Institutional Crowding and the Moments of Momentum

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

November 18, 2018

Abstract

Several theoretical studies suggest that crowding in unanchored investment strategies can generate feedback effects that cause asset prices to deviate substantially from fundamental value. We demonstrate that consequent runaway overvaluation can explain momentum crashes if investors hold fixed beliefs regarding peer crowding. However, when investors condition on prices to rationally infer crowd size they choose nonlinear demands that short circuit crowding induced feedback, preventing overvaluation and crashes. We use proxies constructed from 13F holdings data and find little evidence of a causal link between crowding and momentum tail risk. Indeed, we find that unattractive higher return moments are predicted by *low* rather than high institutional crowding, consistent with momentum demands rationally considering feedback effects from crowding.

Keywords: Crowded trade, destabilize, momentum, institutional investors, crash risk

*Barroso is at the University of New South Wales, Edelen is at Virginia Tech, and Karehnke is at ESCP Europe. Emails: p.barroso@unsw.edu.au, edelenr@vt.edu, pkarehnke@escpeurope.eu. We thank 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, Manchester's Hedge Fund Conference, the Portuguese Finance Network Conference, Rotterdam's Professional Asset Management Conference, and the TCU Finance conference for their helpful comments and suggestions. We are especially grateful for comments from Jeff Busse, John Griffin, Hening Liu, Rick Sias, Neal Stoughton, and Xuemin Yan. Semyon Malamud was particularly helpful to the paper's development. Errors are the authors' responsibility.

1 Introduction

Several studies argue that arbitrageur capital can destabilize prices relative to fundamental value as a result of incomplete information regarding the market setting. For example, Abreu and Brunnermeier (2003) show that if arbitrageurs cannot coordinate an attack, financial bubbles can be advanced by arbitrageurs themselves, who trade against fundamental value to extrapolate or 'ride' the bubble. Stein (2009) shows that uncertainty regarding the crowding of other arbitrageurs running similar unanchored strategies can lead to overstated beliefs that push prices away from fundamental value. When destabilization is sufficiently extreme crash risk is plausibly high, suggesting a link between crowded trades and higher return moments by way of valuation effects.

A link between crowded trades and higher return moments has been examined in several empirical studies, e.g., Khandani and Lo (2007), Pedersen (2009), and Sias, Turtle, and Zykaj (2017). However, the focus in these studies is liquidity spirals rather than equilibrium valuation effects, and the scope is narrow. Lewellen (2011) and Edelen et al. (2016) consider a broader scope of aggregate institutional holdings and changes in holdings (respectively) as they relate to a broad array of anomaly returns, including momentum. But the focus there is neither higher return moments nor potential price destabilization from crowding. Other studies have used return characteristics to link higher moments of momentum returns to crowding, but the link to trading is indirect. Thus, while there is both a theoretical and empirical basis in the literature for the view that unanticipated crowding might drive prices beyond fundamental value, leading to momentum tail risk, the hypothesis has not been comprehensively evaluated.[1] We directly investigate this hypothesis in the context of momentum, from both a theoretical and an empirical perspective.

[1]Several authors (see, e.g., Khandani and Lo, 2007; Pedersen, 2009) have related crowding to the 'quant meltdown' of 2007 by way of a funding / liquidity spiral. These studies do not argue that crowding precipitated the meltdown; rather, that crowding was a necessary precursor for an isolated funding shock to exacerbate into a systemic crash. That is, crowding contributes to funding / liquidity risk or fragility; not overvaluation. See also Brunnermeier and Pedersen (2009). By contrast, the focus in this study does not consider systemic risk but rather focuses on crowding as a source of valuation errors. For studies relating to return characteristics see Lou and Polk (2013) and Huang (2015), who both argue that crowding by momentum investors potentially explains negative skewness in momentum returns.

We first develop a model similar in spirit to Stein (2009) to demonstrate how different belief mechanisms employed by momentum investors lead to dramatically different predictions regarding the impact of crowding on the valuation of (and returns to) momentum portfolios. In addition to the linear setting considered in Stein's analysis, we consider more sophisticated demands in which arbitrageurs condition on the market clearing relation and prior beliefs for crowding and fundamental value, forming rational posterior beliefs of crowding. These demands reflect arbitrageurs who are aware of—and defend against—crowding induced overvaluation and crash risk.

We use simulations under each specification of beliefs to demonstrate the key role that myopic arbitrageurs (ignorant of the potential for crowding-induced crashes) play in theoretically justifying unanticipated crowding as a source of crash risk. When arbitrageur beliefs rationally incorporate inferences of crowding from market-clearing prices, the tail risk in simulated momentum returns is virtually identical to the case of known crowding (i.e. the primitive return distribution). But when arbitrageur demands do not condition on the potential adverse effects of peer crowding, simulated momentum returns exhibit substantial negative skewness and excess kurtosis. We conclude that unanticipated crowding is a plausible source of momentum tail risk, but only to the extent that momentum investors are myopic to that possibility in constructing their demands.

The key aspect of the more sophisticated, rational solution is that momentum investors directly compute the conditional mean for fundamental value at *each* price for momentum assets (a single portfolio in our model), rather than (erroneously) presuming that the conditional mean follows a linear extrapolation rule. Obfuscation from unanticipated crowding generates nonlinear feedback, which implies a nonlinear error to any linear inference of fundamental value. Rational adjustment means that conditional expectations nonlinearly relate to price, and therefore so do demands.

Thus, we solve for these demands by computing the conditional mean and variance of fundamental value at a given price for the momentum portfolio; repeating over a grid of prices to interpolate a (nonlinear) solution. Both parameters condition on market clearing and a correct statement of the joint distribution for the two key primitives: crowding (i.e., the number of peers following the same investment strategy) and fundamental value. Demands based on these parameters yield a fixed point on the momentum-portfolio price by construction. Unanticipated crowding imparts noise into that equilibrium price, but because arbitrageurs form correct conditional expectations at every price, there is no potential for destabilizing feedback: arbitrageurs correctly infer the potential that unanticipated crowding has influenced the price.

This result is seen in our simulations of market equilibrium under rational versus myopic beliefs (meaning, complete myopia regarding the possibility of feedback-trading on unanticipated crowding). Our default calibration yields a (log) momentum rate of return (per evaluation period) with the following characteristics when modeled under myopic versus rational beliefs, respectively:

- Mean of -2.4% versus 3.0%, and a minimum of $-39,000\%$ versus -2.6%,
- Standard deviation of 174% versus 1.6%,
- Skewness of -151.3 versus 0.4, and kurtosis of $30,000$ versus 3.0,
- Certainty equivalent return of *complete loss* versus 2.5%.

Thus, unanticipated crowding can lead to virtually unbounded crash risk in strategy returns if investors following the strategy are myopic to its influence on pricing. But unanticipated crowding contributes no crash risk in an otherwise identical setting with rational beliefs. (The benchmark skewness in this setting with known crowding is 0.6; kurtosis 3.1; and standard deviation 1.4%.)

The implied tail risk in the above simulations under myopic beliefs has a devastating effect on momentum investors. For example, the expected rate of profit for momentum investors maintaining rational beliefs in the above simulations is 3.44% per evaluation period (e.g., quarter if matched to the 13F data). This compares favorably to a benchmark expected rate of return (case of momentum investing with known crowding) of 3.65%. By contrast, under myopic beliefs the expected profit is virtually certain complete loss due to the high probability of momentum crash.

