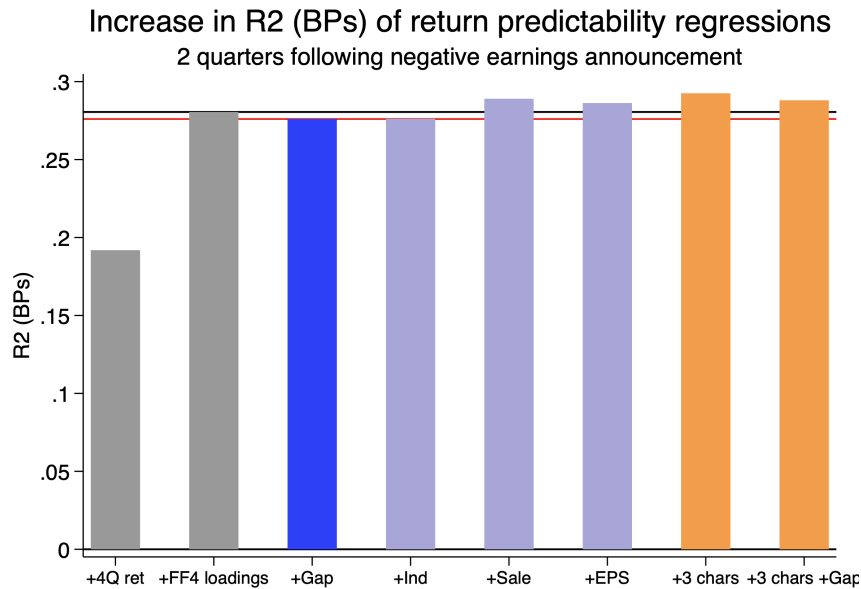


Figure A8: Increase in return prediction R^2 , by non-price measure, negative news

Notes: Figure A8 shows the explanatory power of our holdings gap measure in predicting the 2-quarter post-announcement returns following negative earnings announcements, as specified in Equation (33). The first column shows (in basis points) the adjusted R^2 of using just the past 4 quarter returns including the announcement quarter return. The second column shows the adjusted R^2 of adding the stock's four FF4 loadings in addition to the returns. The third column shows the R^2 after adding the 4-quarter path of the holdings gap measure in addition to the returns and FF4 loadings. The fourth, fifth, and sixth columns show the the R^2 after separately adding, respectively, industry winners, sales, and EPS growth in addition to returns and FF4 loadings. The seventh column shows the R^2 of adding all three of these variables in addition to returns and FF4 loadings. The final column shows the total explanatory power of adding in all three non-price variables and the holdings gap measure in addition to returns and FF4 loadings.

I Explanatory power, sorted portfolio exercise

In this section, we explore another way of testing the relative power of our measure compared to other non-price information measure in predicting investor overreaction. In Section 5 and Appendix F, we test this at a higher frequency, comparing the gain in R^2 predicting post-announcement 2-quarter returns, either in-sample or out of sample. We show that our measure has higher predictive power than the other non-price information, although the gain in explanatory power is quite small, given the difficulty in predicting individual stock-level returns with significant explanatory power.

In this section, we take an alternative approach, and instead compare the difference in the future reversals of momentum portfolios when conditioned on each variable, as analyzed in Section 4.1. For each non-price variable Z , we compare two quantities: the explanatory power of our holdings gap measure conditional on Z and the explanatory power of Z conditional on our holdings gap measure. If our holdings gap aggregates different non-price drivers of overreaction, then the additional explanatory power of Z should be lower than the additional explanatory power of our holdings gap variable.⁵¹

To test the above hypothesis, we compute the returns of two triple-sorted portfolios. The portfolios are formed by triple sorting stocks in each month t into (a) the decile of cumulative log returns over the past 12 months excluding the most recent month $t - 1$ (Carhart, 1997), (b) the quintile of the holdings gap measure in the quarter before month t 's quarter, which is defined following Equation (25), and (c) the quintile of the stock's non-price characteristic Z .

Denote as Gap-purged high Z momentum portfolio as going long winners (stocks in the top return decile) with high Z and going short losers with low Z , holding fixed the quintile of the holdings gap:

$$r^{Z,high} = \frac{1}{5} \sum_{g=1}^5 r(10, g, 5) - r(1, g, 1), \quad (78)$$

where $r(d, g, z)$ is the return of stocks in the d -th decile of past returns, g -th quintile of the holdings gap, and z -th quintile of the non-price characteristic Z . Similarly, denote Gap-purged low Z momentum portfolio as:

$$r^{Z,low} = \frac{1}{5} \sum_{g=1}^5 r(10, g, 1) - r(1, g, 5). \quad (79)$$

Conversely, we denote as the Z -purged high Gap momentum portfolio as going long winners (stocks in the top return decile) with high holdings gap and going short losers with low holdings

⁵¹If our measure was truly a sufficient statistic, then there should be no additional explanatory power of Z , but given that our measure is noisy, one would expect Z to have additional explanatory power conditional on the holdings gap.

