Crowding and Tail Risk in Momentum Returns

Pedro Barroso, Roger M. Edelen, and Paul Karehnke*

October 17, 2020

Abstract

Several theoretical studies suggest that coordination problems can cause arbitrageur crowding to push asset prices beyond fundamental value as investors feedback trade on each others' demands. Using this logic we develop a crowding model for momentum returns that predicts tail risk when arbitrageurs ignore feedback effects. However, crowding does not generate tail risk when arbitrageurs rationally condition on feedback. Consistent with rational demands, our empirical analysis generally finds a negative relation between crowding proxies constructed from institutional holdings and expected crash risk. Thus our analysis casts both theoretical and empirical doubt on crowding as a stand-alone source of tail risk.

*Barroso, pedro.barroso@ucp.pt, Universidade Católica Portuguesa Católica Lisbon School of Business and Economics; Edelen, edelenr@vt.edu, Virginia Tech Pamplin College of Business; and Karehnke, pkarehnke@escp.eu, ESCP Business School (corresponding author). We are grateful to Semyon Malamud for comments that were particularly helpful to the paper's development. We are also grateful for helpful comments and suggestions from an anonymous referee, Jeff Busse, Jennifer Conrad (the editor), John Griffin, Hening Liu, Steve Satchell, Rick Sias, Neal Stoughton, Xuemin Yan, and seminar participants at the University of Arizona, California Riverside, New South Wales, Newcastle, Paris-Dauphine, Technology at Sydney, Virginia Tech, EDHEC, INSEAD, Tennessee's Smokey Mountain Finance Conference, The Annual IDC/Herzliya Conference and Annual Quantitative Trading Symposium, the Financial Risks International Forum, Manchester's Hedge Fund Conference, the Portuguese Finance Network Conference, Rotterdam's Professional Asset Management Conference, and the TCU Finance conference. Pedro Barroso acknowledges the support from FCT - Portuguese Foundation of Science and Technology for the project UID/GES/00407/2019. The authors are solely responsible for errors.

I. Introduction

What is the role of crowding in generating tail risk in investment strategies? A growing theoretical literature argues that financially constrained arbitrageurs can generate tail risk in asset prices by way of fire sales—forced exit from an otherwise profitable trade. The constraints underlying these theories derive from a variety of sources, including delegated portfolio management (Shleifer and Vishny, 1997), segmented margin accounts (Gromb and Vayanos, 2002), self-imposed loss limits (Morris and Shin, 2004), and illiquidity in funding markets (Brunnermeier and Pedersen, 2009). Crowding potentially plays an interactive role in these theories by setting the conditions under which a forced reversal of investment positions generates extreme price impact and transient undervaluation. Indeed, Brown, Howard, and Lundblad (2019) provide empirical support for the hypothesis that concentrated positions amplify tail risk in times of market distress.

The aim of this study is rather different. We explore whether crowding *per se* generates tail risk, without conditioning on market distress. As developed below, such an analysis requires a setting in which arbitrageurs follow a strategy based on prices and incomplete information about peer actions; and a setting where crowding induced tail risk is plausible. The momentum strategy fits that setting well. It uses no exogenous signal of value that could be used to coordinate arbitrage, and it is a statistically robust, widely-documented, and popular investment strategy that has well-known crash tendencies (Barroso and Santa-Clara, 2015; Daniel and Moskowitz, 2016). Related studies have shown that incomplete information about the actions of peers (i.e., the crowd) can induce arbitrageurs to amplify bubbles (Abreu and Brunnermeier, 2003) and push prices away from fundamental value (Stein, 2009). We contribute to this literature by formally developing a theory of crowding and momentum tail risk, and conducting an empirical analysis of that theory using direct measures of institutional crowding applied to 13f data.

Our theoretical setting uses standard assumptions for preferences and the structure of information to highlight the central role of investors' beliefs. We first demonstrate how unobserved crowding can explain momentum tail risk using the logic of Stein (2009). If the demands of peer momentum investors are indistinguishable from the informed trade that the strategy seeks to exploit, then unanticipated entry by those peers can create a feedback component to each momentum investor's demand. Depending on how beliefs are structured, this feedback component can spiral into grossly inflated aggregate momentum demands, driving the market price of momentum stocks far beyond their fundamental value. The extreme valuation reverses with the arrival of fundamental information, yielding large negative momentum returns. This motivates a crowding hypothesis for momentum crashes.

We then demonstrate that this crowding hypothesis requires myopic momentum investors. If momentum investors ignore feedback effects then, under standard assumptions, their demands relate linearly to market prices. However, linear demands generate precisely the feedback effects that momentum investors ignore in forming linear demands. Because those feedback effects are nonlinear, rationality requires a nonlinear mapping from market prices to the inferred fundamental values. This in turn implies nonlinear demands. In our rational solution to market equilibrium, we solve for these demands using conditional expectations derived from a fixed-point equilibrium condition. We find that the resulting demands eliminate crowding-induced momentum tail risk.

This sets up the primary test in our empirical analysis. If we can relate crowd size to tail risk in momentum returns, then we ought to also be able to find feedback effects in momentum investors' demands. Conversely, if we are unable to link crowding to negative skewness in momentum returns, then we should find that momentum investors exit from the momentum strategy when feedback risks are high. This unwinding of feedback effects via risk-managed demands is the main insight from the rational expectations equilibrium, which predicts that momentum investors negate crowding effects by reducing demands when feedback risks are high.

Our empirical analysis uses 13f data to construct proxies for momentum crowding by directly linking changes in institutional holdings to past returns, following the analysis of mutual fund momentum trading in Grinblatt, Titman, and Wermers (1995). However, our proxies incorporate persistence in trading to better distinguish investment strategy from spurious trade-return correlations. Also, we consider a variety of perspectives on identifying crowd size, including the number of momentum investing institutions, their size, and the aggregate intensity of their trade. We consider a perspective that focuses on the momentum factor mimicking portfolio and a perspective that focuses on individual momentum stocks. To our knowledge, this represents the most direct and comprehensive construction of proxies for institutional momentum investing in the literature. Several other studies (most notably Lou and Polk, 2013) use returns-based approaches to infer crowding. We provide evidence on the efficacy of returns-based procedures relative to our direct proxies based on institutions' trading strategy.

We find strong evidence that crowding predicts negatively mean momentum returns, which is consistent with theoretical prediction under all belief specifications considered. However, we find little evidence that (unanticipated) crowding predicts momentum tail risk. To be meaningful, tail risk implies negative skewness, elevated volatility, and excess kurtosis. Our proxies for momentum crowding generally relate negatively to all three,[1] often with statistical reliability. This result is not consistent with crowding providing a stand-alone explanation for momentum crashes, but it is consistent with our rational expectations model of momentum investing. In particular, our finding of reduced participation by institutional momentum investors when feedback risks are high supports the rational model's premise. Thus, we reject crowding as a stand-alone explanation for momentum tail risk.

A link between crowded trades and higher return moments for quantitative investment

[1] To be precise, crowding is associated with lower volatility, less excess kurtosis, and less negative skewness.

strategies such as momentum has been examined in the context of the 'quant meltdown' of 2007 (e.g., Khandani and Lo, 2007; Pedersen, 2009). It is important to note that these studies analyze the interactive effect of crowded strategies on funding liquidity risks; they do not argue that crowding per se led to the meltdown. This is a different model of tail risk (deriving from funding liquidity) than that considered here (deriving from feedback effects). Our analysis concludes that crowding does not provide a stand-alone explanation for momentum tail risk. We do not consider the role of funding liquidity, or how funding interacts with the extent of institutional participation in the momentum strategy.

Our analysis makes an important contribution to the literature on risk management in investment strategies (e.g., Barroso and Santa-Clara, 2015; Daniel and Moskowitz, 2016). Our theory provides intuition in the context of a rational expectations model, and our empirical evidence demonstrates how momentum institutional investors pull back from the strategy in toxic environments. Our assumptions are conventional. We use log normally distributed cash flows and constant relative risk aversion (CRRA) preferences with a second-order approximation in the demand optimization (as in Campbell and Viceira, 2002, and Peress, 2004)—effectively mean-variance preferences scaled to wealth. Our information structure is similar to Hong and Stein (1999) and Stein (2009), from which the crowding hypothesis arguably originates. Finally, our fixed-point methodology does not put restrictive conditions on the distribution of cash flows or the mapping of cash flow expectations into demands. Thus, it is robust to alternative specifications.

On the empirical side, several arguments speak in favor of our crowding measures. First, the measures draw on a large body of literature that shows that institutional investors are momentum traders (Lewellen, 2011; Edelen, Ince, and Kadlec, 2016). Second, we verify with transition matrices that our institution level measures capture meaningful momentum trading. Finally, our first main empirical results shows that our crowding proxies indeed strongly negatively relate to momentum returns—even after controlling for momentum's volatility (Barroso and Santa-Clara, 2015) and dynamic risk factors (Grundy and Martin, 2001).

We try to reconcile our results with studies that use returns-based measures of crowding. For the momentum gap measure proposed by Huang (2015), we show that orthogonalizing this variable to our crowding measures leaves the ability to predict momentum tail risk unaffected. This supports our second main empirical conclusion that momentum's crash risk cannot be attributed to crowding by institutional investors, and it opens the door to alternative explanations for momentum crashes.[2]

There is much related literature on the subject. We provide a detailed survey in Internet Appendix IA.A. Section II develops the model and Section III develops and analyzes its result using a simulation approach. Section IV presents the empirical analyses and Section V concludes the study.

II. Model

A. Setting

In practice momentum is a dynamic strategy that conditions on past (realized) returns. To focus on the effect of crowding, we collapse the analysis into a single call auction as in Stein (2009).[3] Thus, we model a two-period setting where the first period generates a formation-period return and the second period generates an evaluation-period return. In the first period informed investors observe a signal of differential value for good versus bad stocks, and in the second period

[2]See, for example, Daniel, Jagannathan, and Kim (2019) who argue that momentum crashes arise because of the higher effective leverage of past loser stocks (in the sense of Merton, 1974) following market downturns.

[3]Since investors can submit a demand schedule that conditions on the market clearing price at the time of the auction, they can effectively condition on past returns as of that time—even though demand schedules are submitted prior to the auction. Our simplified setting therefore retains the key feature of the momentum strategy.

the fundamental value of all stocks is revealed. As developed below, this setting gives rise to momentum returns similar to Stein (2009).

The setting at time 0 is symmetric in that all stocks have equal market value and all investors hold the same portfolio—an equal-weighted investment in all stocks. There are three types of stocks; good, bad, and neutral, but at time 0 no investor knows their identity. There are three types of investors: informed, momentum, and counterparty. During the formation period (time 0 to 1) informed investors observe the type of each stock, and a valuation signal δ described below. This signal indicates the differential payoff on good versus bad stocks. All good stocks have the same positive signal and all bad stocks have the same negative signal.

The fundamental value $P_{j,2}$ of each stock j is revealed to all investors at time 2 to be

$$\ln P_{j,2} = \ln P_{j,0} + \chi + \iota_j \frac{\delta + \epsilon}{2}, \tag{1}$$

where $P_{j,0}$ is the time-0 price that is normalized to 1 without loss of generality; χ is a normally distributed mean-zero innovation common to all stocks with variance σ_χ^2; ι_j identifies stock j's type (1 for good, -1 for bad, and 0 for neutral); and $\delta + \epsilon$ is the realized valuation spread on good minus bad stocks that consists of the signal δ and a normally distributed mean-zero disturbance ϵ with variance σ_ϵ^2. The random variables χ, δ, and ϵ are independent. χ generates market risk for investors, whereas $\delta + \epsilon$ generates a momentum return as developed below.

Informed investors see the vector of stock identifiers $\{\iota_j\}$ in the formation period, and their signal δ that is drawn from a lognormal distribution. Thus, δ (which can be thought of as the ex ante valuation spread on good stocks minus bad) is positive, whereas $\delta + \epsilon$ (the ex post spread) can be negative.

Informed investors seek to trade in the time 1 call auction. However, because of their signal homogeneity, their demands can be condensed to an equal-weighted portfolio that is long all good stocks and short all bad stocks. This defines the momentum portfolio. Momentum investors infer the portfolio from the differential price pressure of informed demands.

Both counterparty and momentum investors remain uninformed in the formation period. Thus, neither would be willing to counter demands in the time 1 auction if they were rational and knew that informed investors were active. To support market clearing we therefore define a counterparty investor type that irrationally fixates on historical data and ignores the existence of δ (effectively assuming $\delta = 0$).[4] As a result, these investors wrongly interpret informed investors' price pressure as noise, so they trade counter to that price pressure in the time 1 call auction.

Our third investor type, momentum investors, recognizes that the above two investor types generate equilibria with momentum opportunities if informed investors are risk averse. Thus, momentum investors formulate a model of market equilibrium in the time 1 call auction and use that model to try to capture momentum opportunities. In doing so they incorporate varying degrees of rationality into their model. They know the distribution of random variables.

Note that every investor trades a single basket—the momentum portfolio. Informed investors do so because of their homogeneous signal. Counterparty and momentum investors do so because they react to the pricing signal caused by (homogeneous) informed investor demands. We therefore frame our analysis in terms of the momentum portfolio as a single, composite asset.

Let f denote the formation-period log return on the momentum portfolio (to be solved for). The evaluation period log return is then $\delta + \epsilon - f$. Let $m \equiv E\left(\delta + \epsilon - f | \delta, f\right) = \delta - f$ denote the expected momentum return that informed investors leave on the table in the time 1 call auction. Only informed investors see m.

[4]Counterparty investors can also be thought of as contrarian traders anchoring on past prices. See, for example, Hvidkjaer (2006) for empirical evidence on contrarian trading and George and Hwang (2004) for evidence showing that traders likely anchor on past prices.

B. Demand Schedules

All investors have power utility preferences with relative risk aversion γ, choosing a time 1 demand schedule for the momentum portfolio to maximize[5]

$$E\left[u\left(K_2\right)\right] = E\left[\frac{K_2^{1-\gamma}}{1-\gamma}\right], \tag{2}$$

where K_2 denotes their time 2 capital. We use the second-order approximation approach of Campbell and Viceira (2002, Internet Appendix) to restrict investors' attention to the first two moments, yielding[6]

$$Demand = \frac{E_{\text{type}}\left[m+\epsilon\right]}{\gamma Var_{\text{type}}\left[m+\epsilon\right]} K_{\text{type},0}, \tag{3}$$

where $K_{\text{type},0}$ denotes the time 0 initial capital of the indicated investor type $\in$ {I, C, M}, and the details of the derivation are in Internet Appendix IA.B. The expectations can be simplified to $E_{\text{I}}\left[m+\epsilon\,|\delta, f\right] = \delta - f$, $E_{\text{C}}\left[m+\epsilon\,|f\right] = -f$, and $E_{\text{M}}\left[m+\epsilon\,|f\right] = \delta^E - f$. The variances are $Var_{\text{I}}\left[m+\epsilon\,|\delta, f\right] = \sigma_\epsilon^2$, $Var_{\text{C}}\left[m+\epsilon\,|f\right] = \sigma_\epsilon^2$, and $Var_{\text{M}}\left[m+\epsilon\,|f\right] = \delta^V + \sigma_\epsilon^2$. The placeholders $\delta^E = E_{\text{M}}\left[\delta\,|f\right]$ and $\delta^V = Var_{\text{M}}\left[\delta\,|f\right]$ will thereby be solved for below.

Momentum investors infer expected momentum returns, m, by way of the market clearing price. This inference is imperfect and potentially quite complicated since market equilibrium depends on unobservable capital allocations and momentum investors' own demands.[7]

[5]We assume that $\gamma > 1$ without loss of generality throughout the paper.

[6]The advantage of this scheme over CARA utility is to provide a natural role for investors' capital to frame the analysis of crowding. We acknowledge that the analysis would probably not differ substantially if we were to begin with CARA preferences.

[7]To this point our analysis largely parallels Stein (2009). Our analysis of varying degrees of rationality, including a fully rational fixed-point equilibrium, marks the point of departure.

C. Equilibrium

Summing demands across the three investor types and equating to supply (zero) gives the market clearing condition:

$$f = \frac{1}{D}\left(\delta k_{\mathrm{I}} + \frac{\delta^E}{1+\frac{\delta^V}{\sigma_\epsilon^2}} k_{\mathrm{M}}\right), \tag{4}$$

where $D = \left(1 - \frac{\delta^V}{\sigma_\epsilon^2+\delta^V} k_{\mathrm{M}}\right)$ and $k_{\mathrm{type}} = K_{\mathrm{type}}/\left(K_{\mathrm{C}} + K_{\mathrm{I}} + K_{\mathrm{M}}\right)$ indicates the fraction of capital from each investor type. The random variable k_{M} is the momentum-investor relative crowd size.[8] The initial capital allocations k_{I}, k_{C}, and k_{M} are randomly drawn from a symmetric Dirichlet distribution (discussed in more detail in Section III).