While the case of myopic beliefs demonstrates how unanticipated crowding can generate momentum crashes, such beliefs make for an untenable strategy. One alternative is to scale back the myopia inherent in linear demands by lessening the slope in the linear inference of predictable returns given price. Stein (2009) models arbitrageurs taking this approach to protecting against crash risk. We conduct a similar analysis in our setting using a grid search over a range of assumed slopes. We find that when momentum investors follow the optimum such strategy, crash risk is substantially eliminated. However, this approach is inefficient: linearity implies foregoing profitable opportunity when rational inference of crowding is low to moderate, and continued feedback trading when rational inference of crowding is high. As a result of this inefficiency, the expected rate of profit from the optimal linear strategy is a relatively modest 0.65%; i.e., capturing only 17% of potential profits of 3.65% in the known-crowding case (versus a 95% effectiveness using rational beliefs). Moreover, some residual equilibrium crash risk remains.

In summary, if investors are myopic to crowd risk then their demands generate the very risk they ignore: unanticipated crowding in the momentum strategy generates momentum crashes. Conversely, if investors' demands rationally account for the potential impact of unanticipated crowding on prices, then those self-protecting demands attenuate the influence of unanticipated crowding to the point that the equilibrium distribution of momentum returns precludes crashes. In short, theoretical predictions of crowding-induced momentum crashes do not derive from crowding uncertainty *per se*, but rather myopic consideration of that crowding uncertainty. We don't take a stand on the extent of crowding myopia in practice, instead treating it as an empirical matter.

Our empirical analysis uses 13F data to construct proxies for momentum crowding by directly linking changes in institutional holdings to the now-standard 12 - 1 momentum prescription of, i.e., Jegadeesh and Titman (1993); and to past returns generally as in the analysis of mutual fund momentum trading in Grinblatt et al. (1995). Our measures incorporate three innovations. First, we incorporate persistence to better distinguish investment strategy from spurious trade-return correlations. Second, we consider both anticipated and unanticipated crowding measures. Third, we distinguish between crowding by peer institutions (number of momentum investors)—which we argue is the dominant source of uncertainty—versus trading intensity (capital allocated to momentum) which is arguably relatively homogeneous across investors. 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 this procedure relative to direct inference from trading behavior.

We find strong evidence that crowding predicts negatively mean momentum returns, which is consistent with theoretical prediction under all belief specifications. However, we find little in the way of reliable evidence that crowding positively predicts tail risk. To be meaningful, tail risk implies negative skewness, elevated volatility, and excess kurtosis; not just negative skewness. We find that our proxies for momentum crowding generally relate negatively to all three higher moments,[2] often with statistical reliability. This surprising result is not consistent with a causal role for crowding in momentum crashes, but it is consistent with rational momentum investors optimizing their demands to account for time-varying toxicity in market conditions. We provide a broad range of supporting evidence for this assessment of arbitrageur beliefs, based on nuanced differences in the proxies. We conclude that the evidence best supports the theoretical analysis that presumes rational beliefs. In that theoretical setting, crowding does not play a material role for higher moments of momentum returns.

There is much related literature on the subject. We provide a detailed survey in Appendix A. Section 2 develops the model and Section 3 develops and analyzes its result using a simulation approach. Section 4 presents the empirical analyses and Section 5 concludes the study.

2 Model

Section 2.1 lays out the assumptions and setting of the model and Section 2.2 develops four solutions to the equilibrium, differing by how momentum investors form their beliefs.

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

2.1 Setting

To save space we work directly with the momentum portfolio, which is developed from individual stocks in Appendix B. The key components in that development are as follows. There are two periods: the momentum portfolio formation period and the evaluation period in which momentum returns are realized. The formation period begins with all investors holding the market portfolio. An informed subset of investors observe a common signal δ of differential fundamental value for a random subset of winner and loser stocks. The realized differential return is $d = \delta + \epsilon$ where ϵ is a mean-zero disturbance with variance σ_ϵ^2. As informed investors trade on their signal in the formation period, they identify the momentum portfolio.

In practice, momentum investors infer informed traders' signal by observing strictly past returns; generally a six to twelve month period. We abstract from this literal chronology by modeling a call auction on the portfolio formation date (end of formation period) in which all agents condition demands on the market-clearing price. Payoffs are realized after the market clears, which forms the evaluation period. Let f denote the formation-period return on winner minus loser stocks.[3] We refer to f as the price of the momentum portfolio. The realized (evaluation period) momentum return is

$$d - f = m + \epsilon, \tag{1}$$

where the momentum return $m = \delta - f$ is the expected return conditional on all available information as of the portfolio formation date, integrating out the disturbance ϵ. Most of our analysis pertains to m, as ex post residuals are not relevant to empirical predictions.

The price of the momentum portfolio is determined by balancing the demands of three investor groups. Informed investors hold an initial capital stock K_I. They observe δ and hold beliefs

$$E_I\left[m + \epsilon\,|\delta, f\right] = \delta - f, \qquad Var_I\left[m + \epsilon\,|\delta, f\right] = \sigma_\epsilon^2. \tag{2}$$

[3]More specifically, the mean of the log of the call-auction price minus the initial price for winner stocks, minus the same calculation for loser stocks. Because each stock's private signal is +/- δ, the return on each stock is the same except possibly sign.

Momentum investors hold capital stock K_M. They do not observe δ but they attempt to infer it from the market clearing price, f. Let $\delta^E = E_M[\delta|f]$ and $\delta^V = Var_M[\delta|f]$ denote their expectations (further discussed in Section 2.2)

$$E_M[m+\epsilon|f] = \delta^E - f, \qquad Var_M[m+\epsilon|f] = \delta^V + \sigma_\epsilon^2. \tag{3}$$

Counterparty investors hold capital stock K_C. They trade counter to the price-pressure of informed and momentum investors, fixating beliefs on historical public information with $E_C d = 0$. Thus:

$$E_C[m+\epsilon|f] = -f, \qquad Var_C[m+\epsilon|f] = \sigma_\delta^2 + \sigma_\epsilon^2. \tag{4}$$

Counterparty investors obviously lose in expectation; their role is to clear the market without burdening the analysis with a noise-trader framework. All investors hold power utility preferences with relative risk aversion γ, maximizing[4]

$$E[u(K)] = E\left[\frac{K^{1-\gamma}}{1-\gamma}\right] \tag{5}$$

on the portfolio-formation date, where K is wealth following the evaluation period. We use the second-order approximation of Campbell and Viceira (2002, Internet Appendix) to derive

$$Demand = \frac{E_{type}[m+\epsilon]}{\gamma Var_{type}[m+\epsilon]} K_{type}, \tag{6}$$

where K_{type} is the beginning of period capital of a given investor type. Details are in Appendix C. Summing across the three investor types and equating to zero supply gives the market clearing condition:

$$f = \frac{1}{D}\left(\delta k_I + \frac{\delta^E}{1+\frac{\delta^V}{\sigma_\epsilon^2}} k_M\right), \tag{7}$$

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

where $D = \left(1 - \frac{\delta^V}{\sigma_\epsilon^2+\delta^V} k_M\right)$ and $k_{type} = K_{type} / \left(\frac{\sigma_\epsilon^2}{\sigma_\epsilon^2+\sigma_\delta^2} K_C + K_I + K_M\right)$ indicates the fraction of capital from each investor type (with a nuisance adjustment multiplying counterparty capital). Crowding in the momentum strategy is captured by k_M, which is assumed random (likewise, k_I, k_C).[5]

2.2 Solutions

We consider four solutions to Eq. (7) that differ in the specification of momentum-investor beliefs. In the base case the capital allocated to momentum is known, so momentum investors correctly formulate linear beliefs for $m|f$. In the second case crowding is stochastic but momentum investors ignore this fact and treat it as fixed at the expected value, again applying linear beliefs to $m|f$. This is the case of extreme myopia where endogenous investment decisions are not recognized as a source of noise in market prices. In the third case momentum investors maintain linearity in their beliefs, but they choose the slope (rather than their demands in each market occurrence) based on the average gains from predictable returns and the average loss from feedback trading on capital uncertainty. In the fourth case momentum investors use the common prior probability distributions for δ, k_M, and k_I to form a conditional distribution for $m|f$ that correctly assimilates equilibrium crowding effects.