gap, holding fixed the quintile of Z

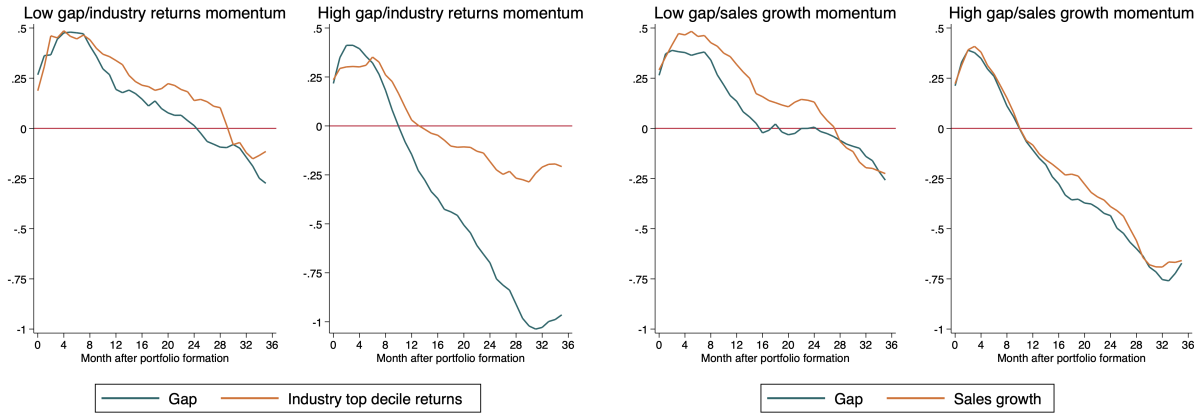
$$r^{Gap,high} = \frac{1}{5} \sum_{g=1}^5 r(10, 5, z) - r(1, 1, z), \quad (80)$$

Similarly, we define the Z -purged low holdings-gap momentum as:

$$r^{Gap,low} = \frac{1}{5} \sum_{g=1}^5 r(10, 1, z) - r(1, 5, z). \quad (81)$$

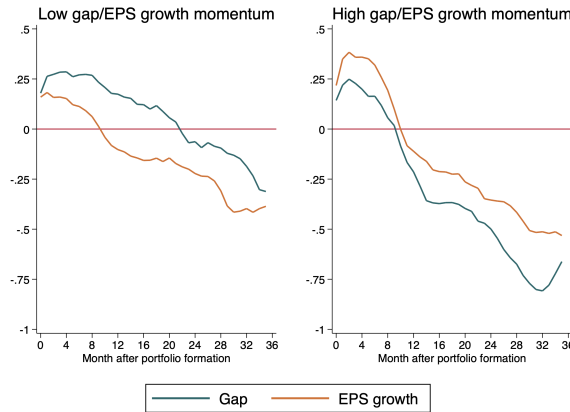
We compare the Sharpe ratios of these four portfolios in Figure A9. Observations are at the monthly level. In the left panel, we plot the Sharpe ratio of the low Gap-purged Z portfolio (in orange) and the low Z -purged Gap portfolio (in green). In the right panel, we plot the Sharpe ratio of the high Gap-purged Z portfolio (in orange) and the high Z -purged Gap portfolio. We compute Sharpe ratios by first regressing, for each horizon, the unweighted mean returns of each sorted portfolio on the returns of the market portfolio (the CRSP value-weighted portfolio), SMB (size), and HML (value) (Fama and French, 1993) over the same horizon. We then take the mean of these adjusted returns over the cumulative three-month Treasury returns over the given horizon, and divide by the standard deviation of this excess return series.

As the three panels in Figure A9 show, the high holdings gap momentum portfolio predicts the strongest long-run reversals compared to all three non-price information characteristics: in other words, our holdings gap measure has greater residual prediction over long-run reversals when controlling for other non-price variables than vice versa. Of all the non-price characteristics, our measure has a similar degree of power to the sales-growth measure, with long-run returns having a Sharpe ratio of roughly -0.75 over 3 years (or in other words, betting against high momentum overreactive stocks will have a Sharpe ratio of 0.75 over 3 years).



(a) Holdings gap vs industry winners

(b) Holdings gap vs sales growth



(c) Holdings gap vs eps growth

Figure A9: Explanatory power: long-run reversals

Notes: Figure A9 plots the Sharpe ratios of (1) low Gap-purged Z portfolios vs low Z-purged Gap portfolios, and (2) high Gap-purged Z portfolios vs high Z-purged Gap portfolios, where the four portfolios are defined by Equations (79), (81), (78), and (80) respectively. Observations are at the monthly level. We compute Sharpe ratios by first regressing, for each horizon, the unweighted mean returns of each sorted portfolio on the returns of the market portfolio (the CRSP value-weighted portfolio), SMB (size), and HML (value) (Fama and French, 1993) over the same horizon. We then take the mean of these adjusted returns over the cumulative three-month Treasury returns over the given horizon, and divide by the standard deviation of this excess return series.