We consider four solutions to equation (4), beginning with a perfect-information base case and continuing through various assumptions on momentum investor rationality.

Known crowding. With known crowding (i.e., the realization of random variables k_{M} and k_{I} are observed prior to trading) momentum investors conjecture a linear equilibrium

$$f = \lambda\delta, \tag{5}$$

with $\lambda \equiv k_{\mathrm{I}} + k_{\mathrm{M}}$. Beliefs $\delta^E = \lambda^{-1} f$ and $\delta^V = 0$ lead to a self-fulfilling linear solution to equation (4). Hence, momentum investors' conjecture fully reveals the private signal of informed investors.

Myopic beliefs. This case assumes that momentum investors know the mean of the distribution of capital allocations, but they do not observe realizations; nor do they attempt to infer realizations from market prices. They trade as if capital allocations were at their unconditional

[8] Note that crowding is not the same as aggregate demand by momentum investors for the momentum portfolio, which follows from equation (3). This is an important distinction for the empirical section: Crowding uncertainty derives from not knowing how many peers are following the same strategy. To quote Stein (2009) (pg. 1530) "...each arbitrageur is uncertain about *how many* [emphasis not added] others will act."

mean values, conjecturing a linear equilibrium pricing relation

$$f = \lambda\delta, \qquad \text{where} \qquad \lambda \equiv Ek_{\mathrm{M}} + Ek_{\mathrm{I}}, \tag{6}$$

as in the known-crowding case. The market clears at

$$f = \lambda\left(\frac{k_{\mathrm{I}}}{\lambda - k_{\mathrm{M}}}\right)\delta = \lambda\left(\frac{k_{\mathrm{I}}}{Ek_{\mathrm{I}} - (k_{\mathrm{M}} - Ek_{\mathrm{M}})}\right)\delta. \tag{7}$$

The bracketed multiplier in equation (7) equals one when realized capital allocations happen to equal their expected values, but it generally adds a noise component to f with unanticipated crowding (as in Stein, 2009). As momentum investors react to this false momentum signal, their demands generate an asymmetric crowding-induced feedback effect. This distortion is rather immaterial when the crowd size is abnormally low, but an abnormally high crowd size can drive the price of the momentum portfolio far beyond fundamental value. The result is a momentum crash when fundamental value is revealed.[9]

Optimal linear beliefs. This is an intermediate stop on the road to rationality, considered in Stein (2009). As in the myopic case, momentum investors make no attempt to infer realized capital allocations from the market clearing price, despite the asymmetric (and occasionally extreme) impact that crowding can have on that price. That is, as the market clearing price f becomes more extreme, it becomes more likely that its source is crowding-induced feedback effects rather than information. When momentum investors restrict their strategy space to linear beliefs, $\delta^E = \lambda^{-1}f$, they ignore this characteristic of market equilibrium. This is the case considered here.

[9]There is an equilibrium at a finite *negative* value for f when $k_M > \lambda$. This corresponds to a reversal of identification on winners and losers, putting momentum investors on the wrong side of the trade (heavily buying losers and selling winners, with informed traders taking the other side and making enormous profits). We consider this case in Internet Appendix IA.C but note here that it too predicts large negative momentum returns.

Thus, momentum investors maintain a linear strategy, as in the myopic case, but they apply an ad hoc attenuation to λ^{-1} to maximize ex ante profits (conditional on linear beliefs). This yields an equilibrium that is structurally similar to equation (7), but with enough attenuation the denominator in the bracketed multiplier can be bounded away from zero. This prevents catastrophic losses from crowding-induced feedback effects.

Rational beliefs. To be rational, when beliefs δ^E and δ^V condition on f they must generate demands that clear the market at f. That is, they provide a fixed point solution. Optimal linear demands do not satisfy this condition.

Rationality is easily attained in the case of known crowding. Here we show that it is also attainable with unknown crowding. The only requirement is that momentum investors account for the asymmetric signal that f provides regarding the presence of crowding-induced feedback effects. The random variables in the system are δ, k_{I}, and k_{M}. We write their joint density as $g(\delta) \cdot h(k_{\mathrm{M}}, k_{\mathrm{I}})$, since δ is independent of k_{M} and k_{I} but the k's themselves are dependent (with k_{C} they sum to one and are nonnegative). Momentum investors compute δ^E and δ^V given this joint density, observation of f, and the definition of conditional expectations:

$$(8) \qquad \delta^E = \int_0^\infty \delta\, p(\delta|f)\, d\delta, \qquad \text{and} \qquad \delta^V = \int_0^\infty \left(\delta - \delta^E\right)^2 p(\delta|f)\, d\delta,$$

where, as shown in Internet Appendix IA.D,

$$(9) \qquad p(\delta|f) = \frac{\frac{g(\delta)}{\delta} \int_0^1 h\left(k_{\mathrm{M}}, \frac{1}{\delta}\left(fD - \frac{\delta^E}{1+(\delta^V/\sigma_\epsilon^2)} k_{\mathrm{M}}\right)\right) D dk_{\mathrm{M}}}{\int_0^\infty \frac{g(\delta)}{\delta} \int_0^1 h\left(k_{\mathrm{M}}, \frac{1}{\delta}\left(fD - \frac{\delta^E}{1+(\delta^V/\sigma_\epsilon^2)} k_{\mathrm{M}}\right)\right) D dk_{\mathrm{M}} d\delta}.$$

We use Matlab's FSolve function to jointly solve Eqs. (8) and density function (9) at a given value for f. We then repeat over a fine grid of plausible values for f to discretely approximate the mappings $f \to \delta^E$ and $f \to \delta^V$. We then interpolate to approximate beliefs at arbitrary values of

f.[10]

III. Simulations

In this section we analyze the equilibrium under each specification of beliefs by solving equation (4) 100,000 times with 100,000 random draws of the market conditions, i.e., δ, k_{M}, and k_{I}, in each of the four belief cases. We then use those 100,000 equilibria per case to evaluate momentum return moments and belief consistency. We use a concentration parameter $\alpha = 3$ in the symmetric Dirichlet distribution for capital proportions, $\mathrm{Dir}(\alpha)$. This provides a relatively diffuse prior belief for crowding, while maintaining a natural dependence among fractions of a whole with equal expected value 1/3 and vanishing probability that $k_{\text{type}} = 0$ or 1.[11] We use a mean (standard deviation) of -2.405 (0.125) in the log-normal distribution for δ,[12] and 0 (0.125) in the normal distribution for ϵ. These calibrations are designed to fit summary characteristics of momentum returns.[13]

Table 1 provides descriptive statistics of simulated momentum returns and Figure 1 provides plots of momentum investors' inference of the informed signal, δ^E, and the expected

[10]We do not presume that investors literally solve Eqs. (8) in practice. Rather, the assertion is that the market equilibrium that results from their trades aligns with this formula. Likewise, the Black-Scholes formula priced options before the equation was derived and published (Black and Scholes, 1972).

[11]In Internet Appendix IA.E we find qualitatively similar predictions under various permutations on α and other characteristics of the presumed setting, such as a uniform distribution for δ.

[12]These values imply an average δ of 9.1% with standard deviation of 1.14%. The log-normal distribution has the advantage that the differential dividend is limited to positive values, and this distributional assumption is consistent with the evidence in Andersen, Bollerslev, Diebold, and Ebens (2001).

[13]The model's aim is to analyze whether market equilibrium in the presence of unanticipated crowding generates tail risk in momentum returns. The most salient way to do this is to presume a setting where there is no tail risk in the primitive distribution for cash flows, and analyze whether or not market equilibrium creates it. However, if we were to simulate a setting with tail risk in the residual, our simulated equilibria would be unaffected, as the model essentially employs mean-variance demands scaled to wealth (given our use of the Campbell and Viceira, 2002, approximation).

momentum returns m given that signal. The first four columns in Table 1 and plots in Figure 1 correspond to the base case (known crowding), the rational, the myopic, and the optimal linear setting; left to right respectively. To construct the plots, simulation trials are ranked into 100 bins according to the conditioning variable on the horizontal axis and the 1000 trials within each bin are averaged to approximate a conditional expectation. Table 1 also includes a fifth column tabulating summary characteristics of actual momentum returns for comparison.

The first row of plots in Figure 1 examines the rationality of beliefs—whether or not momentum investors' presumed belief of δ (δ^E = horizontal axis) is consistent with the average realized δ (vertical axis) in the corresponding simulations. Rationality implies an identity; i.e., momentum investors are correct on average. A wider (narrower) spread on the vertical axis implies that momentum investors successfully infer more (less) information.

The plots in the second row present the expected momentum return m as a function of beliefs, δ^E. A negative m means that momentum investors have systematically pushed the price of momentum stocks beyond their fundamental value. Hence these plots identify settings in which momentum investing generates destabilizing crowding effects.

[Insert Figure 1 and Table 1 near here]

Base case (known crowding). The base case presumes that momentum investors know the degree of crowding from competing momentum investors at the time of trade. This allows perfect coordination of demands so there are no crowding effects. We use the resulting distribution of momentum returns as our benchmark for evaluating the more interesting cases in which investors must form beliefs for unknown crowding.

From Panel A of Table 1, column 1, note that expected momentum returns m in this base case are close to normally distributed with some positive skewness, and are never negative—momentum investors never overshoot informed investors' signal because they can perfectly infer it from the market clearing price. Nevertheless, from Panel B of Table 1, realized momentum returns ($m + \epsilon$) are frequently negative. From the fifth column, note that modeled realized returns in the base case roughly match the empirical mean and variance of quarterly momentum returns, indicating that our assumed distribution of model inputs is well calibrated. However, in contrast to the empirical distribution of momentum returns, modeled returns are essentially normally distributed with no excess tail risk.

Table 1 also presents the simulated profitability of momentum, both as a mean abnormal portfolio return ("mean profit") and a certainty-equivalent return. Profitability takes into account both the momentum return and the magnitude of the position taken in the momentum portfolio. Mean profits in the base case are slightly less and the certainty equivalents slightly more than the corresponding empirical mean and certainty equivalents.

Plot A.1 of Figure 1 presents an identity relation between the mean of informed investors' valuation spread δ and momentum investors' expectation of it δ^E. This confirms that prices fully reveal private information in the base case. Plot A.2 shows that the model does not embed any tail risk in momentum returns absent crowding. In particular, there is no random draw of market conditions that yields negative expected momentum returns.

Rational beliefs. The second columns in Figure 1 and Table 1 present the rational case. From Plot B.1 of Figure 1, mean beliefs δ^E map identically into the mean δ underlying those beliefs, consistent with rational expectations. However, the dispersion of inferences (vertical axis) is narrower than in Plot A.1, indicating that momentum investors cannot infer δ as precisely as when crowd size is known. Plot B.2 indicates that even in the highest bin for δ^E momentum investors do not systematically overshoot fundamental value and generate a negative expected momentum return. From Panel A of Table 1, expected momentum returns exhibit no negative skewness, excess kurtosis, or incremental volatility, just as in the base case. From Panel B, realized momentum returns and profitability are only slightly lower than in the base case. Overall, these results provide our main theoretical result, summarized as

Result 1. *When momentum investors rationally condition on the momentum portfolio price to infer crowd size, unanticipated crowding does not generate momentum tail risks.*

Myopic beliefs. Plot C.1 of Figure 1 indicates a strongly concave relation between beliefs δ^E and fundamental value δ when momentum investors hold myopic beliefs. In particular, when the randomly drawn market conditions generate beliefs of a high momentum return, those beliefs systematically and grossly overstate the corresponding signal of informed investors. From Plot C.2, this generates an expected momentum return m on the order of -400% in the highest beliefs bin. From the third column of Table 1, momentum returns exhibit substantial excess volatility and kurtosis with extreme negative skewness. We summarize this as

Result 2. *When momentum investors myopically employ a linear strategy, making no effort to infer crowd size from the momentum portfolio's price, unanticipated crowding can lead to extreme tail risk in momentum returns.*

Contrasting Results 1 and 2 makes it clear that unanticipated crowding *can* generate momentum tail risk to an essentially unbounded degree, if momentum investors employ myopic strategies. Yet, when momentum strategies are based on rational expectations, crowding-induced tail risk is completely avoided. Thus, predictions of crowding-induced tail risk depend crucially on how the momentum crowd acts when tail risk is high. If unanticipated crowding does (does not) account for momentum tail risk, then we ought to observe an elevated (reduced) crowd size when tail risk is high. Before turning to our empirical analysis of this matter, we consider one more simulation (as in Stein, 2009) for completeness.

Optimal linear beliefs. In this setting momentum investors base their demands on linear beliefs $\delta^E = \lambda^{-1} f$ as in the myopic case. However, rather than completely ignoring crowding effects, investors choose a value for λ^{-1} between one (which implies momentum demands are zero for all f) and the completely myopic case of 1.5 ($= (Ek_{\mathrm{M}} + Ek_{\mathrm{I}})^{-1}$ where $Ek_{\mathrm{M}} = Ek_{\mathrm{I}} = 1/3$) to yield the highest average realized utility from equation (2). We use a grid search to find that value to be $\lambda^{-1} = 1.12$.

From Plot D.1 of Figure 1, fundamental value δ rises monotonically with beliefs δ^E, but the relation is nowhere near an identity as required under rationality. Investors still succumb to feedback effects and overstate beliefs when they are high (i.e., $\delta^E > \delta$), and they understate beliefs at just about all other levels. As a result, they sacrifice substantial profit relative to the rational case. For example, from column 4 of Table 1 certainty equivalent returns are around 70 percent lower. However, feedback effects of unanticipated crowding are greatly attenuated. While there are instances of negative m in Plot D.2, they are not extreme. The higher moments of momentum returns in Table 1 are well behaved. In short, by proportionately restricting trade, a linear momentum strategy can be made to attenuate crowding-induced tail risk, albeit in a grossly inefficient way.

IV. Empirical Section

We base our empirical analysis on quarterly holdings from the Thomson Reuters Institutional 13f database starting in the first quarter of 1980 and ending in the third quarter of 2015. Stock data are from the Center for Research in Security Prices (CRSP) database using price and share adjustment factors, restricted to CRSP share codes 10 and 11 and a listing on AMEX, NYSE, or Nasdaq. The momentum return at time t is defined as the return of winners (stocks in the top 10% using NYSE cutoffs, sorting on returns from months $t - 12$ to $t - 2$) minus the return of losers (stocks in the bottom 10% similarly constructed). Returns are value-weighted within each decile and taken from Kenneth French's online data library.

A. Crowding Proxies

We construct crowding proxies by first applying an algorithm to designate institution i a momentum investor in quarter q, then aggregating this designation using various weighting schemes.

A.1. Momentum Designation

Following Grinblatt et al. (1995) we base our designation of a momentum investor on the following score:[14]

$$\texttt{SCORE}_{i,q} = \sum_{j=1}^{J} \left(\omega_{i,j,q} - \omega_{i,j,q-1}\right) r_{j,q-1}, \tag{10}$$

where $r_{j,q}$ is the quarter q return on stock j and ω is a portfolio weight with

$$\omega_{i,j,q} - \omega_{i,j,q-1} = \frac{w_{i,j,q} P_{j,q-1}}{\sum_{j=1}^{J} w_{i,j,q} P_{j,q-1}} - \frac{w_{i,j,q-1} P_{j,q-1}}{\sum_{j=1}^{J} w_{i,j,q-1} P_{j,q-1}}, \tag{11}$$

and $w_{i,j,q}$ indicates shares held in stock j by institution i at the end of quarter q and $P_{j,q-1}$ is the price of stock j at the end of quarter $q-1$. We fix the prices in equation (11) to avoid passive changes in portfolio weights induced by returns.

A positive $\texttt{SCORE}$ implies trading aligned with a momentum strategy. However, a single quarter of alignment in trading is surely a noisy indication of strategy. Thus, we use:

[14] In previous versions of the paper we also include an alternative scoring procedure based on the now standard $12-1$ momentum strategy outlined in Jegadeesh and Titman (1993). The procedure is largely redundant and yields similar results so we restrict our attention to this procedure as it has precedence in the literature. The corresponding results are available upon request.

Momentum designation. *Institution i is designated a momentum investor in quarter q if it has a positive* SCORE *from equation* (10) *in each quarter $q-3$ through q. We denote this $\mathbb{1}_{\texttt{MOM}_{i,q}} = \mathbb{1}_{\sum_{l=0}^{3} \mathbb{1}_{\texttt{SCORE}_{i,q-l}>0}=4}$, where $\mathbb{1}$ is the indicator function.*

Since 13f filings do not contain short positions, this definition could fail to identify a momentum institution whose long positions underweight winners and overweight losers while its short positions overweight losers. We believe this to be implausible as it requires extreme inconsistency between the momentum tilt of long and short positions and it implies large short positions. Mutual funds dominate our sample and the short side of their portfolios is typically small (Ang, 2014). Moreover, if a momentum investor buys winners and shorts losers they will still be scored a momentum investor under our procedure.