2.2.1 Known crowding

First consider the case of known k_M and k_I. If momentum investors conjecture a linear equilibrium

$$f = \lambda\delta \tag{8}$$

with $\lambda \equiv k_I + k_M$, then beliefs $\delta^E = \lambda^{-1} f$ and $\delta^V = 0$ lead to a self-fulfilling linear solution to Eq. (7).[6] That is, Eq. (8) gives the resulting dependence of price on δ, and $\lambda^{-1} f$ and 0 are the

[5]Note that crowding uncertainty is not the same as aggregate demand for the momentum portfolio. This distinction is important for the empirical section. Crowding uncertainty derives from not knowing how many others are following the same strategy; how they trade conditional on that strategy is common knowledge since they all solve the same problem).

[6]Note that Eq. (8) implies that f reveals δ.

conditional expectation and variance, respectively.

2.2.2 Myopic beliefs (crowding uncertainty that momentum investors ignore)

We now assume that capital proportions are stochastic, but momentum investors treat both as constants equal to their expected values. Hence they form beliefs using

$$f = \lambda_E \delta, \qquad \text{where} \qquad \lambda_E \equiv Ek_M + Ek_I. \tag{9}$$

We refer to this belief mechanism as 'myopic beliefs,' under which the market clears at the price

$$f = \lambda_E \left(\frac{k_I}{\lambda_E - k_M} \right) \delta = \lambda_E \left(\frac{k_I}{Ek_I - (k_M - Ek_M)} \right) \delta. \tag{10}$$

Eq. (10) identifies the problem with failing to account for capital uncertainty in forming momentum demands. While the bracketed multiplier in Eq. (10) equals one when crowding in the momentum strategy happens to equal its expected value, confirming the Eq. (9) conjecture, it generally does not, contributing a stochastic element to the realized coefficient on δ. This has important implications for the distribution of momentum returns.

This stochastic multiplier (the bracketed term in Eq. (10)) grows without bound as unanticipated crowding $k_M - Ek_M$ approaches Ek_I. Consider the extreme case of $k_M - Ek_M = Ek_I$, or $k_M = \lambda_E$. Recalling that $\delta^E = \lambda_E^{-1} f$ and $\delta^V = 0$ under myopic beliefs, Eq. (7) becomes

$$f(1 - k_M) - \delta k_I = f(1 - \lambda_E) \frac{k_M}{\lambda_E}, \tag{11}$$

where the left-hand side is the supply of the momentum portfolio to momentum investors and the right-hand side is their demand. Thus $k_M = \lambda_E$ implies that demand exceeds supply at all values of f. Momentum investors apply unrelenting upward pressure on the price of the momentum portfolio, never realizing that they are mostly—and eventually entirely[7]—feedback trading on each

[7]There is an equilibrium at a finite *negative* value for f when $k_M > \lambda_E$. This corresponds to a reversal of identifi-

others' same mistake. This potentially drives the market valuation of winners and losers far beyond their fundamental value, leading to a momentum crash. At less extreme values $k_M - Ek_M < Ek_I$ kurtosis and negative skewness in momentum returns can nevertheless be quite large.

Result 1 *If momentum investors are myopic to the possibility that unanticipated crowding has contributed to the formation-period valuation of the momentum portfolio, as expressed with beliefs Eq.* (9)*, then momentum returns can have arbitrarily large tail risk in the form of high variance, negative skewness, and excess kurtosis.*

2.2.3 Optimal linear beliefs

Momentum investors can avoid—or at least mitigate—the destabilizing effects of capital uncertainty by lessening the slope coefficient λ^{-1} in the $\delta^E = \lambda^{-1} f$ expression of linear beliefs. This gives a linear equilibrium similar to the left-side expression in Eq. (10), with λ replacing λ_E. A larger λ (greater attenuation of beliefs), lowers the possibility that k_M is large enough to cause the denominator to approach (or fall below) zero, therefore lowering variance, skewness, and kurtosis in momentum returns. However, from Eq. (6), a larger λ also means more moderate momentum investing and therefore lower profits in non-crash periods. We do not consider an analytical solution for the optimum scaling back of linear demands but note that the optimum presumably implies some sacrifice of profit and some residual exposure to higher moments. We determine the optimum with a grid search using the common market setting considered later in the analysis, and find both outcomes to be the case.

2.2.4 Rational beliefs

The conditional probability distribution that investors use to form δ^E and δ^V at a given value of f should produce demands that clear the market at the price f. Here we outline the solution to

cation 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 to enormous profit). We consider this case in Appendix D but note here that it too predicts extreme negative momentum returns.

those demands. The random variables in the system are δ, k_I, and k_M. We write their joint density as $g(\delta)\,h(k_M, k_I)$, noting that δ is independent of k_M and k_I but that the k's themselves are clearly dependent. Momentum investors seek to compute δ^E and δ^V given this distribution and observation of f:

$$\begin{aligned}
\delta^E &= \int_0^\infty \delta p_1(\delta|f)\,d\delta = \int_0^\infty \delta \frac{p_3(\delta, f)}{p_2(f)} d\delta, \\
\delta^V &= \int_0^\infty \left(\delta - \delta^E\right)^2 \frac{p_3(\delta, f)}{p_2(f)} d\delta,
\end{aligned} \tag{12}$$

where numerical subscripts serve to distinguish the functional form of each probability density function (pdf). To solve for these pdfs we replace the primitive random variable k_I with the observable random variable f using Eq. (7). This exchange of variables yields a joint pdf for the market clear price f equal to

$$p_4(\delta, k_M, f) = g(\delta)\,h\left(k_M, \frac{1}{\delta}\left(fD - \frac{\delta^E}{1 + \frac{\delta^V}{\sigma_d^2}} k_M\right)\right)\frac{D}{\delta}, \tag{13}$$

where D is as in Eq. (7) (see Appendix E). Integrating k_M, and then δ, out of Eq. (13) gives

$$\begin{aligned}
p_3(\delta, f) &= \frac{g(\delta)}{\delta} \int_0^1 h\left(k_M, \frac{1}{\delta}\left(fD - \frac{\delta^E}{1 + \frac{\delta^V}{\sigma_d^2}} k_M\right)\right) D dk_M, \\
p_2(f) &= \int_0^\infty \frac{g(\delta)}{\delta} \int_0^1 h\left(k_M, \frac{1}{\delta}\left(fD - \frac{\delta^E}{1 + \frac{\delta^V}{\sigma_d^2}} k_M\right)\right) D dk_M d\delta.
\end{aligned} \tag{14}$$