A.2. Measures of Momentum Investing

We conduct our analyses using two sets of three measures for crowd size (i.e., proxies for k_M)—six in total. The three measures focus on: the count of institutions following a momentum strategy (denoted CNT); their assets under management (denoted AUM); and their trading (more precisely, quarterly change in holdings) (denoted TRD). The two sets differ with respect to the scope of crowding. Measures constructed at a factor level examine institutions' crowding into a momentum factor-mimicking portfolio (no consideration of individual securities). Measures constructed at a security level first examine the crowd size in each component security of the momentum factor-mimicking portfolio, then aggregate across the portfolio.

As in Stein (2009), our theoretical analysis points to coordination among momentum investors as the key driving force behind crowding-induced tail risk in momentum returns. With perfect coordination, feedback effects from crowding are eliminated. On the one hand, it seems reasonable to conjecture that the ability to coordinate should depend on the number of momentum investors involved in the coordination. This suggests that a count-based measure of crowd size might best predict crowding effects. On the other hand, our rational-expectations theory shows how momentum investors can coordinate implicitly, via price. Thus, rational momentum investors can avoid feedback effects irrespective of the number of competing peers. Thus, we consider a count based measure to help distinguish these predictions.

We construct a count-based measure at the factor level using the fraction of momentum institutions in quarter q:

$$\texttt{CNT\_F}_q = \frac{1}{N_q} \sum_{i=1}^{N_q} \mathbb{1}_{\texttt{MOM}_{i,q}}. \tag{12}$$

Scaling by N_q, the total count of institutions in quarter q, makes the measure comparable across quarters. CNT_F only considers overall momentum investing, but crowding effects might be more salient when momentum investors buy (or sell) the same individual stocks. This motivates an alternative stock-level measure of crowding. Thus, we define the crowd size at the individual security level and then aggregate across momentum securities using:

$$\texttt{CNT\_S}_q = \sum_{j=1}^{J} \left(\bar{\omega}_{j,q} - \underline{\omega}_{j,q}\right) \frac{\sum_{i=1}^{N_q} \mathbb{1}_{w_{i,j,q}>0} \mathbb{1}_{\mathrm{MOM}_{i,q}}}{\sum_{i=1}^{N_q} \mathbb{1}_{w_{i,j,q}>0}}, \tag{13}$$

where $\bar{\omega}_{j,q}$ is the weight of security j in the winner leg of the factor mimicking portfolio (relative market capitalization otherwise zero), $\underline{\omega}_{j,q}$ is the corresponding weight in the loser leg, and the second term in the sum indicates the fraction of institutional holders of the security that follow a momentum strategy. A high (low) ratio indicates a crowded market for a winner (loser) security.[15]

[15] Because these fractions can take on extreme values at the individual security level, we Winsorize at the 1 and 99 percent level.

The difference, summed across securities, yields CNT_S as an aggregate measure of crowding.

While lack of coordination is key to crowding effects, the dollar magnitude of momentum investment should also play an important role as it directly ties to supply and demand conditions. Thus we also construct proxies for k_M weighting by assets under management, and by the dollar volume of momentum trade. To do so we first construct the following five variables by institution-quarter:[16]

$$\texttt{HOLD}_{i,q} = \sum_{j=1}^{J} w_{i,j,q} P_{j,q}, \qquad \texttt{WHOLD}_{i,q} = \sum_{j=1}^{J} w_{i,j,q} P_{j,q} \mathbb{1}_{\iota_{j,q}=1}, \qquad \texttt{LHOLD}_{i,q} = \sum_{j=1}^{J} w_{i,j,q} P_{j,q} \mathbb{1}_{\iota_{j,q}=-1},$$

$$\texttt{WTRD}_{i,q} = \sum_{j=1}^{J} \left(w_{i,j,q} - w_{i,j,q-1}\right) P_{j,q-1} \mathbb{1}_{\iota_{j,q}=1}, \qquad \texttt{LTRD}_{i,q} = \sum_{j=1}^{J} \left(w_{i,j,q} - w_{i,j,q-1}\right) P_{j,q-1} \mathbb{1}_{\iota_{j,q}=-1},$$

where $\iota_{j,q} = \pm 1$ indicates a winner or loser security as in the model ($\iota_{j,q} = 0$ if not an extreme decile of past returns). We define

$$\texttt{AUM\_F}_q = \frac{\sum_{i=1}^{N_q} \texttt{HOLD}_{i,q} \mathbb{1}_{\texttt{MOM}_{i,q}}}{\sum_{i=1}^{N_q} \texttt{HOLD}_{i,q}}, \tag{14}$$

$$\texttt{TRD\_F}_q = \frac{\sum_{i=1}^{N_q} \texttt{WTRD}_{i,q} \mathbb{1}_{\texttt{MOM}_{i,q}}}{\sum_{i=1}^{N_q} \texttt{WHOLD}_{i,q}} - \frac{\sum_{i=1}^{N_q} \texttt{LTRD}_{i,q} \mathbb{1}_{\texttt{MOM}_{i,q}}}{\sum_{i=1}^{N_q} \texttt{LHOLD}_{i,q}}. \tag{15}$$

Since these measures reference aggregate momentum capital we denote them factor-level measures. AUM_F is essentially CNT_F weighted by institution size, whereas TRD_F weights by the dollar volume of momentum trading (quarterly change in holdings) separately in the winner and loser portfolios.

[16] Again, these variables Winsorize at the 1 and 99 percent level.

We also construct the security-level measures

$$\texttt{AUM\_S}_q = \sum_{j=1}^{J} \left(\bar{\omega}_{j,q} - \underline{\omega}_{j,q}\right) \frac{\sum_{i=1}^{N_q} w_{i,j,q} P_{j,q} \mathbb{1}_{\texttt{MOM}_{i,q}}}{\sum_{i=1}^{N_q} w_{i,j,q} P_{j,q}}, \tag{16}$$

$$\texttt{TRD\_S}_q = \sum_{j=1}^{J} \left(\bar{\omega}_{j,q} - \underline{\omega}_{j,q}\right) \frac{\sum_{i=1}^{N_q} \left(w_{i,j,q} - w_{i,j,q-1}\right) P_{j,q-1} \mathbb{1}_{\texttt{MOM}_{i,q}}}{\sum_{i=1}^{N_q} w_{i,j,q} P_{j,q}}. \tag{17}$$

These are again analogs to CNT_S using weights equal to either the dollars allocated to the momentum security, AUM_S, or the dollar trading in that security, TRD_S.[17] In all six cases we consider both levels (generically referenced as $\texttt{CROWD}_q$) and changes (generically referenced as $\Delta\texttt{CROWD}_q$) to capture anticipated and unanticipated components.[18] We use a GARCH(1,1) specification of expected volatility in the time series of crowding to capture uncertainty, referenced as CROWD_EVOL.

One final comment on count versus capital based measures. There is a close analogy in microstructure considerations of the impact of trading on prices (Jones, Kaul, and Lipson, 1994), and more general considerations of institutional trading's impact on prices as in Sias, Starks, and Titman (2006) and Edelen et al. (2016). In both settings it turns out that count-based measures provide a more effective proxy than capital-based measures. Hence, even though capital-based measures are closer to supply and demand conditions, count-based measures have a priori precedence and are perhaps more closely tied to coordination.

B. Descriptive Statistics

Table 2 provides summary statistics for the 13f data in Panel A and the six momentum proxies in Panel B. For the purpose of Panel A only, we define a *consistent* momentum investor as

[17] Again, fractions Winsorize at the 1 and 99 percent level.

[18] Thus CROWD either collectively refers to all variants (e.g., AUM, TRD, CNT crossed with _F and _S), or to a specific variant if so noted (as in the tables).

any institution that scores a momentum quarter as in equation (10) in at least two-thirds of available quarters. (We do not use this momentum trader definition in our later empirical analysis because it is forward looking.) From Panel A, 16% (986/6,360) of institutions are classified as consistent momentum investors. On average these institutions have higher turnover (26% compared to 20%); manage more assets (1.7 billion versus 1.5 billion); and hold more stocks (181 versus 136) than their counterparts.

[Insert Table 2 near here]

From Table 2 Panel B, 10.8% of institutions are designated momentum investors ($\mathbb{1}_{\texttt{MOM}_{i,q}} = 1$) in the typical quarter (`CNT`-factor measure) and 14.5% of institutional capital is held by momentum investors (`AUM`-factor measure). Both measures are highly persistent (83% and 77%, respectively). For the typical momentum stock only 4.3% of institutional holders are classified as momentum investors (`CNT`-security measure), but 7.3% of the aggregate institutional investment in the stock is held by momentum investors (`AUM`-security measure). These security-level measures are slightly less persistent than their factor-level counterparts (62% and 69%, respectively). The `TRD` measures indicate that momentum investors trade about 3.4% of the institutional capital traded in the typical momentum stock (using `TRD`-security) or the momentum portfolio (3.5% using `TRD`-factor). Because most crowding variables show strong persistence, we estimate volatility with a GARCH(1,1) specification (Bollerslev, 1986) using residuals from an AR(1) regression of quarterly crowding observations.

Table 3 summarizes regressions of momentum returns using the Fama and French (1993) three factor model (abbreviated FF3), and a dynamic version of the same model (dynamic FF3 or DFF3). The dynamic model is motivated by the evidence in Grundy and Martin (2001) of time-varying risk exposure in the momentum portfolio. It includes regressors with an interactive indicator variable for a positive prior-year factor return. Note that the dynamic model provides a substantial improvement in adjusted R-squared over the traditional Fama and French model. Also note that the Grundy and Martin (2001) argument applies similarly to all three Fama and French factors and two out of three interactive terms are significant. Also notice from the moments of momentum returns reported previously in Table 1 of Section III that momentum has substantial crash risk (high excess kurtosis with pronounced left-skewness) in our sample (which includes the momentum crash of March-May 2009).

[Insert Table 3 near here]

Table 4, Panel A shows the persistence of a momemtum $\texttt{SCORE}_{i,q}$ and a momentum designation, $\mathbb{1}_{\texttt{MOM}_{i,q}}$. First note that $\texttt{SCORE}_{i,q}$ has a transition probability of 0.54 at both one and four quarter horizons, implying a 19% greater likelihood of a positive momentum score four quarters ahead than the unconditional probability of 0.45. An institution's momentum designation $\mathbb{1}_{\texttt{MOM}_{i,q}}$ persists four quarters ahead with probability 0.34, which is 3.32 times the 0.10 unconditional probability reported in the table.[19]

In any given quarter, some institutions trade with and some against momentum (from Table 4, Panel A the unconditional mean $\texttt{SCORE}_{i,q} = 0.45$).[20] A possible concern is that institutions can be classified as momentum investors even if none of them trades on relative strength signals. An investor acting on signals uncorrelated with past returns has a 50% chance of a momentum score in any given quarter. Yet, only deliberate momentum investors acting on a relative strength signal can create feedback effects and the resulting tail risk. Accidental momentum investors who trade independently of (but coincident with) past returns do not generate feedback. This highlights the

[19]The 10% figure presented here differs from the 10.8% figure in Table 2 because the latter is a pooled mean, rather than a steady-state probability (estimated from the quarterly transition matrix).

[20]This is less than the 59% estimate that Grinblatt et al. (1995) provide for mutual funds. The difference probably reflects that our dataset includes hedge funds which are contrarians on average (Grinblatt, Jostova, Petrasek, and Philipov, 2016).

importance of using a measure such as $\mathbb{1}_{\texttt{MOM}_{i,q}}$ that identifies institutions that consistently follow a momentum strategy. From Table 4, 10% of all institution-quarters indicate a momentum strategy—i.e., the institution has traded in line with past returns for each of four consecutive quarters. An investor who trades uncorrelated with past returns has an expected $\mathbb{1}_{\texttt{MOM}_{i,q}}$ of $0.5^4 = 6.25\%$. Hence, the group of observed momentum investors is indeed higher than what chance alone would predict, suggesting a subset of institutional investors deliberately follows the momentum strategy. Interestingly, as institutions as a whole approximately hold the market portfolio (Lewellen, 2011), this also suggests some institutions have a contrarian bias, as in the setup of our model.

Table 4, Panel B evaluates predictability of trading in the stocks comprising the $12-1$ momentum factor mimicking portfolio used for momentum returns. Specifically, let $\texttt{MOM\_BUY}_{i,q}$ denote an indicator for

$$\sum_{j=1}^{J} (\omega_{i,j,q} - \omega_{i,j,q-1})\, \iota_{j,q} > 0, \tag{18}$$

where, as before, $\iota_{j,q} = +1$ (-1) when the $12-1$ past-return of stock j is in the top (bottom) decile ($\iota_{j,q} = 0$ otherwise). $\texttt{MOM\_BUY}_{i,q}$ identifies investor i as a net buyer of momentum stocks in quarter q. From Panel B, a momentum designation in quarter q predicts $\texttt{MOM\_BUY}_{i,q+4}$ with probability 0.69. This is substantially higher than the unconditional probability (an untabulated 0.4862), as seen with the $q+4$ likelihood ratio of 1.42. It is also substantially higher than the predictability using $\texttt{SCORE}_{i,q}$ (a likelihood ratio of 1.15). This shows that a consistent pattern of trading on relative strength over four consecutive quarters (i.e., $\mathbb{1}_{\texttt{MOM}_{i,q}}$) improves the identification of future momentum trading.

[Insert Table 4 near here]

Table 4 also documents the rapidly changing composition of the momentum portfolio.

Winners have a 56% chance of remaining winners the following quarter, but at four quarters the probability is only 16%, which is actually less than the 23% chance of becoming a loser. Persistence is higher with losers, with 31% retaining that classification after four quarters.

C. Crowding and Momentum Conditional Expected Returns

Table 5 shows the results of predictive regressions of momentum returns on the various crowding measures. In all regressions the return on the momentum portfolio is dated $q+1$. In Panel A we include as control regressors the dynamic Fama-French benchmark factors (see Table 3). Because the results in Panel A (with return benchmarking) are broadly consistent with the results in Panel B (without such controls), we focus our discussion on Panel A. In all regressions we include realized volatility of momentum returns as a control, computed from squared daily momentum returns in quarter q, which Barroso and Santa-Clara (2015) show strongly predicts (negatively) momentum returns.[21] The regression sample for momentum returns begins in the third quarter of 1981 and ends in the fourth quarter of 2015 to account for the six quarters of data required to compute all regressors.

[Insert Table 5 near here]

We find that both $\texttt{CROWD}_{q\text{-}1}$ and $\Delta\texttt{CROWD}_q$ measures consistently and significantly negatively predict momentum returns. This supports the idea that the information not yet incorporated into prices from trading is decreasing in crowd size.[22] It also demonstrates that our measures have power to detect crowding effects on the first moment of momentum returns.

The table confirms the Barroso and Santa-Clara (2015) result on lag volatility of momentum returns even with the anticipated volatility of crowding CROWD_EVOL included as a

[21]In unreported results we also controlled for the bear market states proposed by Cooper, Gutierrez, and Hameed (2004). Using this control in our sample period did not change our results.

[22]We discuss in the context of Figure IA.3 in the internet appendix that this pattern is consistent with our model.

regressor. Thus, the Barroso and Santa-Clara (2015) result does not appear to be related to crowding uncertainty. Indeed, if anything crowding uncertainty appears to positively predict momentum returns, but this result is not robust across proxies.

D. Crowding and Tail Risk in Momentum Returns

This section considers tail risk in momentum returns. We first conduct a probit analysis of tail probabilities. We then use a sorting procedure based on lagged crowding measures.

Table 6 presents the probit analysis of tail probabilities to assess the link between crowding and negative skewness in momentum returns. In particular, we examine the conditional probability of a momentum return in the pth percentile of the unconditional distribution, where p is either 5% or 10%. Note that the mean dependency documented in Table 5 implies that crowding predicts a leftward shift in the conditional distribution of momentum returns. Thus, even with no effect on higher moments, crowding is expected to increase these tail probabilities. To address this we use a Wald test applied to a bivariate probit analysis that considers the increase in left-tail probability relative to the decrease in right-tail probability. Our null hypothesis is that crowding fattens the left tail of momentum returns only because it shifts the entire distribution (including the mean) leftward, with no elevation in negative skew.

[Insert Table 6 near here]

Panel A (B) of Table 6 considers 5% (10%) tails of the unconditional distribution of momentum returns. The more extreme 5% tail is arguably more relevant in identifying a momentum "crash," but the 10% tail provides more data. Across Panels A and B the crowd measure (levels or changes) is associated with an increase in tail probability in 22 of 24 cases. The increase is never statistically reliable using the 5% tail (Panel A), but it is statistically reliable in 4 out of 12 specifications using the 10% tail (Panel B). Nevertheless, the p-values (in square brackets) for the difference in left versus right tail effect are far from significant in every case. Thus, we find no evidence that crowding increases tail risk.

[Insert Figure 2 near here]

Figure 2 plots the time series of our crowding measures. Because the academic discovery of momentum can be attributed to either Jegadeesh and Titman (1993) or Levy (1967), it may not be surprising that we observe no increase in crowding after 1993. This result is also consistent with Grinblatt et al. (1995) who find evidence of pervasive momentum investing in a sample of 155 mutual funds over the 1975-1984 sample period, a time interval set entirely before the publication of the seminal study of Jegadeesh and Titman (1993). We also notice that some of our crowding variables have a statistically significant time trend, but on average more than 85% of their variation is not explained by the trend.[23]

The measures indicate that the momentum strategy was indeed crowded during the internet bubble. Piazzesi and Schneider (2009) find similar evidence of increased trend following behavior during the "housing bubble" of 2007-2009. On the other hand, no striking pattern is discernible before or during the major momentum crash of 2009. If anything, momentum investing by 13f institutions seems to have retracted prior to that crash.

Baltzer, Jank, and Smajlbegovic (2019) find that institutional selling of loser stocks peaked in Germany during the Great Recession just before the 2009 momentum crash. This is consistent with the momentum strategy being crowded at the time, suggesting a causality that our results appear to contradict. But an important difference to keep in mind is that our study only considers institutions that appear to consistently trade on the momentum signal as part of a deliberate strategy. Their study considers the trading of all institutions, including those that sell losers for

[23] In unreported results, we also found that controlling for a time trend does not materially change any of our inferences.

reasons other than reaction to past returns (e.g., a correlation between financial-crisis losers and contractual constraints on institutional holdings). The latter does not generate feedback effects as in the crowding models of Stein (2009) or this paper.

While the notion of tail risk surely involves pronounced left skewness, to be meaningful tail risk must also be accompanied by high volatility and excess kurtosis.[24] In Table 7 we examine all three higher moments of momentum returns by sorting calendar quarters based on lagged crowding. We include a sort based on lagged realized volatility of momentum returns for comparison.

The evidence in Table 7 does not support crowding as a source of tail risk in momentum returns. Across the three moments and twelve specifications of `CROWD` and Δ`CROWD` we find only one instance of a significant relation consistent with crowding-induced tail risk.[25] By contrast, we find nine instances of a significant relation consistent with momentum investors *avoiding* momentum tail risk (i.e., crowding measure correlated with lower volatility, more positive skewness, and lower kurtosis).

[Insert Table 7 near here]

Contrast these results with the last column, where we condition on lagged volatility in momentum returns. Here, we see statistically reliable prediction of negative tail risk in all moments. This evidence is consistent with Barroso and Santa-Clara (2015) who find that a volatility-managed momentum strategy has much smaller crash risk than original momentum. This contrast between predictability of crash risk using return volatility and predictability using direct measures of crowding (particularly the contrast in direction of predictability) suggests that many momentum

[24] A large kurtosis combined with left skewness is much more meaningful if volatility is also high. A low volatility directly reduces the denominator in these quantities inflating their values.

[25] We find significantly more negative skewness using the change in AUM measure at the security level; t-statistic -2.28.

investors incorporate tail-risk predictability into their strategy (e.g., via lag return volatility). This would be consistent with the rational beliefs case of our model, which in turn suggests (as in Result 1) that feedback effects from crowding do not explain the tendency for high lag volatility to predict crashes.

Note that tail-risk avoidance is seen most prominently in count based (`CNT`) measures of crowding. Indeed, estimates using our capital based measures (`AUM` and `TRD`) are almost always statistically indistinguishable from zero whereas `CNT` based measures point to tail-risk avoidance in 8 of 12 instances. This suggests that tail-risk avoidance takes the form of marginal momentum investors pulling away from the strategy in volatile environments.

We agree with much of the literature that it is tempting to interpret the predictive power of volatility for momentum returns as indirect evidence of crowding effects. Indeed, our study was originally motivated from just this perspective. However, a careful examination of this crowding hypothesis—from both a theoretical perspective based on rational beliefs and from an empirical perspective using direct proxies for institutional momentum investing—yields the opposite conclusion. 13f momentum investors seem to identify potential destabilizing effects of crowding, and adjust their strategy participation accordingly.

E. Crowding and the Volatility of Momentum Returns

If volatility clustering in momentum returns derives from crowding rather than factors exogenous to momentum investors' actions, then we should find that momentum volatility is positively predicted by measures of crowding. To examine this, Table 8 presents predictive regressions for realized volatility of momentum returns computed from raw and risk-adjusted (dynamic FF3) daily returns over the quarter.

[Insert Table 8 near here]

Consistent with the preceding results in the literature, Table 8 finds that lagged realized volatility has strong predictive power for subsequent volatility, with t-statistics ranging between eight and twelve across regressions. However, we find no evidence that crowding positively predicts elevated return volatility. Indeed, the relation is significantly negative in many cases. This result suggests that expectation of momentum-strategy risk affects institutions' willingness to participate in the strategy. Note that this interpretation implies that forward-looking institutions observe a wider information set than just lagged volatility in momentum returns, which we have controlled for.

The explanatory power of crowding measures in Table 8 pales in comparison to that of lagged realized volatility. This is perhaps not surprising, as lag dependent variables capture all persistent characteristics of the setting. Moreover, lag volatility is estimated using daily data whereas crowding measures are based on holdings data observed at a quarterly frequency so the precision of estimates differs greatly. What the crowding measures have going for them is that they are explicit economic measures brought to bear on the data from theory. Lag dependent variables offer little economic insight beyond persistence. We believe that this fact makes up for the shortcoming in predictive power.

F. Determinants of Crowding

Table 9 presents regressions of crowding measures $\texttt{CROWD}_q$ on lagged one-year momentum returns (denoted 1YR_RET) and one-year momentum-return volatility (denoted 1YR_VOL) computed from daily observations. Crowding is significantly negatively related to volatility lagged one quarter, using five out of the six proxies considered. Combining inferences from Tables 8 and 9, we conclude that reductions in crowding seem to anticipate high future volatility in momentum returns including but not limited to a quick response to past volatility.

[Insert Table 9 near here]

Chabot, Ghysels, and Jagannathan (2014) use a comprehensive sample period of 140 years to show that momentum crash risk increases after periods of good recent momentum returns. Piazzesi and Schneider (2009) use survey data to study momentum investing in the US housing market and find a substantial increase in the number of momentum investors towards the end of the housing boom. Taken together, these studies suggest that good past returns in momentum draw investors to the strategy and increase its crash risk. Table 9 shows that one-year returns indeed predict positively crowding in momentum, consistent with such a mechanism. However, higher past one-year momentum returns do not predict higher crash risk in our sample.[26] Hence, we are unable to relate our evidence of a dependence of crowding on past returns to the results in these papers' linking past returns to future crash risk.

Cooper et al. (2004) show that momentum returns are stronger in bull markets. However, in unreported results, we find that momentum crowding is not related to lagged market states after controlling for lagged momentum returns and volatility. Thus, the well-documented predictive ability of market states for momentum returns does not seem to stem from crowding in the strategy.

G. Comparison with Return-based Measures of Crowding

Lou and Polk (2013) propose a proxy for crowding defined as comomentum, a measure of abnormal co-movement of stocks in the momentum portfolio. In support of this proposition, they document a positive relation between comomentum and aggregate institutional ownership of the winners portfolio. Huang (2015) argues that the momentum gap, defined as the cross-sectional dispersion of formation period returns, also proxies for crowding. He supports this by showing that it is related to the difference in institutional ownership for winner versus loser portfolios. Both studies find that the indirect, returns-based proxies for crowding negatively relate to momentum

[26] In unreported results we found no significant relation between past one-year momentum returns and crashes in probit regressions controlling for momentum's volatility.

returns. Finally, volatility in momentum could also be hypothesized to arise from investor crowding, potentially representing a third returns-based proxy.

We have already seen that the relation between volatility and institutional crowding is more consistent with investors using using past and anticipated future volatility as a signal to exit the strategy than investors causing volatility by way of their crowding. The question we address here is how other returns-based proxies relate to our direct measures of institutional crowding. We focus on the momentum gap due to its simplicity; because it is a strong predictor of risk and return for momentum; and because of its proximity to the theory.

[Insert Table 10 near here]

Table 10 mirrors Table 7, except that the focus is momentum gap (denoted `GAP`) and momentum gap orthogonalized to our crowding measures (generically denoted $\texttt{GAP}^{\perp}$). If momentum gap's predictability stems from institutional crowding then orthogonalizing it should attenuate its predictability. In each column, we rank all months in our sample into terciles according to the sorting variable listed in the column heading, as of the preceding quarter. In the first column the ranking variable is the momentum gap. Consistent with Huang (2015), we find that a high momentum gap forecasts significantly higher volatility, negative skewness, and excess kurtosis. That is, momentum gap forecasts tail risk.

In the remainder of Table 10 we sort into terciles based on the momentum gap orthogonalized to each of our six proxies for institutional crowding, and to two additional variables: ΔMOM_INST from Huang (2015), defined as the percentage difference in aggregate institutional ownership between past winners and losers; and WIN_INST from Lou and Polk (2013), defined as the aggregate institutional ownership of the winner decile.[27] $\texttt{GAP}^{\perp}$ retains substantially all of the

[27]The WIN_INST measure is subject to two concerns. First, it shows a strong time trend as the asset management industry grows over the sample implying that momentum is mechanically more crowded over time. Yet institutions

predictive power for tail risk of momentum gap itself (column 1). We conclude that the momentum gap is a strong predictor of momentum tail risk for reasons unrelated to institutions' crowding in the strategy.

We also orthogonalize GAP to the volatility control used in many of our preceding empirical exercises. Since GAP is a measure of (cross-sectional) dispersion in returns, it should closely relate to our time-series measure of volatility (indeed, the correlation is 0.73). However, $\texttt{GAP}^{\perp}$ remains a reliable predictor of future volatility: predicted volatility averages 35.6% following periods of high $\texttt{GAP}^{\perp}$ compared to 21.0% following periods of low $\texttt{GAP}^{\perp}$. The t-statistic for the difference is 3.2. However, predicted skewness and predicted kurtosis is no longer reliably different for high versus low periods of $\texttt{GAP}^{\perp}$.

V. Conclusion

We provide a model of crowding with momentum investors who attempt to infer informed investors' private signal from prices. The model is similar to Stein (2009) in setting, but our analysis differs in its exploration of and emphasis on rational arbitrageur beliefs. Our primary result is that predictions of destabilizing effects from unanticipated crowding require a myopic, linear specification of beliefs in which momentum investors do not adequately account for the potential destabilizing effect of crowding on prices. With rational (generally nonlinear) beliefs, the potential destabilizing effects of crowding are internalized into demands. This mitigates feedback effects that would otherwise lead to destabilized prices and results in stable momentum returns. In short, our theory shows that crowding is not a viable explanation for momentum crashes in general, but that crowding with myopic momentum investors can provide that prediction.

as a whole approximately hold the market (Lewellen, 2011) so institutional holdings of losers must show a similar increase over time. Second, it is a measure based on holdings and therefore has high inertia.

Our empirical contribution is twofold. First, we directly examine proxies for momentum investing by institutional investors, in contrast to much of the literature that focuses on indirect inferences of crowding from return covariances or volatility. Second, we directly examine the implications of optimal versus myopic beliefs, documenting a generally inverse relation between momentum investing and future tail risk in momentum returns.

Across the empirical analyses, we consistently find evidence of crash-avoidance behavior rather than destabilizing feedback trading. Consistent with our theory under rational beliefs, we find no evidence that crowding by momentum investors deteriorates the higher moments of momentum returns (that is, causes crashes), despite a clear impact on the first moment of returns. We do find that past volatility in momentum returns identifies crashes, as in prior studies. But we also find that momentum investors both control for this result as well as condition on other sources to anticipate and back away from periods of instability.

References

D. Abreu and M. K. Brunnermeier. Bubbles and crashes. *Econometrica*, 71(1):173–204, 2003.

T. G. Andersen, T. Bollerslev, F. X. Diebold, and H. Ebens. The distribution of stock return volatility. *Journal of Financial Economics*, 61:43–76, 2001.

A. Ang. *Asset Management: A Systematic Approach to Factor Investing*. Oxford University Press, New York, 2014.

M. Baltzer, S. Jank, and E. Smajlbegovic. Who trades on momentum? *Journal of Financial Markets*, 42:56–74, 2019.

P. Barroso and P. Santa-Clara. Momentum has its moments. *Journal of Financial Economics*, 116 (1):111–120, 2015.

F. Black and M. Scholes. The valuation of option contracts and a test of market efficiency. *The Journal of Finance*, 27(2):399–417, 1972.

T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

G. Brown, P. Howard, and C. Lundblad. Crowded trades and tail risk. Working Paper SSRN, 2019.

M. Brunnermeier and L. H. Pedersen. Market liquidity and funding liquidity. *The Review of Financial Studies*, 22(6):2201–2238, 2009.

J. Y. Campbell and L. M. Viceira. *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*. Oxford University Press, 2002.

B. Chabot, E. Ghysels, and R. Jagannathan. Momentum trading, return chasing, and predictable crashes. NBER Working Paper, 2014.

M. J. Cooper, R. C. Gutierrez, and A. Hameed. Market states and momentum. *The Journal of Finance*, 59(3):1345–1365, 2004.

K. Daniel and T. J. Moskowitz. Momentum crashes. *Journal of Financial Economics*, 122(2): 221–247, 2016.

K. Daniel, R. Jagannathan, and S. Kim. A hidden markov model of momentum. Working Paper SSRN, 2019.

R. M. Edelen, O. S. Ince, and G. B. Kadlec. Institutional investors and stock return anomalies. *Journal of Financial Economics*, 119(3):472–488, 2016.

E. F. Fama and K. R. French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33:3–56, 1993.

T. J. George and C.-Y. Hwang. The 52-week high and momentum investing. *The Journal of Finance*, 59(5):2145–2176, 2004.

M. Grinblatt, S. Titman, and R. Wermers. Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior. *The American Economic Review*, 85(5): 1088–1105, 1995.

M. Grinblatt, G. Jostova, L. Petrasek, and A. Philipov. Style and skill: Hedge funds, mutual funds, and momentum. Working Paper SSRN, 2016.

D. Gromb and D. Vayanos. Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of Financial Economics*, 66(2-3):361–407, 2002.

B. D. Grundy and J. S. M. Martin. Understanding the nature of the risks and the source of the rewards to momentum investing. *The Review of Financial Studies*, 14(1):29–78, 2001.

H. Hong and J. C. Stein. A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance*, 54:2143–2184, 1999.

S. Huang. The momentum gap and return predictability. *Working Paper SSRN*, 2015.

S. Hvidkjaer. A trade-based analysis of momentum. *Review of Financial Studies*, 19(2):457–491, 2006.

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

C. M. Jones, G. Kaul, and M. L. Lipson. Transactions, volume, and volatility. *The Review of Financial Studies*, 7(4):631–651, 1994.

A. E. Khandani and A. W. Lo. What happened to the quants in august 2007? *Journal of Investment Management*, 5(4):29–78, 2007.

R. A. Levy. Relative strength as a criterion for investment selection. *The Journal of Finance*, 22(4): 595–610, 1967.

J. Lewellen. Institutional investors and the limits of arbitrage. *Journal of Financial Economics*, 102 (1):62–80, 2011.

D. Lou and C. Polk. Comomentum: Inferring arbitrage activity from return correlations. Working Paper London School of Economics, 2013.

R. C. Merton. On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance*, 29(2):449–470, 1974.

S. Morris and H. S. Shin. Liquidity black holes. *Review of Finance*, 8(1):1–18, 2004.

W. K. Newey and K. D. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708, 1987.

L. H. Pedersen. When everyone runs for the exit. *The International Journal of Central Banking*, 5: 177–199, 2009.

J. Peress. Wealth, information acquisition, and portfolio choice. *The Review of Financial Studies*, 17(3):879–914, 2004.

M. Piazzesi and M. Schneider. Momentum traders in the housing market: Survey evidence and a search model. *American Economic Review*, 99(2):406–11, 2009.