The expressions for beliefs, conditional on f, are

$$\delta^E = \frac{\int_0^\infty g(\delta) \int_0^1 h\left(k_M, \frac{1}{\delta}\left(fD - \frac{\delta^E}{1+\frac{\delta^V}{\sigma_d^2}} k_M\right)\right) D dk_M d\delta}{\int_0^\infty \delta^{-1} g(\delta) \int_0^1 h\left(k_M, \frac{1}{\delta}\left(fD - \frac{\delta^E}{1+\frac{\delta^V}{\sigma_d^2}} k_M\right)\right) D dk_M d\delta}, \tag{15a}$$

and

$$\delta^V = \frac{\int_0^\infty \frac{(\delta-\delta^E)^2}{\delta} g(\delta) \int_0^1 h\left(k_M, \frac{1}{\delta}\left(fD - \frac{\delta^E}{1+\frac{\delta^V}{\sigma_d^2}} k_M\right)\right) D dk_M d\delta}{\int_0^\infty \delta^{-1} g(\delta) \int_0^1 h\left(k_M, \frac{1}{\delta}\left(fD - \frac{\delta^E}{1+\frac{\delta^V}{\sigma_d^2}} k_M\right)\right) D dk_M d\delta}. \tag{15b}$$

In the simulations, we solve for δ^E and δ^V by conditioning on a given value of f,[8] repeating on a fine grid over the plausible range of values for f. This yields a discrete approximation to the continuous belief mappings $f \to \delta^E$ and $f \to \delta^V$. We then approximate the mapping at arbitrary values of f by interpolation.

3 Simulations

In this section we analyze the equilibrium under each specification of beliefs by solving Eq. (7) under each of 100,000 random draws for the triple δ, k_M and k_I. We then use those 100,000 equilibria to evaluate momentum return moments and consistency of beliefs. The crowding variables k_I and k_M (and the derived k_C) follow a symmetric Dirichlet distribution, Dir(α), which provides a natural dependence among capital allocations as fractions of a whole with equal expected value 1/3. We use concentration parameter $\alpha = 3$ to reflect a relatively diffuse prior belief for crowding, subject to the desideratum of vanishing probability that $k_{type} = 0$ or 1.[9] We use a lognormal distribution for δ with $\mu = -2.405$ and $\sigma = 0.125$.[10] Finally, ϵ is drawn from a zero-mean normal distribution with $\sigma_\epsilon = 0.125$.

Table 1 provides descriptive statistics of momentum returns and Figure 1 provides plots of beliefs and conditional mean returns. Each of the four columns of plots in Figure 1 correspond to

[8]In particular, we use Matlab's FSolve function to jointly locate the roots of the LHS minus RHS for Eqs. (15a) and (15b) at a given f.

[9]In an internet appendix we find qualitatively similar predictions under various permutations on α and other characteristics of the presumed setting, such as a uniform distribution for δ.

[10]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 et al. (2001).

a different assumption of beliefs. To construct each plot, simulation trials are ranked into 100 bins according to the conditioning variable (horizontal axis). Averages are then computed within each bin to approximate the conditional expectation of the variable indicated on the vertical axis.

[Insert Figure 1 and Table 1 near here]

Known crowding. In the simplest specification (Panel A) k_M and k_I are random but directly observed by momentum investors prior to trading. Plot A.1 of Figure 1 shows the mean for fundamental value δ as a function of beliefs δ^E. It is an identity, as expected since momentum investors effectively observe δ via f. From the third row of plots, A.3 depicts m decreasing with realized crowding ($k_M - Ek_M$) as competition bids away expected return. The relation is linear, since there is no unanticipated crowding (or feedback effects, or momentum crashes). This can also be seen in Table 1, Panel A where standard deviation and excess kurtosis are low, and skewness is slightly positive. The values provide a benchmark for considering the more realistic cases in Panels B through D.

Myopic beliefs Plot B.1 of Figure 1 indicates a strongly concave relation between beliefs δ^E and fundamental value δ, with investors substantially overstating value at the highest levels of beliefs. These errant beliefs derive from a failure to recognize that the source of the high f that they are linearly extrapolating is likely feedback from unanticipated crowding. The potential for a catastrophic outcome from this myopia is demonstrated in Plot B.3, where momentum crashes involve loss rates in excess of -300%.[11] Even more extreme crashes (on the order of -400%) are identified by conditioning on investors' beliefs δ^E as in Plot B.2. From Panel B of Table 1, momentum returns exhibit substantial volatility and extreme negative skewness and excess kurtosis. These crashes imply a certainty equivalent utility of assured total loss. In short, these results make it clear that feedback effects from unanticipated crowding can explain momentum crashes, provided investors are sufficiently myopic in their beliefs.

[11]Notice that the loss in the highest $k_M - Ek_M$ bin is smaller than in the second highest bin because momentum returns are here determined by the $-f$ equilibrium described in Appendix D.

Optimal linear beliefs. In this setting we presume that investors temper their myopia but continue to form beliefs as a linear extrapolation of past returns, rather than as a proper conditional expectation of value (as in the next and final case). We let λ range from the known-capital case of $2/3$ up to 1 ($\lambda = 1$ implies momentum demands are 0 at all f), and then select the value with the highest average realized utility, from Eq. (5). The presumed setting yields $\lambda^{-1} = 1.12$.

The first-order effect of optimally choosing λ is is to eliminate the run-away overstatement of beliefs seen with myopic beliefs. From Plot C.1 of Figure 1, fundamental value δ now at least monotonically increases with beliefs δ^E. However, the relation is nowhere near the identity required under rationality. Investors still overstate fundamental value when beliefs are high (fueling tail risk), and they now understate value when beliefs are low to moderate (sacrificing profit). Momentum returns are predictably negative with extreme beliefs, as seen in Plot C.2, and some negative skewness and excess kurtosis remains in Panel C of Table 1. Volatility of returns is higher than in the base (observed k_M) case, and there are rare cases of negative tail events severe enough to match momentum crashes observed in practice. Nevertheless, as seen in Plots C.2 and C.3, momentum returns are now largely free of crash risk and from Table 1 the higher moments of momentum returns are greatly attenuated relative to the previous case.

Note that equilibrium momentum returns are high (4.2% versus 3.0% in the base case), yet the expected profit to momentum investors is much lower at 0.65% versus 3.65%, respectively. Thus, protecting against crowding risk while adhering to a linear constraint comes at substantial cost. In trading off against this high cost, optimal linear beliefs leave some crash risk on the table. As we will see in the case of rational beliefs, this sacrifice of profit and the associated residual crash risk is entirely unnecessary. Investors can do a much better job of protecting against crowding risks, while at the same time garnering profits that almost match the base case of crowding certainty.

Rational beliefs. On virtually all counts market equilibrium with rational beliefs (and crowding uncertainty) behaves quite similar to the case of no crowding uncertainty. There are no aberrant higher moments, and all statistics relating to the first moment (mean returns, certainty equivalent returns, and profit) essentially match the base case.

First note the efficiency of beliefs as indicated in Plot D.1 of Figure 1: the mapping from beliefs δ^E to the actual mean δ is an identity as required under rationality. Likewise, Plot D.3 indicates an approximate linear relation between unanticipated momentum capital and momentum returns with no evidence of feedback from unanticipated crowding. From Plot D.2, expected profit is positive at all levels of belief, and remains so even in the most extreme realizations of unanticipated crowding (top bin of Plot D.3). Second, note that equilibrium momentum returns exhibit no negative skewness or excess kurtosis (Table 1, Panel D) and little incremental volatility. The worst-case return is -2.55%: crashes are not predicted under any of the 100,000 simulated market conditions (δ, k_I, and k_M).[12] Third, note that profitability of the momentum strategy with unanticipated crowding (and rational beliefs) is nearly equal to the setting with known crowding, and five times larger than in the case of optimal linear beliefs. Certainty equivalent return differences between known and unanticipated crowding are less than 0.1% for all values of risk aversion (γ) considered.

We conclude that unanticipated crowding does not provide a sufficient theoretical basis for momentum crashes; arbitrageurs must also be myopic to the possibility that they are feedback trading on their own price pressure. We summarize this as

Result 2 *Crowding induced crash risk in momentum returns is eliminated if momentum investors rationally incorporate tail risk into their conditional beliefs of fundamental value, $E(\delta|f)$. However, momentum returns remain negatively related to unanticipated momentum capital $k_M - Ek_M$ under all specifications of beliefs.*

4 Empirical section

We base our empirical analysis on quarterly holdings from the Thomson Reuters Institutional 13F database starting in Q1 of 1980 and ending in Q3 of 2015. Stock data are from CRSP using price and share adjustment factors, restricted to CRSP share code 10 and 11 and a listing on AMEX,

[12] Recall that each simulation solves for the expected momentum return under a given market state (fundamentals and crowding), integrating out the noise term ϵ. It is in this sense that we refer to a predicted return under a given state.

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, taken from Kenneth French's online data library.

4.1 Crowding proxies

To construct our crowding measures we first score the trading of institution i in quarter q based on alignment with a momentum strategy. Second, we label i a momentum institution in quarter q if recent scoring is consistently high. Third, we use the fraction of momentum institutions in quarter q as our measure of crowd size at that time, i.e., k_M.

Step one. Define two momentum scores; both relate to an inner product of change in portfolio weight and past returns. The first is taken from Grinblatt et al. (1995) and denoted

$$\text{GTWscore}_{i,q} = \sum_{j=1}^{J} \left(\omega_{i,j,q} - \omega_{i,j,q-1}\right) r_{j,q-1}, \tag{16}$$

where ω is a portfolio weight and $r_{j,q}$ is a quarterly stock return. Prior quarter prices are used in computing weight changes[13] to capture only active trading:

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

where w indicates shares held. The strategy implicit in GTWscore is a departure from the now-

[13]Note that 13F filings do not report short positions, hence weights apply only to long positions. However, used as an overlay to a broadly diversified investment strategy momentum investing implies overweighting winner stocks and underweighting loser stocks. Thus, changes in long portfolio weight should track sensitivity to both winner and loser stocks.

standard 12 - 1 momentum strategy, e.g., Jegadeesh and Titman (1993). Thus, we also define

$$\text{BEKscore}_{i,q} = \sum_{j=1}^{J} \left(\omega_{i,j,q} - \omega_{i,j,q-1}\right) \iota_{j,q} \tag{17}$$

using an indicator for past 12 - 1 return decile, where $\iota_{j,q} = 0$ unless stock j is a top-decile winner ($= 1$) or bottom-decile loser ($= -1$). This implies equal weighting of 20% of all stocks. By contrast, GTWscore implies continuous weighting of all stocks.

Step two. A single quarter of trading alignment does not make for a momentum *strategy*: a random portfolio change has a 50% chance of being so labeled. Moreover, to the extent that institutional trading impacts prices on the one hand, and is persistent on the other, both GTWscore and BEKscore are upward biased (with reverse causality giving the appearance of feedback trading even if trade motive is exogenous). This distinction is critical since there are no crowding-induced crashes *sans* feedback trading. Thus, we indicate (i, q) to be a momentum investor if $\mathbb{1}_{\sum_{l=0}^{3} \mathbb{1}_{\text{GTWscore}_{i,q-l}>0}=4} = 1$ (or BEK), i.e., a positive score in each of quarters $q - 3$ through q.

Step three. Aggregate by quarter to form a crowding measure. Note that each institution playing the momentum game optimizes a similar problem, conditioning on similar information (past prices). Thus, demand intensity is not unanticipated; what *is* unanticipated is how many institutions are playing the game. This implies that crowding uncertainty primarily relates to the count of momentum-investing institutions.[14] Hence we define

$$\text{GTW\_4qtr}_q = \frac{1}{N_q} \sum_{i=1}^{N_q} \mathbb{1}_{\sum_{l=0}^{3} \mathbb{1}_{\text{GTWscore}_{i,q-l}>0}=4}, \qquad \text{GTW\_1qtr}_q = \frac{1}{N_q} \sum_{i=1}^{N_q} \mathbb{1}_{\text{GTWscore}_{i,q}>0}, \tag{18}$$

$$\text{BEK\_4qtr}_q = \frac{1}{N_q} \sum_{i=1}^{N_q} \mathbb{1}_{\sum_{l=0}^{3} \mathbb{1}_{\text{BEKscore}_{i,q-l}>0}=4}, \qquad \text{BEK\_1qtr}_q = \frac{1}{N_q} \sum_{i=1}^{N_q} \mathbb{1}_{\text{BEKscore}_{i,q}>0}, \tag{19}$$

[14] More formally, from Eg. (6) using M_q as the count of momentum institutions, write aggregate momentum demands as $\sum_{i=1}^{M_q} \text{Demand}_{i,q} = \sum_{i=1}^{M_q} \left(\frac{E_{M,q}[m+\epsilon]}{\gamma Var_{M,q}[m+\epsilon]} \overline{K_{M,q}} \right) \frac{K_{i,q}}{\overline{K_{M,q}}} = \phi_q M_q$, where $\overline{K_{M,q}}$ is the average institution's capital and ϕ_q is the representative momentum investors' optimal demand intensity. If the latter is common knowledge, then M_q is the right proxy.

where N_q is the total count of institutions in quarter q. Our focus is the 4qtr measures on the left. We include the one-quarter versions listed on the right for completeness. Note that count is scaled by number of institutions to make the measures comparable across quarters.

We also consider two capital based proxies for crowding to incorporate trading intensity:

$$\text{BEKcap\_4qtr}_q = \frac{\sum_{i=1}^{N_q} Cap_{i,q} \mathbb{1}_{\sum_{l=0}^{3} \mathbb{1}_{BEK_{i,q-l}>0}=4}}{\sum_{i=1}^{N_q} K_{i,q}}, \quad \text{BEKcap\_1qtr}_q = \frac{\sum_{i=1}^{N_q} Cap_{i,q} \mathbb{1}_{BEK_{i,q}>0}}{\sum_{i=1}^{N_q} K_{i,q}}, \tag{20}$$

where $Cap_{i,q} = \sum_{j=1}^{J} P_{j,q} w_{i,j,q} \iota_{j,q}$ and $K_{i,q} = \sum_{j=1}^{J} P_{j,q} w_{i,j,q}$. In a multivariate setting, the BEKcap measures identify variation in momentum trading intensity across quarters. Finally, we consider both levels (referenced as $\text{Crowd}_{q\text{-}1}$) and changes (referenced as ΔCrowd_q) to capture anticipated and unanticipated components. We use a GARCH(1,1) specification of expected volatility in the crowding time series to capture crowding uncertainty (referenced as $\hat{\sigma}_{\text{Crowd}}$).