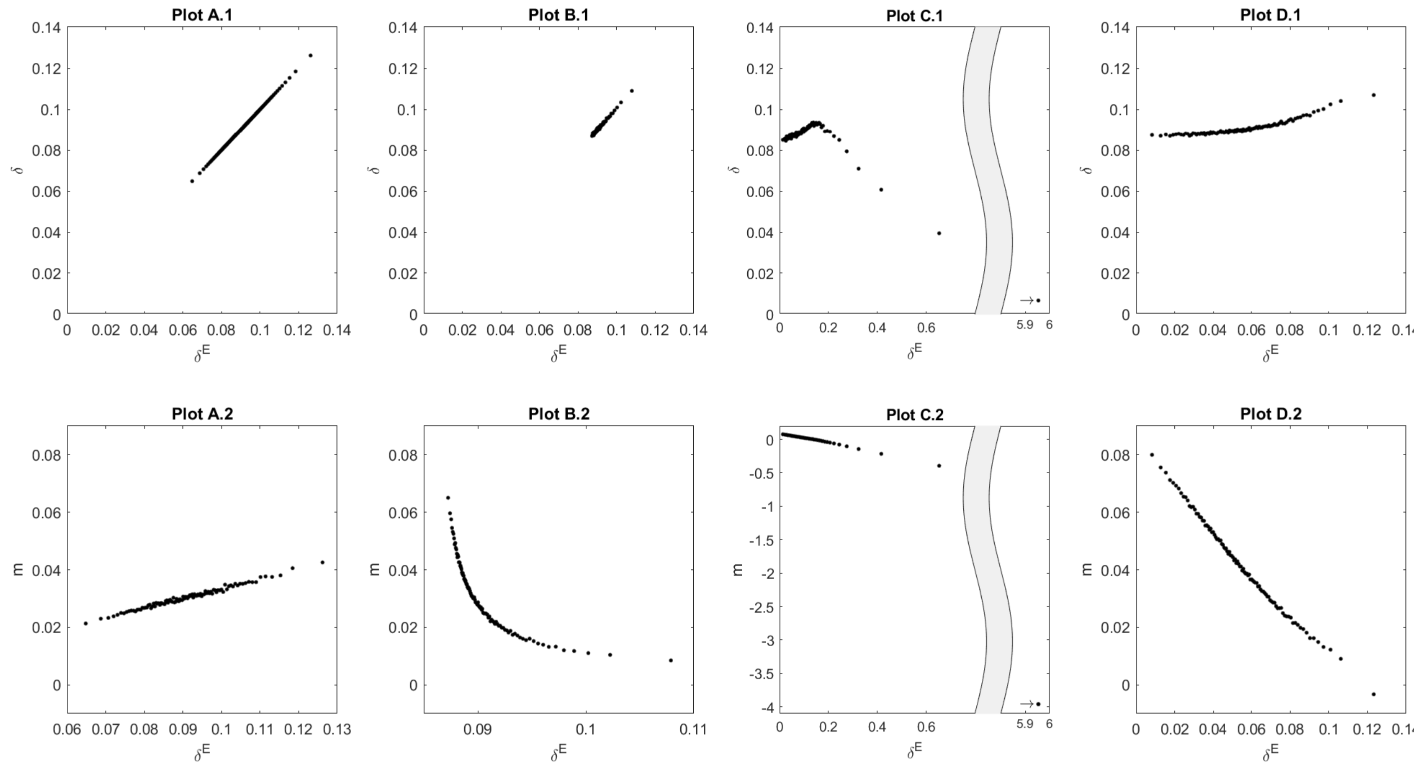
A. Shleifer and R. W. Vishny. The limits of arbitrage. *The Journal of Finance*, 52(1):35–55, 1997.

R. Sias, L. Starks, and S. Titman. Changes in institutional ownership and stock returns: Assessment and methodology. *The Journal of Business*, 79(6):2869–2910, 2006.

J. C. Stein. Presidential address: Sophisticated investors and market efficiency. *The Journal of Finance*, 64(4):1517–1548, 2009.

H. White. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4):817–838, 1980.

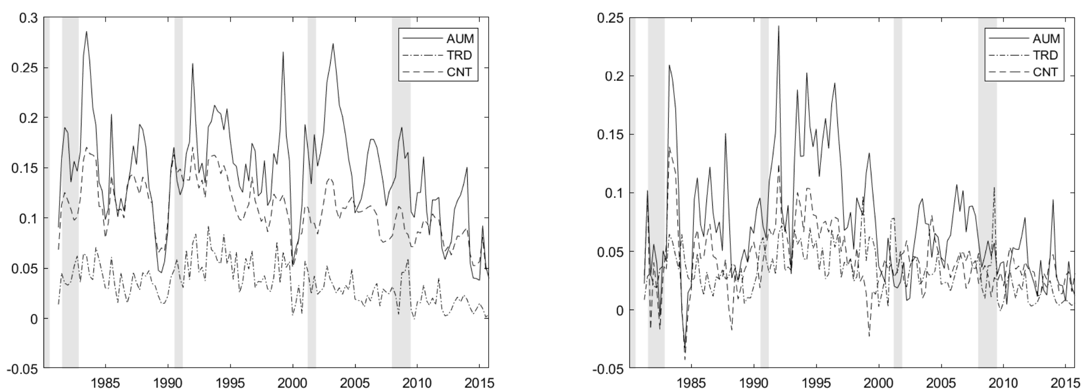
Figure 1: Beliefs and Expected Momentum Returns in Simulations



The simulations use 100,000 independent random draws of $\{k_I, k_M, \delta\}$, where k_I and k_M are informed and momentum capital, respectively, and δ is the signal of differential fundamental value for winners minus losers. k_I and k_M (and k_C) follow a Dirichlet distribution with concentration parameters 3, and $Ek_I = Ek_M = 1/3$. δ follows a log-normal distribution with $\mu = -2.405$ and $\sigma = 0.125$ implying an average δ of 9.1% with standard deviation of 1.14%. The market clearing formation period return f is solved for each $\{k_I, k_M, \delta\}$ pair by iteration using different specifications for momentum traders' beliefs δ^E: known crowding in Plots A.1-2, rational beliefs in Plots B.1-2, myopic beliefs in Plots C.1-2, and optimal linear beliefs in Plots D.1-2. The expected momentum return is then $m = \delta - f$. The 100,000 values of $\{\delta, m\}$ in each simulation run are ranked into 100 equally populated bins according to the horizontal-axis variable, and the plots represent the averages for the indicated variables within these bins.

Figure 2: Measures of Crowding

Panel (a) and (b) report the crowding measures at the factor and security level, respectively, constructed with 13f holdings data in the period from 01/1981 to 09/2015. The shaded areas denote NBER recessions.



(a) factor level

(b) security level

Table 1: Momentum Returns in Simulations

The table reports unconditional return statistics for the simulations described in Figure 1. Descriptive statistics of expected momentum returns in Panel A and realized momentum returns in Panel B are across all simulation. Mean, stdev, skew, kurt, min, and max refer to average, standard deviation, skewness, kurtosis, minimum, and maximum, respectively. The simulations are computed under the indicated belief specification, and the concentration parameters of the Dirichlet distribution are 3. Optimal linear beliefs maximize the utility of a constant relative risk aversion (CRRA) investor with $\gamma = 2$, and this optimal λ is reported in the row λ^{-1}. Profits are likewise the expected portfolio returns of a $\gamma = 2$ investor, and certainty equivalents are calculated for $\gamma = 2, 4, 10$, where portfolio weights are calculated as in equation (3). Momentum returns are given by $m + \epsilon$ where ϵ is randomly drawn from a zero-mean normal distribution with standard deviation 0.125. The last column of the table reports the corresponding statistics for the quarterly momentum returns in our sample period.

	Belief specification				Empirical
	known	rational	myopic	optimal linear	
λ^{-1}			1.50	1.12	
Panel A. Expected momentum returns m					
mean	3.0%	3.0%	-2.4%	4.2%	
stdev	1.4%	1.6%	174.2%	2.0%	
skew	0.6	0.4	-151.3	-0.3	
kurt	3.1	3.0	29218.7	10.8	
min	0.05%	-2.55%	-38957.17%	-53.10%	
max	10.26%	11.53%	13.16%	13.28%	
Panel B. Realized momentum returns $m + \epsilon$					
mean	3.1%	3.1%	-2.4%	4.2%	3.5%
stdev	12.6%	12.6%	174.7%	12.7%	12.5%
skew	0.0	0.0	-150.1	0.0	-0.9
kurt	3.0	3.0	28910.6	3.0	7.0
mean profit	3.65%	3.44%	-4863.08%	0.65%	4.04%
cer(2)	2.62%	2.53%	-100.00%	0.74%	1.43%
cer(4)	1.30%	1.25%	-100.00%	0.37%	0.84%
cer(10)	0.52%	0.50%	-100.00%	0.15%	0.35%

Table 2: Descriptive Statistics

In Panels A and B the indicated variable is computed by institution (i.e., 13f filer) and then summarized across institutions. Qtrs, med, stdev, and mgd refer to quarters, median, standard deviation, and managed, respectively. Assets are in units of $100 million and turnover is quarterly. Momentum investors are those institutions whose trading aligns with past returns at least 2/3s of the time. ΔCROWD and CROWD refer to the AUM, TRD, or CNT measures as indicated in the column header, either at the factor level or security level as defined in Section IV.A. CROWD_EVOL is the estimate of volatility from a GARCH(1,1) model.

Panel A. Institutions

	All			Mom Investors			not Mom Investors		
	mean	med	stdev	mean	med	stdev	mean	med	stdev
#Qtrs of data	37.8	28.0	31.8	32.5	20.5	30.9	38.8	29.0	31.9
#Qtrs missing	3.9	1.0	9.9	3.9	0.0	10.0	3.9	1.0	9.9
#Stocks held	143.2	62.9	275.5	180.6	87.8	275.3	136.4	58.3	275.0
Assets mgd	15.2	2.0	102.4	16.8	2.0	74.3	14.9	2.0	106.8
Turnover	0.21	0.16	0.17	0.26	0.21	0.18	0.20	0.15	0.16
#Institutions	6,360			986			5,374		

Panel B. Crowding variables

	AUM			TRD			CNT		
	mean	stdev	ar(1)	mean	stdev	ar(1)	mean	stdev	ar(1)
factor level:									
ΔCROWD	0.000	0.035	0.025	0.000	0.019	-0.439	0.000	0.017	0.109
CROWD	0.145	0.052	0.770	0.035	0.020	0.518	0.108	0.029	0.833
CROWD_EVOL	0.033	0.003	0.819	0.017	0.002	0.859	0.015	0.002	0.977
security level:									
ΔCROWD	0.000	0.039	-0.144	0.000	0.021	-0.328	0.000	0.025	-0.193
CROWD	0.073	0.050	0.692	0.034	0.020	0.456	0.043	0.029	0.623
CROWD_EVOL	0.034	0.011	0.947	0.018	0.003	0.329	0.020	0.007	0.984

Table 3: Factor Model

The table contains the factor exposures of quarterly momentum returns on the Fama-French three factor model (FF3) and a dynamic extension in which the three factors are interacted with dummies for positive past annual factor returns (DFF3). Alphas are monthly and t-statistics (in parentheses) use White (1980) standard errors. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	alpha	mkt	SMB	HML	Dmkt	DSMB	DHML	Adj R2
FF3	0.016***	-0.35**	-0.48**	-0.59**				12%
	(4.39)	(-1.98)	(-2.06)	(-2.06)				
DFF3	0.014***	-0.85***	-0.68***	-0.95***	0.81***	0.48	1.02**	25%
	(4.26)	(-2.93)	(-2.71)	(-2.68)	(2.77)	(1.22)	(2.14)	

Table 4: Transition Frequencies

Panel A tabulates the probability of remaining in the classification indicated in the row heading, in the quarter indicated in the column header. We require that the later quarter does not contain a missing observation. Likelihood refers to the conditional relative to unconditional probability of the indicated classification. Panel B tabulates the probability of being a net buyer of momentum stocks in the quarter indicated in the column heading (`MOM_BUY` = net buying of winner decile stocks plus net selling of loser decile stocks > 0), conditional on the classification in the row heading. Likelihood is here taken relative to the unconditional probability of being a net buyer of momentum stocks, which is 0.4862. Panel C tabulates the transition probabilities for stocks in the winner (Win), loser (Los), or middle (mid) deciles of the momentum ranking. 'All q' refers to the unconditional probability of classification.

Panel A. Institutions' type

	probabilities			likelihood	
	$q+1$	$q+4$	All q	$q+1$	$q+4$
$\texttt{SCORE}_{i,q} = 1$	0.54	0.54	0.45	1.20	1.19
$\mathbb{1}_{\texttt{MOM}_{i,q}} = 1$	0.71	0.34	0.10	7.05	3.32

Panel B. Institutions' trading

	probabilities			likelihood		
	q	$q+1$	$q+4$	q	$q+1$	$q+4$
$\texttt{SCORE}_{i,q} = 1$	0.68	0.57	0.56	1.39	1.16	1.15
$\mathbb{1}_{\texttt{MOM}_{i,q}} = 1$	0.78	0.70	0.69	1.60	1.44	1.42

Panel C. Stock returns

	$q+1$			$q+4$			All q
	Win	mid	Los	Win	mid	Los	
Winner	0.56	0.42	0.02	0.16	0.60	0.23	0.13
mid	0.08	0.82	0.09	0.12	0.74	0.14	0.67
Loser	0.02	0.33	0.65	0.17	0.52	0.31	0.19

Table 5: Momentum Factor Returns on Crowding Measures

Each column represents a predictive regression of quarterly momentum factor returns (1981–2015) on crowding. Panel A presents a dynamic factor specification that controls for the Fama and French (1993) three factor model with the factors interacted with dummies for positive past annual factor returns and the specification in Panel B uses no risk-factor controls. Each panel considers the three indicated crowding measures computed both at the factor level and the security level. The regressor $\texttt{CROWD}_{q\text{-}1}$ refers to the level of the crowding measure at the end of quarter $q-1$; $\Delta\texttt{CROWD}_q$ refers to the change over quarter q; and $\texttt{CROWD\_EVOL}_q$ is the estimate of volatility from a GARCH(1,1) model. 'Realized vol. of Mom rets.' is a control variable equal to the realized volatility of daily momentum returns over the previous quarter. Intercepts are not tabulated. The t-statistics (in parentheses) are computed with White (1980) standard errors. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Level of analysis:	factor			security		
Measure of crowding:	AUM	TRD	CNT	AUM	TRD	CNT
Panel A. Dynamic FF3 model						
$\Delta\texttt{CROWD}_q$	-0.21***	-0.34*	-0.33*	-0.07	-0.41**	-0.04
	(-2.64)	(-1.94)	(-1.84)	(-0.95)	(-2.05)	(-0.31)
$\texttt{CROWD}_{q\text{-}1}$	-0.24***	-0.34*	-0.58***	-0.18***	-0.36**	-0.26**
	(-4.32)	(-1.87)	(-4.35)	(-3.00)	(-2.39)	(-2.52)
$\texttt{CROWD\_EVOL}_q$	3.06**	2.30	6.60***	0.19	1.58	0.20
	(2.47)	(1.31)	(3.75)	(0.75)	(1.08)	(0.56)
Realized vol.	-0.27***	-0.28**	-0.25**	-0.29**	-0.28***	-0.28**
of Mom rets.	(-2.62)	(-2.54)	(-2.21)	(-2.45)	(-2.63)	(-2.30)
Adj-rsquare	40.7%	31.8%	37.7%	32.6%	34.0%	31.7%
Panel B. No risk controls						
$\Delta\texttt{CROWD}_q$	-0.19*	-0.33	-0.29	-0.08	-0.44*	-0.07
	(-1.94)	(-1.51)	(-1.36)	(-1.18)	(-1.69)	(-0.62)
$\texttt{CROWD}_{q\text{-}1}$	-0.23***	-0.44**	-0.50***	-0.18***	-0.41**	-0.26**
	(-3.83)	(-2.10)	(-3.39)	(-2.79)	(-1.96)	(-2.48)
$\texttt{CROWD\_EVOL}_q$	2.79*	2.21	4.55**	0.01	0.25	-0.06
	(1.83)	(1.06)	(2.27)	(0.03)	(0.14)	(-0.16)
Realized vol.	-0.29	-0.31*	-0.30	-0.33*	-0.27	-0.32*
of Mom rets.	(-1.61)	(-1.75)	(-1.56)	(-1.77)	(-1.51)	(-1.69)
Adj-rsquare	16.6%	9.0%	12.0%	9.7%	10.2%	8.7%