In short, the permutations are: GTW versus BEK formulations; count versus capital metrics; 4qtr versus 1qtr indicators for persistence in strategy; and levels versus changes. Our primary variable of interest is changes using 4qtr persistence on the BEK strategy with a count metric.

4.2 Descriptive statistics

Table 2 provides summary statistics for the 13F data (in Panel A); for the proxies for momentum investing (in Panel B); and for momentum returns (in Panel C). In Panel A we label an institution a *consistent* momentum investor if they follow a momentum strategy in two-thirds of the quarters for which we have the necessary data. We find that 23% (1,435/6,360) of institutions are consistent momentum investors. By contrast, Grinblatt et al. (1995) find that 59% of mutual funds are momentum investors, defined as having a positive average GTWscore.[15] Momentum institutions have higher turnover (26% compared to 20%); manage more assets (2.27 billion versus 1.28 billion);

[15]Grinblatt et al. (1995) also define momentum investors using returns and trading over the same quarter. Using this definition, they find that 77% of mutual funds are momentum investors.

and hold more stocks (204 stocks on average versus 125) than their counterparts.

[Insert Table 2 near here]

From Table 2, Panel B, approximately 50% of institutional investors are classified as momentum investors in a given quarter using either BEK_1qtr or GTW_1qtr. The more relevant 4qtr measures average 12.5% and 10.8%, respectively, compared to a $0.5^4 = 6.25\%$ random-trading null. Most crowding variables show strong persistence using the coefficient in an AR(1) regression. Given this (and below) evidence of persistence, we estimate the volatility of crowding using residuals from an AR(1) regression with a GARCH(1,1) specification (Bollerslev, 1986).

Table 2, Panel C summarizes regressions of momentum returns using the Fama-French 3 factor model (abbreviated FF3) and a dynamic version of the same model (dynamic FF3).[16] The latter is motivated by the evidence in Grundy and Martin (2001) that the momentum portfolio has strongly time-varying risk exposure. In the dynamic FF3 specification we include regressors with an interaction indicator variable for a positive prior-year factor return. In unreported results we find that momentum has substantial crash risk in our sample (high excess kurtosis with pronounced left-skewness), which includes the momentum crash of March-May 2009.

Table 3, Panel A considers the persistence of the four momentum classifications in more detail. The probability of maintaining the current classification in the following quarter is 71% for both BEK_4qtr and GTW_4qtr. The four quarter ahead probabilities average 33% for these measures, or about three times their respective unconditional probabilities (listed under column 'All q'). Table 3, Panel B relates each of the four measures to contemporaneous and predictive values for BEK_1qtr; the idea being that BEK_1qtr tracks realized momentum trading. The probability of a positive BEK_1qtr four quarters ahead is about 68% for the 4qtr measures, suggesting they provide meaningful predictors of momentum trading.

[16] We often make reference to FF3 or dynamic FF3 models or residuals. In fact, in return regressions the dependent variable is the momentum factor return with FF3 or dynamic FF3 factors are included as controls. For the crash and volatility analyses residuals from a FF3 or dynamic FF3 models are used.

[Insert Table 3 near here]

Table 3 also indicates the rapidly changing composition of the momentum portfolio. Winners have a 56% chance of remaining winners the following quarter, but at four quarters the likelihood 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.

4.3 Crowding and conditional expected returns on the momentum factor

Table 4 shows the results of predictive regressions of momentum returns on the various crowding measures. All momentum trading measures are lagged (in this and subsequent tables) to ensure that there is no overlap between the measurement of the independent variable and the momentum return. For example, we use the change in BEK_1qtr_q to predict momentum returns in quarter $q + 1$. As a control we include lagged realized volatility of momentum computed from squared daily momentum returns in the previous quarter. Barroso and Santa-Clara (2015) show that this strongly predicts (negatively) momentum returns.[17] Because computing the regressors requires up to six quarters of data, the regression sample begins in Q3 1981 and ends in Q4 2015.

[Insert Table 4 near here]

We find that both anticipated crowding (proxied with $\text{Crowd}_{q\text{-}1}$) and unanticipated crowding (proxied with ΔCrowd_q) negatively relate to the mean of momentum returns. The relation is generally significant with the 4qtr proxies (Panel A) but not so with the noisier 1qtr proxies (Panel B).[18] These results are as predicted. We generally do not see any reliable relation with BEKcap proxies, where the coefficient estimate is often (insignificantly) positive.

The difference between the two proxies (count-based vs. Cap) is the intensity of the momentum bet. Hence the pattern of results in Table 4 suggests that the *intensity* of momentum investing

[17] In unreported results we also controlled for the bear market states proposed by Cooper et al. (2004). Using this control in our sample period did not change our results.

[18] Unless otherwise noted the significance levels discussed refer to two-tailed tests even when the model provides a clear prediction for the sign of the coefficient.

positively relates to future returns, whereas the *number of competitors* (i.e., crowding) negatively relates to future returns. This is consistent with rational beliefs, where demand intensity recoils from crowding-induced feedback with its associated lower returns. It is inconsistent with myopic beliefs, where crowding-induced feedback generates a component of demand intensity that negatively relates to returns. This is the first of several pieces of evidence we provide that supports the rational beliefs hypothesis over the myopic beliefs alternative.

Finally, from Table 4 the anticipated volatility of crowding, $\hat{\sigma}_{\text{Crowd}}$, is positively related to the expected returns at the 1% level in some specifications, but the relation is generally insignificantly positive. A positive relation implies that uncertainty in the number of competing momentum investors inhibits participation in the strategy, lowering momentum demand and therefore raising evaluation-period returns. But that inference is cloudy at best. In our later consideration of the second moment (volatility, in Table 7) we find evidence (again cloudy) consistent with this story.

4.4 Crowding and negative tail events in momentum returns

The focus of the study is the relevance of unanticipated crowding for the pronounced left tail in the distribution of momentum returns. We assess this link in Table 5 with a probit analysis of tail probabilities using 4qtr proxies and both raw and dynamic FF3 returns. First, note that the mean dependencies documented in Table 4 suggests that crowding shifts the distribution of momentum returns leftward. Even if crowding had no effect on higher moments, the probit analysis might indicate tail risk. We therefore use a bivariate probit analysis that considers the shift in probabilities for both the left and right tail, and then report on the difference to identify crash-like tendencies. The null hypothesis is that fattening of the left tail is due only to a downward shift in the mean, rather than an elevation in negative skew.

[Insert Table 5 near here]

The p-values in square brackets from Table 5 provides the Wald test of a difference in left versus right tail effects. Using BEK proxies we find a statistically reliable positive relation between

$\Delta\text{Crowd}_\text{q}$ and 10% left-tail probabilities. This suggests that crowding indeed contributes to tail risk, but the evidence is weaker with GTW and using 5% tails, perhaps because of small-sample difficulties. The Wald tests indicate that the impact of crowding on tail risk is largely due to the shift in mean returns documented in Table 4 rather than increased negative skew. Few p-values are significant at conventional levels, with the results for $\hat{\sigma}_{\text{Crowd}}$ coming the closest. Several others are significant at a 20% level but generally the evidence of crowding induced tail risk is very weak. One might also note that in all cases the strongest signal comes from the count-based BEK proxy. The Cap versions are never close to relevant. Thus, to the extent that there is any evidence of a relation between crowding and momentum tail risk it relates to unanticipated number of competitors, not unanticipated trade intensity.

[Insert Figure 2 near here]

Figure 2 plots the time series of our crowding measures. The measures shown suggest 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 and argue that the actions of a small cluster of momentum investors can exert considerable influence on prices. 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.

4.5 Tail risk

While the notion of tail risk surely involves pronounced left skewness, to be meaningful it must also be accompanied by high volatility and excess kurtosis.[19] With this in mind we examine all three higher moments of momentum returns in Table 6 by sorting calendar quarters based on our Crowd proxies. We add sorts based on lagged realized volatility of momentum returns for comparison.

[19] 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.

The evidence in Table 6 does not support crowding as a source of tail risk in momentum returns. Rather, momentum investors seem somewhat successful in avoiding it. For example, consider the unanticipated-crowding proxy ΔCrowd_q: it is statistically unrelated to subsequent values of each of the three contributors to tail risk; volatility, skewness, and kurtosis. If anything, the relation is negative. In particular, using the count-based proxy BEK_4qtr (middle column of results), skewness is significantly more negative in the *bottom* (low) tercile of ΔCrowd_q (t-statistic 4.1). Likewise, excess kurtosis is significantly higher in that low tercile. This pattern is more consistent with the unanticipated component of crowding forecasting—and avoiding—tail risk than causing it. Contrast these results with the case of conditioning on lagged volatility in momentum returns. Here, we see statistically reliable prediction of tail risk in all moments, in the right direction. 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.

[Insert Table 6 near here]

Table 6 offers important insight. There is tail risk in our sample and it does relate to ex ante market conditions. However, crowding (particularly unanticipated crowding) does not seem to be one of them. Indeed, the contrast in Table 6 between the predictability of return volatility and the predictability of direct measures of crowding (particularly the contrast in direction of predictability) strongly suggests that momentum investors form beliefs that incorporate tail risk. Quite likely, those beliefs derive from return volatility. As in the theory, such rational beliefs prevent unanticipated crowding from itself causing tail risk. The evidence here goes further: high levels of momentum investing (top tercile) seem to identify a predictably stable environment, much less self-inflicted tail risk.

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, we embarked upon the project with just this perspective. However, in examing this crowding hypothesis from both a theoretical perspective that considers rational beliefs, and from an empirical perspective that employs direct proxies, we find just the opposite. Optimizing momentum investors identify the potential for crowding to destabilize momentum returns, and they adjust their demands accordingly.

4.6 Crowding and the volatility of momentum returns

If volatility clustering in momentum returns derives from crowding, rather than factors exogenous to arbitrageur actions, we should find that momentum volatility is predicted by measures of crowding. To examine this Table 7 presents predictive regressions for realized volatility of momentum returns computed from raw and risk adjusted (FF3 and dynamic FF3) daily returns over the quarter.

[Insert Table 7 near here]

Consistent with the preceding results, Table 7 finds that lagged realized volatility has strong predictive power for subsequent volatility, with t-statistics between 6.4 and 9.1 across all regressions. However, ΔCrowd_q and $\text{Crowd}_{q\text{-}1}$ provide fairly robust inferences that crowding predicts negatively volatility in momentum returns. The coefficient estimate on both specifications (applied to BEK measures) is statistically significant at the 5% level in 4 of 12 cases across Panels A and B, and is negative in all cases. This result is consistent with expectations of risk in the momentum strategy affecting 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 7 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.

4.7 Determinants of crowding

Table 8 presents regressions of Crowd_q on one-year momentum returns and one-year momentum-return volatility computed from daily observations, lagged at the indicated horizon. To ensure predetermined values for the 4qtr measures, we add one-year returns and volatility lagged five quarters to the analysis. From Panel A, the BEK_4qtr specification of Crowd_q is significantly negatively related to one-year volatility lagged one quarter. Overlap in observation periods obfuscates the sequencing of events here (we have already seen from Table 7 that BEK measures predict negatively future risk in the momentum strategy). However the BEK_1qtr specification of Crowd_q is strongly negatively related to past volatility in momentum returns. Combining inferences from the anticipation (Table 7) and reaction (Table 8) analyses, we conclude that volatility in momentum returns is a determinant of (reduced) crowding, that its impact is largely anticipatory, and that the response is fast (no reliable dependence of BEK_4qtr on 5-quarter lagged volatility).

[Insert Table 8 near here]

Chabot et al. (2014) show in a comprehensive sample period of 140 years that the crash risk of momentum increases after periods of good recent returns in the strategy. 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. A natural question is whether this predictive relation stems from crowding.

Table 8 also shows that one-year returns indeed predict positively crowding in momentum. The coefficients on lagged returns from Table 8 are positive in all regressions and statistically significant at the 1% level in eight cases out of twelve. In the case of overlap between the estimation period for 4qtr measures and returns (i.e., top row), the relation is less positive due to the negative effect of crowding on returns seen in Table 4. However, in the case of returns lagged five quarters, we again see a reliable positive relation using the base BEK_4qtr specification. This is also seen in the 1qtr specifications with one-quarter lagged returns. Unfortunately we are unable to relate this to the evidence in Chabot et al. (2014) because higher past one-year returns do not predict higher crash risk in our sample.[20]

Cooper et al. (2004) show that momentum returns are stronger in bull markets. That evidence supports the interpretation of momentum as partially caused by over-confident and self-attributing investors becoming more so during bull markets (Daniel et al., 1998). In unreported results, we do not find any predictive power of market states for our measures of momentum crowding, once we control for the lagged returns and volatility of the strategy.

Our results are consistent with the momentum strategy becoming more crowded when its recent performance is good—both in terms of high returns and low volatility. As previously discussed, the volatility result suggests that forward-looking rational momentum investors time strategy risk. Chasing momentum returns may also be consistent with our theory. We have not modeled variance in δ (the magnitude of dispersion in private information about fundamental value), but predictable (i.e., persistent) changes in this parameter should yield both higher past momentum returns and larger momentum demands—i.e., return chasing. In our sample strategy returns do not show time-series autocorrelation, but that could be an equilibrium result of the stronger demands.

4.8 Capital versus crowd

We have argued that BEK measures track the number of momentum investors and BEKcap measures track the intensity of their momentum-demands. In our final analysis we consider the two dimensions of crowding in a multivariate setting. As the intent is to summarize, we consider all three moments of momentum returns in a single table using the dynamic FF3 model and the BEK rather than GTW specification for the number of momentum-trading institutions. Results in Table 9 confirm all previous findings.

[Insert Table 9 near here]

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

4.9 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 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 volatility as a signal to avoid tail risk than volatility being caused by 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 6, except that the focus is momentum gap and momentum gap orthogonalized to our crowding measures. 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 value of a sorting variable in the last available quarter. In the first column the ranking variable is the momentum gap. Consistent with Huang (2015), we find that a high momentum gap forecasts higher volatility, negative skewness, and excess kurtosis; all with statistical significance at the 1% level.