Table 6: Crowding and the Left-tail of Momentum Returns

Each column represents a Probit regression using an indicator for next-quarter momentum returns in the bottom 5% (Panel A) or 10% (Panel B) of the full-sample (1981–2015) distribution, where returns are dynamic Fama and French residuals. Each panel considers the six crowding proxies. The regressor $\texttt{CROWD}_{q\text{-}1}$ refers to the level of the crowding measure at the end of quarter $q-1$; $\Delta\texttt{CROWD}_q$ refers to the change over quarter q; and $\texttt{CROWD\_EVOL}_q$ is the estimate of volatility from a GARCH(1,1) model. 'Realized vol. of Mom rets.' is a control variable equal to the realized volatility of daily momentum returns over the previous quarter. Intercepts are not tabulated. T-statistics for the coefficient estimates are reported in parenthesis. Wald test p-values are reported in square brackets, testing the null hypothesis that a regressor's effect on the left tail probability is equal in magnitude and opposite in sign to the (untabulated) effect on the right tail probability. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Level of analysis:	factor			security		
Measure of crowding:	AUM	TRD	CNT	AUM	TRD	CNT
Panel A. Predicting the 5% left tail						
$\Delta\texttt{CROWD}_q$	8.4	12.3	19.2	-1.4	11.8	3.8
	(1.49)	(1.15)	(1.17)	(-0.21)	(1.03)	(0.28)
	[0.35]	[0.92]	[0.60]	[0.22]	[0.20]	[0.90]
$\texttt{CROWD}_{q\text{-}1}$	7.2	14.4	24.5*	4.9	26.4	22.1
	(1.55)	(1.00)	(1.79)	(0.77)	(1.39)	(1.33)
	[0.86]	[0.81]	[0.98]	[0.98]	[0.53]	[0.46]
$\texttt{CROWD\_EVOL}_q$	-1.4	43.2	-213.3	-11.8	-221.3	-91.4
	(-0.02)	(0.41)	(-1.09)	(-0.42)	(-1.30)	(-1.15)
	[0.29]	[0.43]	[0.48]	[0.17]	[0.05]*	[0.11]
Realized vol.	10.1***	9.9***	10.8***	9.9***	12.3***	9.9***
of Mom rets.	(3.22)	(3.18)	(3.10)	(2.92)	(2.87)	(2.94)
	[0.00]***	[0.00]***	[0.00]***	[0.00]***	[0.00]***	[0.00]***
Panel B. Predicting the 10% left tail						
$\Delta\texttt{CROWD}_q$	11.7**	6.7	16.5	0.4	12.5	-4.9
	(2.40)	(0.71)	(1.32)	(0.07)	(1.36)	(-0.61)
	[0.25]	[0.57]	[0.91]	[0.58]	[0.66]	[0.68]
$\texttt{CROWD}_{q\text{-}1}$	10.4***	15.4	21.6**	6.9	26.9**	9.5
	(2.68)	(1.25)	(2.22)	(1.47)	(2.04)	(1.18)
	[0.35]	[0.66]	[0.93]	[0.98]	[0.46]	[0.96]
$\texttt{CROWD\_EVOL}_q$	-11.3	20.7	-187.7	-14.1	-65.7	-25.1
	(-0.17)	(0.23)	(-1.35)	(-0.65)	(-0.84)	(-0.72)
	[0.14]	[0.30]	[0.92]	[0.74]	[0.73]	[0.57]
Realized vol.	11.5***	11.6***	11.6***	11.4***	11.8***	10.8***
of Mom rets.	(3.96)	(3.60)	(3.78)	(3.65)	(3.28)	(3.60)
	[0.00]***	[0.00]***	[0.00]***	[0.00]***	[0.00]***	[0.00]***

Table 7: Conditional Volatility, Skewness, and Kurtosis of Momentum Returns

We split the sample of monthly momentum returns (1981–2015) into terciles according to crowding (CROWD), change in crowding (ΔCROWD) or volatility in the last available quarter for each month. The T1 stands for the bottom tercile, T2 for the second tercile, and T3 for the top tercile. The values in parenthesis are t-statistics (in parentheses) for the difference between T3 and T1 obtained with the delta method, and *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	factor			security			Realized vol.
	AUM	TRD	CNT	AUM	TRD	CNT	of Mom rets.
Panel A. CROWD							
Volatility							
T1	23.8	25.7	32.3	23.1	23.3	31.3	15.3
T2	28.0	22.3	26.8	33.2	30.3	27.7	17.4
T3	25.8	29.5	16.5	20.1	23.8	17.2	38.3
	(0.40)	(0.74)	(-3.63)***	(-1.28)	(0.13)	(-3.23)***	(5.62)***
Skewness							
T1	-1.8	-1.5	-1.7	-0.3	-1.9	-1.7	-0.3
T2	-1.6	-0.7	-1.1	-2.0	-1.9	-1.4	-0.3
T3	-1.1	-1.8	-0.6	-0.1	0.0	0.0	-1.2
	(0.57)	(-0.26)	(1.83)*	(0.51)	(1.59)	(2.66)***	(-2.06)**
Kurtosis							
T1	14.6	10.8	10.6	4.3	15.5	10.9	4.0
T2	11.5	4.9	8.1	11.1	10.7	8.7	4.0
T3	8.9	12.6	4.7	4.3	5.9	3.4	6.7
	(-1.11)	(0.37)	(-2.77)***	(0.06)	(-2.08)**	(-3.40)***	(2.29)**
Panel B. ΔCROWD							
Volatility							
T1	26.0	27.0	25.7	21.6	26.4	25.1	
T2	20.7	26.4	26.3	30.5	25.6	29.2	
T3	30.2	24.6	25.9	25.0	26.1	23.6	
	(0.77)	(-0.51)	(0.04)	(0.93)	(-0.05)	(-0.34)	
Skewness							
T1	-1.8	-1.5	-1.8	0.6	-1.3	-2.4	
T2	-0.5	-1.6	-1.2	-2.1	-1.9	-1.6	
T3	-1.5	-1.4	-1.5	-1.6	-1.3	-0.2	
	(0.25)	(0.03)	(0.24)	(-2.28)**	(-0.03)	(2.27)**	
Kurtosis							
T1	15.1	9.5	15.4	5.2	10.3	15.8	
T2	5.8	13.6	9.0	11.1	14.8	11.4	
T3	9.2	11.2	10.5	11.2	9.4	4.5	
	(-1.32)	(0.40)	(-0.97)	(1.60)	(-0.25)	(-2.43)**	

Table 8: Volatility in Momentum Factor Returns on Crowding Measures

Each column represents a predictive regression of realized volatility in daily momentum factor returns over the next quarter (1981–2015) on crowding. Each panel presents a different dependent variable using daily: (1) residual returns obtained with the Fama and French three factor model with dynamic weights, and (2) raw returns on the momentum portfolio. Each panel considers the six indicated crowding measures. The regressor $\texttt{CROWD}_{q\text{-}1}$ refers to the level of the crowding measure at the end of quarter $q-1$; $\Delta\texttt{CROWD}_q$ refers to the change over quarter q; and $\texttt{CROWD\_EVOL}_q$ is the estimate of volatility from a GARCH(1,1) model. 'Realized vol. of Mom rets.' is a control variable equal to the realized volatility of daily momentum returns/residuals over the previous quarter. Intercepts are not tabulated. The t-statistics (in parentheses) are computed with Newey and West (1987) standard errors with 3 lags. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Level of analysis:	factor			security		
Measure of crowding:	AUM	TRD	CNT	AUM	TRD	CNT
Panel A. Dynamic FF3 model						
$\Delta\texttt{CROWD}_q$	0.03	-0.10	-0.05	0.04	-0.03	0.01
	(0.65)	(-0.72)	(-0.51)	(0.67)	(-0.24)	(0.11)
$\texttt{CROWD}_{q\text{-}1}$	-0.06**	-0.18*	-0.10	-0.03	-0.16**	-0.08*
	(-2.24)	(-1.82)	(-1.21)	(-1.15)	(-2.03)	(-1.80)
$\texttt{CROWD\_EVOL}_q$	0.12	-0.13	-0.75	-0.31**	-0.83	-0.42**
	(0.14)	(-0.12)	(-0.56)	(-2.53)	(-0.86)	(-2.27)
Realized vol.	0.77***	0.77***	0.74***	0.72***	0.81***	0.73***
of Mom rets.	(8.83)	(8.42)	(9.10)	(8.70)	(10.02)	(8.59)
Adj-rsquare	59.4%	59.2%	59.5%	60.2%	60.1%	59.6%
Panel B. No risk controls						
$\Delta\texttt{CROWD}_q$	0.06	-0.04	-0.06	0.01	0.06	-0.06
	(0.84)	(-0.21)	(-0.40)	(0.25)	(0.38)	(-1.00)
$\texttt{CROWD}_{q\text{-}1}$	-0.04	-0.15	-0.05	-0.02	-0.14	-0.12**
	(-1.37)	(-1.32)	(-0.46)	(-0.94)	(-1.28)	(-2.51)
$\texttt{CROWD\_EVOL}_q$	-0.14	-0.46	-1.71	-0.38**	-1.26	-0.50**
	(-0.12)	(-0.34)	(-0.95)	(-2.52)	(-0.83)	(-2.17)
Realized vol.	0.80***	0.80***	0.77***	0.76***	0.84***	0.76***
of Mom rets.	(8.84)	(8.64)	(9.15)	(8.84)	(11.99)	(8.72)
Adj-rsquare	63.3%	63.1%	63.5%	63.7%	64.5%	63.8%

Table 9: Momentum Factor Returns as a Determinant of Crowding

Each column represents a predictive regression of a different crowding measure on lag momentum returns (1YR_RET) and lag momentum realized volatility (1YR_VOL) computed using daily momentum returns. Intercepts are not tabulated. The t-statistics (in parentheses) are computed with Newey and West (1987) standard errors with 3 lags. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Level of analysis:	factor			security		
Measure of crowding:	AUM	TRD	CNT	AUM	TRD	CNT
$\texttt{1YR\_RET}_{q\text{-}1}$	0.72**	0.25***	0.41***	0.29	0.24***	0.17
	(2.57)	(3.40)	(2.69)	(1.32)	(2.62)	(1.34)
$\texttt{1YR\_RET}_{q\text{-}5}$	0.92***	0.40***	0.52***	0.35	0.27***	0.21
	(3.22)	(3.93)	(2.94)	(1.29)	(3.08)	(1.34)
$\texttt{1YR\_VOL}_{q\text{-}1}$	-0.28*	-0.20***	-0.36***	-0.39**	-0.08	-0.18**
	(-1.94)	(-3.31)	(-4.50)	(-2.46)	(-1.34)	(-2.09)
$\texttt{1YR\_VOL}_{q\text{-}5}$	0.41*	0.07	0.18**	-0.10	0.02	0.08
	(1.90)	(1.34)	(2.33)	(-0.68)	(0.41)	(1.22)
Adj-rsquare	10.9%	20.1%	19.1%	12.9%	10.6%	2.2%

Table 10: Robustness: Conditional Volatility, Skewness, and Kurtosis of Momentum Returns

To calculate each column we split monthly momentum returns (1981–2015) into terciles every four quarters according to the level of Huang (2015)'s momentum gap variable in the column labelled GAP, and the momentum gap variable orthogonal to the level of the variables shown for the columns labelled $\text{GAP}^{\perp}$. Variables not previously used are ΔMOM_INST which is the percentage difference in aggregate institutional ownership between past winners and losers (see, Huang, 2015), and WIN_INST which is the aggregate institutional ownership of the winner decile (see, Lou and Polk, 2013). T1 stands for the bottom tercile, T2 for the second tercile, and T3 for the top tercile. The values in parenthesis are t-statistics for the difference between T3 and T1 obtained with the delta method. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	GAP	$\text{GAP}^{\perp}$ (orthogonal to:)								
				factor			security			Realized vol.
		ΔMOM_INST	WIN_INST	AUM	TRD	CNT	AUM	TRD	CNT	of Mom rets.
Volatility										
T1	12.8	12.2	13.5	12.9	12.3	12.7	14.0	14.3	12.9	21.0
T2	19.2	19.9	18.7	19.9	19.7	19.4	20.8	18.9	20.6	17.8
T3	38.6	38.5	38.6	38.3	38.6	38.6	37.4	38.3	37.9	35.6
	(6.42)***	(6.66)***	(6.24)***	(6.30)***	(6.58)***	(6.59)***	(5.63)***	(5.96)***	(6.22)***	(3.23)***
Skewness										
T1	-0.3	-0.4	-0.3	-0.5	-0.4	-0.3	-0.5	-0.4	-0.4	-0.7
T2	0.0	0.0	0.0	-0.6	0.1	-0.1	-0.2	-0.7	-0.4	-0.2
T3	-1.3	-1.3	-1.3	-1.3	-1.3	-1.3	-1.4	-1.3	-1.3	-1.5
	(-2.44)**	(-2.22)**	(-2.64)***	(-2.13)**	(-2.38)**	(-2.47)**	(-1.71)*	(-2.19)**	(-2.21)**	(-1.36)
Kurtosis										
T1	3.3	3.4	3.2	3.7	3.7	3.4	4.3	3.9	3.3	6.4
T2	3.8	3.7	4.2	4.1	3.5	3.8	4.3	4.5	4.6	4.8
T3	6.6	6.6	6.6	7.0	6.7	6.6	7.4	7.0	7.0	8.4
	(2.86)***	(2.79)***	(3.15)***	(2.82)***	(2.69)***	(2.91)***	(2.10)**	(2.64)***	(3.23)***	(1.11)

Internet Appendix to accompany the paper

"Crowding and Tail Risk in Momentum Returns"

(Not for publication)

Abstract

This internet appendix accompanies the paper "Crowding and Tail Risk in Momentum Returns." It contains the literature review in Section IA.A, details on the model's derivation and solution in Sections IA.B to IA.D, and further simulation results in Section IA.E.

IA.A. Existing Literature

Our paper is related to the empirical and theoretical literature on momentum. Momentum was initially documented for US stock returns (Levy, 1967; Jegadeesh and Titman, 1993) and has since been documented for stock returns in most countries (Rouwenhorst, 1998) and across asset classes (Asness, Moskowitz, and Pedersen, 2013). Besides its very high average returns, momentum carries significant downside risk or negative skewness in the form of occasional large crashes (Daniel and Moskowitz, 2016). Existing research also shows that institutional investors are momentum traders, i.e., tilt their portfolios towards momentum stocks (Grinblatt et al., 1995; Lewellen, 2011; Edelen et al., 2016; Baltzer et al., 2019). Our paper contributes to this literature by directly examining whether uncertain institutional participation in the momentum strategy is the source of higher-moment return characteristics.

A recent empirical literature examining the time series properties of momentum finds results broadly consistent with an over-reaction explanation of the effect. The premium is stronger in periods of bull markets (Cooper et al., 2004), high liquidity (Avramov, Cheng, and Hameed, 2016), and high sentiment (Antoniou, Doukas, and Subrahmanyam, 2013). Hillert, Jacobs, and Müller (2014)'s finding that momentum is more pronounced in firms with more media coverage also supports an over-reaction interpretation, as does the evidence in Edelen et al. (2016) regarding institutional purchases in the portfolio-formation period.

On the other hand, the momentum premium is stronger in stocks experiencing frequent but small price changes that are less likely to attract attention (Da et al., 2014) or those characterized by small trades of investors under-reacting to past returns (Hvidkjaer, 2006). Also there is recent evidence that momentum is somehow explained by improvements in firm fundamentals (Novy-Marx, 2015; Sotes-Paladino, Wang, and Yao, 2016; DeMiguel, Martín-Utrera, Nogales, and Uppal, 2020). This evidence suggests momentum investors exploit under-reaction and as such (exogenous increases in) crowding should reduce its premium.

The related theoretical literature on momentum offers theories based on institutional investors and fund flows (Vayanos and Woolley, 2013) or behavioral biases such as over-reaction / self-attribution (see, e.g., Barberis, Shleifer, and Vishny, 1998; Daniel, Hirshleifer, and Subrahmanyam, 1998) or information externalities and gradual diffusion of information (see, e.g., Stein, 1987; Hong and Stein, 1999; Andrei and Cujean, 2017). Our work is most closely related to the latter branch of the literature.

Our model builds on the information externality that the actions of unanticipated momentum investors impose on their peers. Thus, it is closest in development to Stein (2009), but follows in a long line of research relating to arbitrageur information coordination and externalities. This literature dates to Stein (1987) who characterizes the externality, and Scharfstein and Stein (1990) and Froot, Scharfstein, and Stein (1992) who relate it to herding behavior. Hong and Stein (1999) relate the externality to persistence and reversal patterns in returns. A related branch of the literature identifies the positive feedback trading of momentum investors as a source of destabilizing noise in prices, e.g., De Long, Shleifer, Summers, and Waldmann (1990a,b).

More recently, Kondor and Zawadowski (2019) study whether the presence of more arbitrageurs improves welfare in a model of capital reallocation. Trades in the model can become crowded due to imperfect information, but arbitrageurs can also devote resources to learn about the number of earlier entrants. They find that if the number of arbitrageurs is high enough, more arbitrageurs do not change capital allocations, but decrease welfare due to costly learning.

Related empirical research includes Hanson and Sunderam (2014) who construct a measure of the capital allocated to momentum and the valuation anomaly (book-to-market or B/M) using short-interest. They find some evidence that an increase in arbitrage capital has reduced the returns on B/M and momentum strategies. In addition, Lou and Polk (2013) proxy for momentum capital with the residual return correlations in the short and long leg of the momentum strategy and find that momentum profits are lower and crashes more likely in times of higher momentum capital. Baltzer et al. (2019) classify institutions as a whole as momentum traders and find that, in Germany, momentum trading peaked before the crash. While our analysis uses a different approach and insights in proxying for momentum capital, our results on unanticipated momentum capital and momentum returns are generally consistent with these findings, but we do not attribute momentum's crashes to crowding. Finally, Huang (2015) proposes a momentum gap variable, which is defined as the cross sectional dispersion of formation period returns. He shows that this measure predicts momentum returns and crashes, and argues that this is consistent with Stein (2009)'s crowded trade theory. Throughout our analysis, we control for momentum's past volatility, which has a correlation of 0.73 with the momentum gap measure. We also verify in Section IV.G that momentum gap's predictive power for crash risk is unrelated to various institutional measures of momentum crowding. This corroborates our finding that momentum's crashes are not explained by crowding.

We go beyond the usual focus on first moments to study the determinants of the risk of momentum. This relates our work to a recent strand of literature focusing on the predictability of the moments of momentum. Barroso and Santa-Clara (2015) show that the volatility of momentum is highly predictable and it is a useful variable to manage the risk of the strategy. Daniel and Moskowitz (2016) argues the crash risk of momentum is due to the optionality effect of the losers portfolio that resembles an out-of-the-money call option after extreme bear markets. Jacobs, Regele, and Weber (2015) examine the expected skewness of momentum as a potential explanation of its premium. They propose an enhanced momentum strategy but find that managing its risk results in a performance hard to reconcile with a premium for skewness. Grobys, Ruotsalainen, and Äijö (2018) find industry momentum has different risk properties from standard momentum but shows similar gains from risk management. Our results address the question of whether investors condition their exposure to momentum using this new-found predictability. Consistent with the economic case for managing the risk of momentum, we find less crowding in momentum after periods of high volatility.

IA.B. Derivation of Equation (3)

First notice that solving equation (2) is equivalent to solving each of the following (presuming $\gamma > 1$)

$$\max_{\varsigma} \quad \frac{K_{type,2}^{1-\gamma}}{1-\gamma} \cdot E\left[e^{(1-\gamma)\left(r_f+\log\left(1+\varsigma\left(e^{r_p-r_f}-1\right)\right)\right)}\right] \quad \Leftrightarrow \quad \min_{\varsigma} \quad \log E\left[e^{(1-\gamma)\log\left(1+\varsigma\left(e^{r_p-r_f}-1\right)\right)}\right],$$

where ς is the weight on the risky asset portfolio and r_p its log return, and r_f is the log risk-free rate. Second, to solve for the fraction of wealth invested in the risky portfolio, we follow Section

2.1.1 in Campbell and Viceira (2002, Internet Appendix) and approximate the function $g(r_p - r_f) = \log(1 + \varsigma(e^{r_p - r_f} - 1))$ with a second-order Taylor expansion around 0:[28]

$$g(r_p - r_f) \cong \log(1) + \frac{\varsigma e^0}{1 + \varsigma(e^0 - 1)}(r_p - f) + \frac{1}{2}\frac{\varsigma\left[e^0\left(1 + \varsigma\left(e^0 - 1\right)\right) - \varsigma e^{2\cdot 0}\right]}{\left(1 + \varsigma\left(e^0 - 1\right)\right)^2}(r_p - f)^2$$

$$\cong \varsigma(r_p - f) + \frac{1}{2}\left(\varsigma - \varsigma^2\right)\sigma^2, \qquad \text{(IA.B.1)}$$

where $(r_p - f)^2$ is replaced with its conditionally expectation σ^2 as in Campbell and Viceira (2002, Internet Appendix). Using equation (IA.B.1), we can rewrite the maximization problem to

$$\begin{aligned}
&\min_{\varsigma} \quad \log E\left[\exp\left[\frac{1}{2}\left(\varsigma - \varsigma^2\right)(1-\gamma)\sigma_p^2\right] \cdot \exp\left[\varsigma(1-\gamma)(r_p - r_f)\right]\right] \\
\Leftrightarrow &\min_{\varsigma} \quad \frac{1}{2}\left(\varsigma - \varsigma^2\right)(1-\gamma)\sigma_p^2 + \varsigma(1-\gamma)(\mu_p - r_f) + \frac{1}{2}\varsigma^2(1-\gamma)^2\sigma_p^2 \\
\Leftrightarrow &\max_{\varsigma} \quad \varsigma\left(\mu_p - r_f + \frac{1}{2}\sigma_p^2\right) - \frac{1}{2}\varsigma^2\gamma\sigma_p^2,
\end{aligned}$$

which has the solution

$$\varsigma = \frac{\mu_p - r_f + \frac{1}{2}\sigma_p^2}{\gamma\sigma_p^2}.$$

To proceed, we assume that log returns and arithmetic returns are similar such that $\mu_p - r_f + \frac{1}{2}\sigma_p^2 \cong e^{\mu_p - r_f} \cong \mu_p - r_f$. We then determine μ_p and σ_p^2 for a portfolio that consists of the market investment plus a long-short momentum investment. Because the momentum portfolio is self-financing, feasible combinations of the market portfolio and the momentum portfolio are given by the weight vector $\boldsymbol{w}' = \begin{bmatrix} 1 & w_m \end{bmatrix}$, i.e., hold the market portfolio plus a proportionate long-short

[28]See also, e.g., Peress (2004), for the use of this approximate solution to the CRRA portfolio choice problem in a noisy rational expectations setting.

momentum overlay w_m. The optimal risky portfolio w_m solves the constrained optimization

$$\min_{\boldsymbol{w}} \quad \frac{\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}}{2}, \qquad \text{s.t.} \qquad \boldsymbol{\mu}'\boldsymbol{w} = r^* - r_f,$$

where

$$\boldsymbol{w} = \begin{bmatrix} 1 \\ w_m \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} r - r_f \\ E_{type}\left[m+\epsilon\right] \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_\chi^2 & 0 \\ 0 & Var_{type}\left[m+\epsilon\right] \end{bmatrix},$$

and $r^* - r_f$ is a target return premium that traces out the efficient frontier, and r is the required return on the market portfolio. The solution is

$$\text{(IA.B.2)} \qquad w_m = \frac{E_{type}\left[m+\epsilon\right] / Var_{type}\left[m+\epsilon\right]}{\left(r - r_f\right) \Big/ \sigma_\chi^2}.$$

Using equation (IA.B.2), the parameters of the optimal risky portfolio are

$$\text{(IA.B.3)} \qquad \mu_p - r_f = \begin{bmatrix} r - r_f & E_{type}\left[m+\epsilon\right] \end{bmatrix} \begin{bmatrix} 1 \\ w_m \end{bmatrix} = r - r_f + w_m E_{type}\left[m+\epsilon\right], \text{ and}$$

$$\text{(IA.B.4)} \qquad \sigma_p^2 = \boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w} = \sigma_\chi^2 + w_m^2 Var_{type}\left[m+\epsilon\right] = \frac{\sigma_\chi^2}{r - r_f}\left(r - r_f + w_m E_{type}\left[m+\epsilon\right]\right).$$

Thus,

$$\text{(IA.B.5)} \qquad \varsigma = \frac{r - r_f}{\gamma \sigma_\chi^2}.$$

Combining Eqs. (IA.B.2) and (IA.B.5),

$$Demand = w_m \varsigma K_{type,2} = \frac{E_{type}\left[m+\epsilon\right]}{\gamma Var_{type}\left[m+\epsilon\right]} K_{type,2}.$$

IA.C. Negative Market-clearing Price for Momentum Portfolio

In the case of $k_M > \lambda$, the demand of momentum investors increases with a positive f faster than the supply can keep up with, implying an increasingly large buying imbalance as f rises (depicted in Figure IA.3, Plot C.1). This again suggests that momentum investors buy up to their capacity, leading to a subsequent momentum crash.

However, when $k_M > \lambda$ there is also a (finite) negative value for f that clears the market. While we discount this equilibrium as implausible, we note that even here the contrary pricing of winner and loser stocks implies a substantial negative momentum return because the formation-period 'winners' are actually the fundamental losers, and vice versa.

It is not clear how this $f < 0$ equilibrium could be found, because informed investors presumably seed formation-period returns with *buying* of the momentum portfolio (and an initially positive f). Nevertheless, it is a call auction and if they were to bizarrely trade contrary to their private information, seeding a negative value for f, then they might induce momentum investors into selling (buying) so much winner (loser) stock that their bizarre trade is preferred.

IA.D. Probability Density Function Conditional on f

Below we derive the expression for $p\,(\delta|f)$. By the definition of conditional probability, we have

$$p\,(\delta|f) = \frac{p_2(\delta, f)}{p_1(f)},$$

where numerical subscripts distinguish the functional form of each probability density function. To solve for these densities, we exchange the primitive random variable k_I with the observable random variable f. Formally, let

$$F:(\delta,k_I,k_M)\rightarrow f=\frac{1}{D}\left(\delta k_I+\frac{\delta^E}{1+\frac{\delta^V}{\sigma_\epsilon^2}}k_M\right),$$

which characterizes market clearing as in equation (4). We map the primitive random variables into

$$\begin{pmatrix}\delta\\ \\ k_M\\ \\ k_I\end{pmatrix}\rightarrow\begin{pmatrix}\delta\\ \\ k_M\\ \\ F\left(\delta,k_M,k_I\right)\end{pmatrix}.$$

Next, we need

$$\mid \mathrm{J}\mid=\det\begin{pmatrix}\frac{\partial\delta}{\partial\delta} & \frac{\partial\delta}{\partial k_M} & \frac{\partial\delta}{\partial k_I}\\ \\ \frac{\partial k_M}{\partial\delta} & \frac{\partial k_M}{\partial k_M} & \frac{\partial k_M}{\partial k_I}\\ \\ \frac{\partial F^{-1}}{\partial\delta} & \frac{\partial F^{-1}}{\partial k_M} & \frac{\partial F^{-1}}{\partial k_I}\end{pmatrix}=\frac{D}{\delta}.$$

Following a standard result (see, e.g., Theorem 2 in Section 4.4 of Rohatgi and Saleh, 2000), the density is then given by

$$\begin{aligned}p_3\left(\delta,k_M,f\right)&=g\left(\delta\right)h\left(k_M,F^{-1}\right)\mid \mathrm{J}\mid\\ &=g\left(\delta\right)h\left(k_M,\frac{1}{\delta}\left(fD-\frac{\delta^E}{1+\frac{\delta^V}{\sigma_\epsilon^2}}k_M\right)\right)\frac{D}{\delta},\end{aligned}\tag{IA.D.1}$$

where D is as in equation (4). Integrating k_M and then δ out of equation (IA.D.1) gives

$$p_2(\delta,f)=\frac{g\left(\delta\right)}{\delta}\int_0^1 h\left(k_M,\frac{1}{\delta}\left(fD-\frac{\delta^E}{1+\frac{\delta^V}{\sigma_\epsilon^2}}k_M\right)\right)Ddk_M,$$

$$p_1(f)=\int_0^\infty\frac{g\left(\delta\right)}{\delta}\int_0^1 h\left(k_M,\frac{1}{\delta}\left(fD-\frac{\delta^E}{1+\frac{\delta^V}{\sigma_\epsilon^2}}k_M\right)\right)Ddk_Md\delta.$$

We then obtain

$$p(\delta|f) = \frac{\frac{g(\delta)}{\delta}\int_0^1 h\left(k_M, \frac{1}{\delta}\left(fD - \frac{\delta^E}{1+\frac{\delta^V}{\sigma_\epsilon^2}}k_M\right)\right)Ddk_M}{\int_0^\infty \frac{g(\delta)}{\delta}\int_0^1 h\left(k_M, \frac{1}{\delta}\left(fD - \frac{\delta^E}{1+\frac{\delta^V}{\sigma_\epsilon^2}}k_M\right)\right)Ddk_M d\delta}.$$

IA.E. Additional simulation results

This section of the internet appendix investigates the impact of changing the distributional assumptions for δ and higher concentration parameters in the simulation analysis of Section III. It also analyses the relation between expected momentum returns and unexpected momentum capital in the different simulations.

First, we ask whether our results are robust to using a higher concentration parameter for the Dirichlet distribution and a uniform distribution for δ instead of a log-normal distribution. In particular, we let δ follow a uniform distribution on $[0.06,\ 0.12]$, and let $\alpha_i = 12$. The results are reported in Figure IA.1 and Table IA.1. In summary, the results are very similar to those in Section III. In the myopic beliefs case, momentum returns again have pronounced negative skewness, high volatility and large excess kurtosis, and they are well behaved with low volatility, slightly positive skewness, and no excess kurtosis in the rational beliefs case.

Second, we ask whether the beliefs specifications for unknown capital become more similar to the known capital case when $var(k)$ is very small. To achieve this, we set the concentration parameters $\alpha_i = 60$ in the Dirichlet distribution, and leave the setting otherwise identical to the one in the paper. The results in Figure IA.2 and Table IA.2 verify that crashes disappear in the myopic beliefs case once capital uncertainty is negligible. Momentum returns in all four specifications are now well behaved and have similar return characteristics.

Finally, we analyze the relation between unexpected momentum capital $k_M - Ek_M$ and expected momentum returns m. To do this, we rank the simulation trials for the different specifications in Figure 1, Figure IA.1, and Figure IA.2 into 100 bins according to $k_M - Ek_M$, and report the averages within each bin to approximate a conditional expectation. All nine subplots of Figure IA.3 corresponding to different distributional and beliefs assumptions show that the residual information m not incorporated into prices decreases with crowd size. Thus, the model supports the negative relation between momentum capital and expected momentum returns we document empirically.

References

D. Andrei and J. Cujean. Information percolation, momentum and reversal. *Journal of Financial Economics*, 123(3):617–645, 2017.

C. Antoniou, J. A. Doukas, and A. Subrahmanyam. Cognitive dissonance, sentiment, and momentum. *Journal of Financial and Quantitative Analysis*, 48(1):245–275, 2013.

C. S. Asness, T. J. Moskowitz, and L. H. Pedersen. Value and momentum everywhere. *The Journal of Finance*, 68(3):929–985, 2013.

D. Avramov, S. Cheng, and A. Hameed. Time-varying liquidity and momentum profits. *Journal of Financial and Quantitative Analysis*, 51(6):1897–1923, 2016.

M. Baltzer, S. Jank, and E. Smajlbegovic. Who trades on momentum? *Journal of Financial Markets*, 42:56–74, 2019.

N. Barberis, A. Shleifer, and R. Vishny. A model of investors sentiment. *Journal of Financial Economics*, 49:307–343, 1998.

P. Barroso and P. Santa-Clara. Momentum has its moments. *Journal of Financial Economics*, 116 (1):111–120, 2015.

J. Y. Campbell and L. M. Viceira. *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*. Oxford University Press, 2002.

M. J. Cooper, R. C. Gutierrez, and A. Hameed. Market states and momentum. *The Journal of Finance*, 59(3):1345–1365, 2004.

Z. Da, U. G. Gurun, and M. Warachka. Frog in the pan: Continuous information and momentum. *Review of Financial Studies*, 22(7):2171–2218, 2014.

K. Daniel and T. J. Moskowitz. Momentum crashes. *Journal of Financial Economics*, 122(2): 221–247, 2016.

K. Daniel, D. Hirshleifer, and A. Subrahmanyam. Investor psychology and security market under- and overreactions. *The Journal of Finance*, 53:1839–1885, 1998.

J. B. De Long, A. Shleifer, L. H. Summers, and R. J. Waldmann. Noise trader risk in financial markets. *Journal of Political Economy*, 98(4):703–738, 1990a.

J. B. De Long, A. Shleifer, L. H. Summers, and R. J. Waldmann. Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance*, 45(2):379–395, 1990b.

V. DeMiguel, A. Marttín-Utrera, F. J. Nogales, and R. Uppal. A Transaction-Cost Perspective on the Multitude of Firm Characteristics. *The Review of Financial Studies*, 33(5):2180–2222, 04 2020.

R. M. Edelen, O. S. Ince, and G. B. Kadlec. Institutional investors and stock return anomalies. *Journal of Financial Economics*, 119(3):472–488, 2016.

K. A. Froot, D. S. Scharfstein, and J. C. Stein. Herd on the street: Informational inefficiencies in a market with short-term speculation. *The Journal of Finance*, 47(4):1461–1484, 1992.

M. Grinblatt, S. Titman, and R. Wermers. Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior. *The American Economic Review*, 85(5): 1088–1105, 1995.

K. Grobys, J. Ruotsalainen, and J. Äijö. Risk-managed industry momentum and momentum crashes. *Quantitative Finance*, 18(10):1715–1733, 2018.

S. G. Hanson and A. Sunderam. The growth and limits of arbitrage: Evidence from short interest. *Review of Financial Studies*, 27(4):1238–1286, 2014.

A. Hillert, H. Jacobs, and S. Müller. Media makes momentum. *The Review of Financial Studies*, 27 (12):3467–3501, 2014.

H. Hong and J. C. Stein. A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance*, 54:2143–2184, 1999.

S. Huang. The momentum gap and return predictability. *Working Paper SSRN*, 2015.

S. Hvidkjaer. A trade-based analysis of momentum. *Review of Financial Studies*, 19(2):457–491, 2006.

H. Jacobs, T. Regele, and M. Weber. Expected skewness and momentum. Working Paper SSRN, 2015.

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

P. Kondor and A. Zawadowski. Learning in crowded markets. *Journal of Economic Theory*, 184: 104936, 2019.

R. A. Levy. Relative strength as a criterion for investment selection. *The Journal of Finance*, 22(4): 595–610, 1967.

J. Lewellen. Institutional investors and the limits of arbitrage. *Journal of Financial Economics*, 102 (1):62–80, 2011.

D. Lou and C. Polk. Comomentum: Inferring arbitrage activity from return correlations. Working Paper London School of Economics, 2013.

R. Novy-Marx. Fundamentally, momentum is fundamental momentum. Working Paper 20984, National Bureau of Economic Research, 2015.

J. Peress. Wealth, information acquisition, and portfolio choice. *The Review of Financial Studies*, 17(3):879–914, 2004.

V. K. Rohatgi and A. K. M. E. Saleh. *Multiple Random Variables*, chapter 4, pages 102–179. John Wiley & Sons, Inc., 2000.

K. G. Rouwenhorst. International momentum strategies. *The Journal of Finance*, 53(1):267–284, 1998.

D. S. Scharfstein and J. C. Stein. Herd behavior and investment. *Amercian Economic Review*, 80 (3):465–479, 1990.

J. M. Sotes-Paladino, G. J. Wang, and C. Y. Yao. The value of growth: Changes in profitability and future stock returns. Working Paper SSRN, 2016.

J. C. Stein. Informational externalities and welfare-reducing speculation. *Journal of Political Economy*, 95(6):1123–1145, 1987.

J. C. Stein. Presidential address: Sophisticated investors and market efficiency. *The Journal of Finance*, 64(4):1517–1548, 2009.

D. Vayanos and P. Woolley. An institutional theory of momentum and reversal. *Review of Financial Studies*, 26(5):1087–1145, 2013.

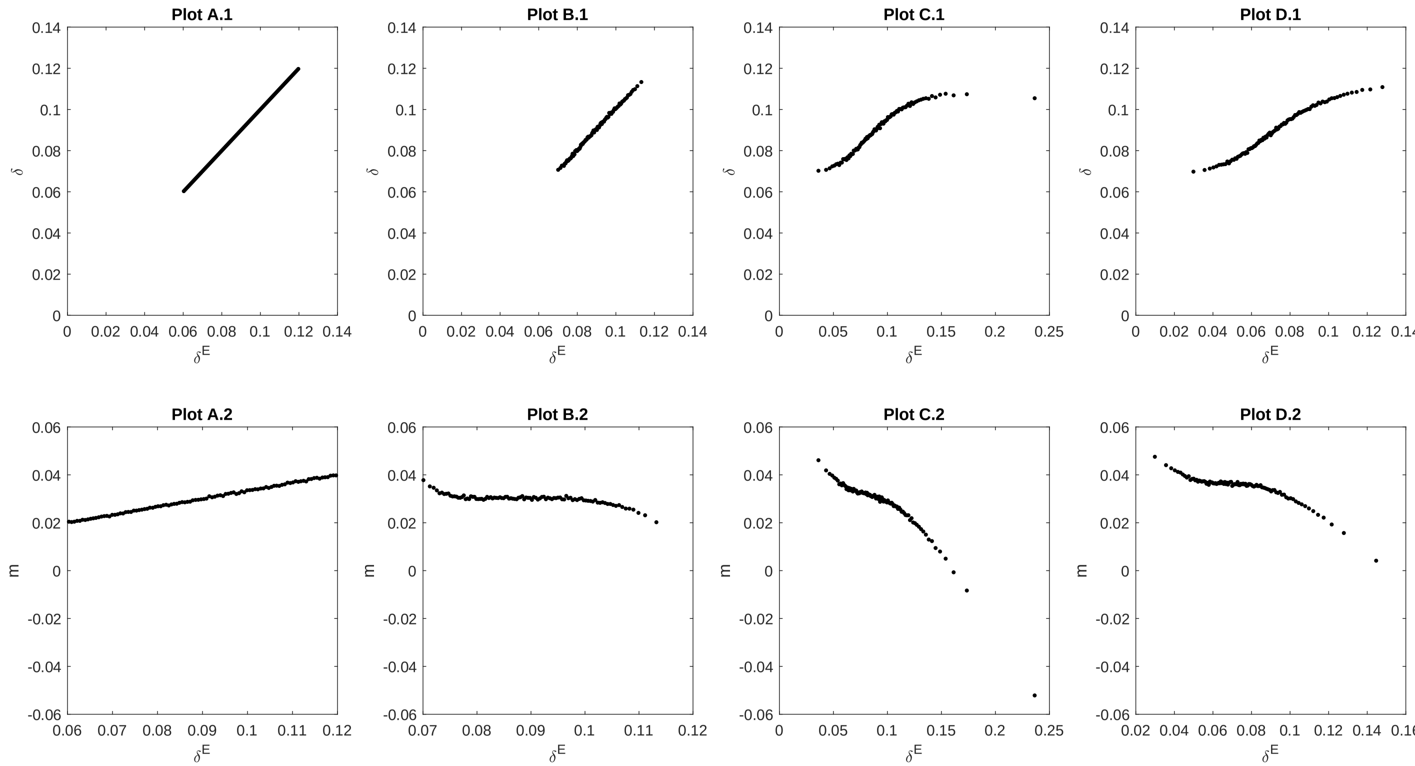


Figure IA.1: Beliefs and expected momentum returns in simulations—uniform distribution

This figure is constructed in the same way as Figure 1 in the paper. In particular, the simulations use 100,000 independent random draws of $\{k_I, k_M, \delta\}$, where k_I and k_M are informed and momentum capital, respectively, and δ is the signal of differential fundamental value for winners minus losers. k_I and k_M (and k_C) follow a Dirichlet distribution with concentration parameters $\alpha_i = 12$, and $Ek_I = Ek_M = 1/3$. δ follows a uniform distribution on [0.06 0.12]. The market clearing formation period return f is solved for each $\{k_I, k_M, \delta\}$ pair by iteration using different specifications for momentum traders' beliefs δ^E: known crowding in Plots A.1-2, rational beliefs in Plots B.1-2, myopic beliefs in Plots C.1-2, and optimal linear beliefs in Plots D.1-2. The expected momentum return is then $m = \delta - f$. The values of $\{\delta, m\}$ are ranked into 100 equally populated bins according to the horizontal-axis variable, and the plots represent the averages for the indicated variables within these bins.

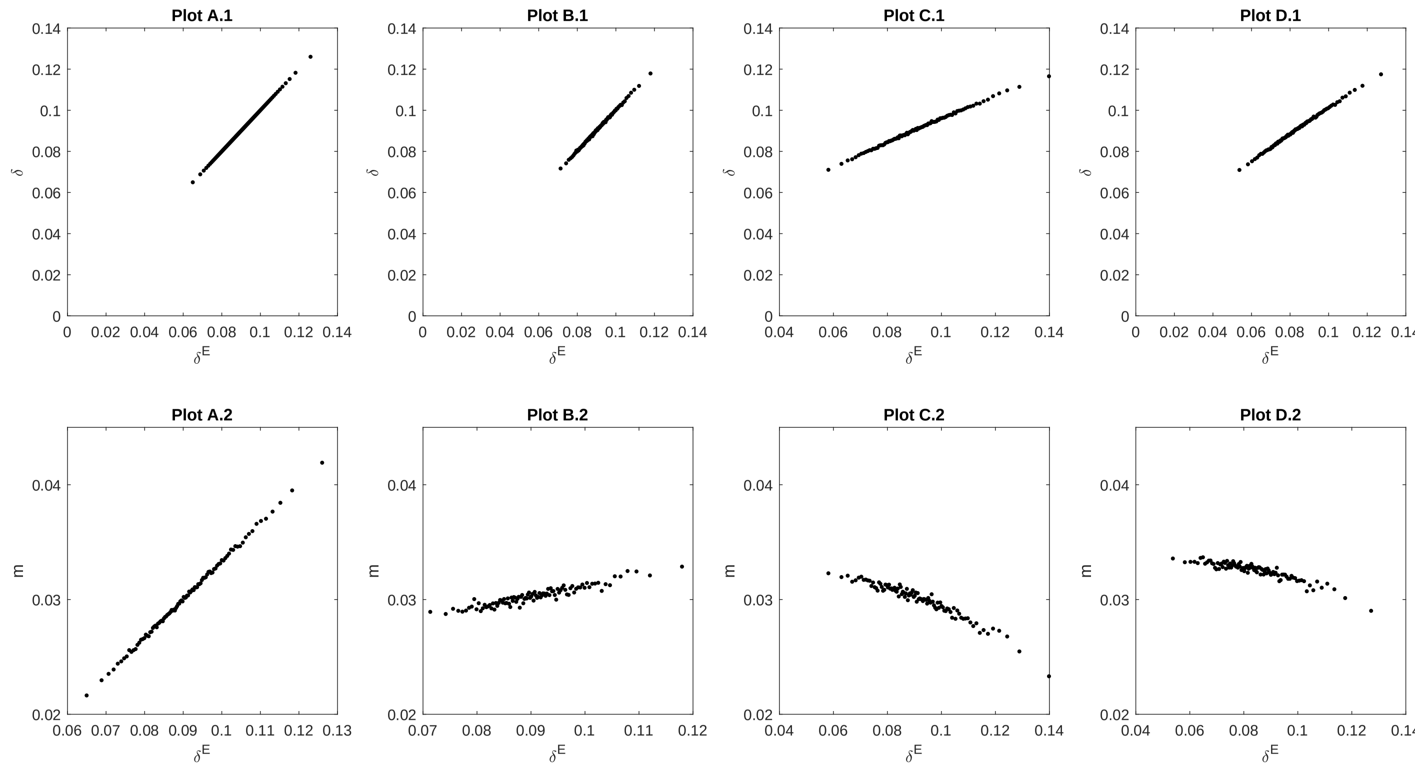


Figure IA.2: Beliefs and expected momentum returns in simulations—very low $var(k)$

This figure is constructed in the same way as Figure 1 in the paper, and the simulations are identical to those in Figure 1 except that k_I and k_M (and k_C) follow a Dirichlet distribution with concentration parameters $\alpha_i = 60$, which still implies $Ek_I = Ek_M = 1/3$. In addition, as opposed to Figure IA.1 δ follows a log-normal distribution with $\mu = -2.405$ and $\sigma = 0.125$ implying an average δ of 9.1% with standard deviation of 1.14% as in the paper.

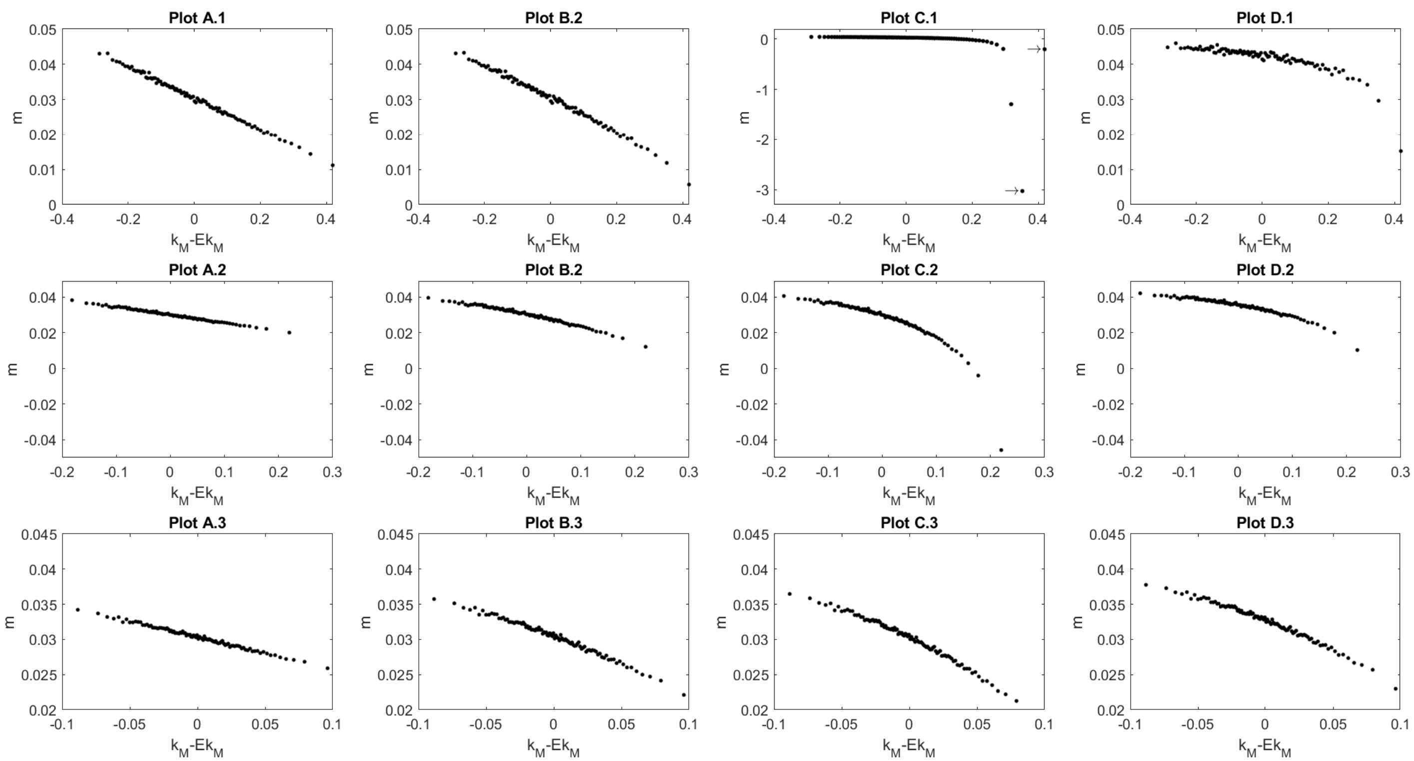


Figure IA.3: Expected momentum returns and unexpected momentum capital

Rows one, two, and three correspond to the simulations of Figure 1, Figure IA.1, and Figure IA.2, respectively. For each simulation $\{m, k_M\}$ are ranked into 100 equally populated bins according to the horizontal-axis variable, and the plots represent the averages for the indicated variables within these bins.

Table IA.1: Momentum returns in simulations—uniform distribution

The table reports unconditional return statistics for the simulations described in the caption of Figure IA.1. Panel A contains the descriptive statistics of expected momentum returns across all simulations. Mean, stdev, skew, kurt, min and max refer to average, standard deviation, skewness, kurtosis, minimum, and maximum, respectively. The simulations are computed under the indicated belief specification, and the Dirichlet distribution has the concentration parameters $\alpha_i = 12$, and δ follows a uniform distribution on $[0.06,\ 0.12]$. Optimal linear beliefs are chosen to maximize the utility of a CRRA investor with $\gamma = 2$, and they are reported in the row λ^{-1}. Profits are likewise the expected portfolio returns of a $\gamma = 2$ investor, and certainty equivalents 'cer(γ)' are calculated for $\gamma = 2, 4, 10$, with portfolio weights calculated as in (3). Realized momentum returns in Panel B are given by $m + \epsilon$ where ϵ is randomly drawn from a zero-mean normal distribution with standard deviation 0.125. Cer(γ) is an arithmetic return, and all other statistics are based on log returns.

Belief spec.	known	rational	myopic	optimal linear
λ^{-1}			1.50	1.34
Panel A. Expected momentum returns m				
mean	3.0%	3.0%	2.8%	3.5%
stdev	0.9%	1.4%	2.7%	1.4%
skew	0.5	0.3	-92.9	-0.1
kurt	3.1	2.9	17386.3	9.6
min	0.61%	-2.04%	-534.41%	-35.57%
max	7.51%	8.56%	9.22%	9.38%
Panel B. Realized momentum returns $m + \epsilon$				
profit	3.11%	2.78%	1.75%	1.99%
cer(2)	2.30%	2.16%	-100.00%	1.85%
cer(5)	1.14%	1.07%	-100.00%	0.92%
cer(10)	0.45%	0.43%	-100.00%	0.37%

Table IA.2: Momentum returns in simulations—very low $var(k)$

The table reports unconditional return statistics for the simulations described in the caption of Figure IA.2 and is constructed in the same fashion as Table IA.1.

Belief spec.	known	rational	myopic	optimal linear
λ^{-1}			1.50	1.44
Panel A. Expected momentum returns m				
mean	3.0%	3.0%	3.0%	3.3%
stdev	0.5%	0.7%	0.8%	0.7%
skew	0.4	0.3	0.2	0.3
kurt	3.4	3.3	3.3	3.3
min	1.39%	0.18%	-0.39%	0.41%
max	6.13%	7.35%	7.12%	7.35%
Panel B. Realized momentum returns $m + \epsilon$				
profit	3.02%	2.91%	2.90%	2.64%
cer(2)	2.29%	2.24%	2.15%	2.24%
cer(5)	1.14%	1.11%	1.07%	1.11%
cer(10)	0.45%	0.44%	0.43%	0.44%