Next we orthogonalize momentum gap with respect to measures of institutional crowding (denoted gap$^{\perp}$). In addition to levels of BEK_4qtr, GTW_4qtr, and BEKcap_4qtr, we consider ΔMom

Inst from Huang (2015), defined as the percentage difference in aggregate institutional ownership between past winners and losers. We also include the Win Inst measure from Lou and Polk (2013), defined as the aggregate institutional ownership of the winner decile. Columns 2 to 6 show that $gap^{\perp}$ retains substantially all of the predictive power of the momentum gap itself. Overall, the evidence in Columns 2 to 6 shows that while the momentum gap is a strong predictor of the performance of the strategy, its predictability does not appear to derive from crowding.

Finally, in the last column of Table 10 we orthogonalize the momentum gap to the volatility control used in many of our empirical exercises. Since the momentum gap is defined as the interquartile range of the return distribution for stocks in a given formation period, it is a measure of (cross-sectional) dispersion in returns that should be closely related to volatility. Indeed, the correlation between momentum gap and volatility is 0.73 in our sample period. Nevertheless, $gap^{\perp}$ remains a reliable predictor of volatility. Volatility averages 35.6% following high $gap^{\perp}$ versus 21.0% following low $gap^{\perp}$, with a t-statistic for the difference of 3.2. This suggests that momentum gap captures different information from that contained in volatility. However, volatility largely strips momentum gap of significance in its forecasting of higher moments of momentum returns.

Momentum gap relates to tail risk. In principal, that relation could be a reflection of crowded trades, or some other destabilizing shock that drives prices far apart in the formation period only to reverse badly in the evaluation period. Our evidence suggests the latter; a conclusion mirrored with our results for volatility. Both volatility and momentum gap predict tail risk *despite* the behavior of institutions, not because of their behavior.

5 Conclusion

We provide a model of momentum investors who infer from past returns regarding the incomplete assimilation of informed investors' private signals of value. 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 non-linear) beliefs, the potential destabilizing effects of crowding are internalized into demands. This mitigates feedbacking behavior 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, only crowding with myopic arbitrageurs can provide that prediction.

Our empirical contribution is twofold. First, we directly examine proxies for momentum trading by institutional investors, in contrast to much of the literature which focuses on indirect inferences of crowding from return covariances or volatility. Second, we directly examine the implications of optimal versus myopic beliefs for patterns in momentum trading.

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.

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 et al., 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). 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 et al., 2016), high sentiment (Antoniou et al., 2013), and low market volatility (Wang and Xu, 2015). Hillert et al. (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. As a whole, this evidence suggests that crowding plausibly explains the higher-moment characteristics of momentum.

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 et al., 2016; DeMiguel et al., 2017). 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., Daniel et al., 1998; Barberis et al., 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 et al. (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 et al. (1990a,b).

More recently, Kondor and Zawadowski (2015) 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. While our analysis uses a different approach and insights in proxying for momentum capital, our result on unanticipated momentum capital and momentum returns is generally consisted with their finding, but we cannot 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 4.9 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 crowded trades of institutions.

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 et al. (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 et al. (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.

B Development of momentum portfolio

Stocks are indexed by j. Each pays a discrete dividend X_j which evolves according to

$$\log\left(X_j/X_{j,0}\right) = \chi + \iota_j \frac{d}{2},$$

where $X_{j,0}$ is the beginning of period dividend, χ is a random zero-mean innovation common to all stocks with variance σ_χ^2 that generates the market return; the indicator ι_j selects the momentum portfolio, taking on the value 1 or -1 for 10% of all stocks (in each leg); and d generates the differential return on the two groups of stocks, with variance σ_ϵ^2 and mean δ where δ is the private information in the market. All investors know σ_ϵ^2 and σ_χ^2. Let $\boldsymbol{X}$ denote the vector of all stocks' dividends, which is unknown prior to the end of the period, at which point it becomes known to all. We refer to stocks with $\iota_j = 1\,(-1)$ as winner (loser) stocks.

At the beginning of the period information is symmetric, hence each investor holds the market portfolio. This results in a vector of public-information valuations $\boldsymbol{P}_0 = \boldsymbol{X}_0/r$ where r denotes an unmodelled required return on the market portfolio. At some intermediate time within the period (portfolio formation date) a subset of investors managing beginning-of-period capital K_I observes δ and $\boldsymbol{\iota}$ and trades. This trading identifies the momentum portfolio by generating a formation period return on winner and loser stocks via informed investors' demands. Informed investors expect an end of period price increase of $e^{\frac{1}{2}\left(\delta+\sigma_\chi^2\right)+\frac{1}{8}\sigma_d^2}$ for each winner stock and a price decrease of $e^{\frac{1}{2}\left(-\delta+\sigma_\chi^2\right)+\frac{1}{8}\sigma_d^2}$ for each loser stock. Neutral stocks experience no trading, because all market participants maintain the same expected end of period dividend increase of $e^{\frac{\sigma_\chi^2}{2}}$ for such stocks.

Uninformed (i.e., momentum and counterparty) investor begin with homogeneous expectations of dividends for all stocks, equal to their beginning of period values. Thus, the returns on each winner; loser; and neutral stock are homogeneous within type. Moreover, by presuming the same average information signal on winner stocks ($\delta/2$) as loser stocks ($-\delta/2$), the long-short portfolio of winners minus losers (i.e., the momentum portfolio), can be treated as a single asset. The information signal on this long-short portfolio, δ, parameterizes the private information in the market.

C Derivation of Eq. (6)

First notice that solving Eq. (5) is equivalent to solving each of the following (presuming $\gamma > 1$)

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

where ς is the weight on the portfolio of risky asset 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 $g\left(r_p - r_f\right) = \log\left(1 + \varsigma\left(e^{r_p-r_f} - 1\right)\right)$ using a second-order Taylor expansion around 0:[21]

$$\begin{aligned} g\left(r_p - r_f\right) &\cong \log(1) + \frac{\varsigma e^0}{1+\varsigma\left(e^0-1\right)}\left(r_p - f\right) + \frac{1}{2}\frac{\varsigma\left[e^0\left(1+\varsigma\left(e^0-1\right)\right) - \varsigma e^{2\cdot 0}\right]}{\left(1+\varsigma\left(e^0-1\right)\right)^2}\left(r_p - f\right)^2, \\ &\cong \varsigma\left(r_p - f\right) + \frac{1}{2}\left(\varsigma - \varsigma^2\right)\sigma^2, \end{aligned} \tag{C.1}$$

where $\left(r_p - f\right)^2$ is replaced with its expected value σ^2. Using Eq. (C.1), we can then rewrite the maximization problem to

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

[21] 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.

which has solution

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

To proceed, we will 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 in the context of a portfolio comprised 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 momentum overlay w_m. The optimal risky portfolio is then that choice of w_m that solves the constrained optimization

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

using weights $\boldsymbol{w}' = \begin{bmatrix} 1 & w_m \end{bmatrix}$, where

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

and $r^* - r_f$ is a target return premium that traces out the efficient frontier (recall r is the required return on the market portfolio). This has solution

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

Using Eq. (C.2), the parameters of the optimal risky portfolio are

$$\mu_p - r_f = \begin{bmatrix} r - r_f & E_{type}[m + \epsilon] \end{bmatrix} \begin{bmatrix} 1 \\ w_m \end{bmatrix} = r - r_f + w_m E_{type}[m + \epsilon],$$

and

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

Taking the ratio gives

$$\varsigma = \frac{r-r_f}{\gamma\sigma_\chi^2}. \tag{C.3}$$

Combining Eqs. (C.2) and (C.3),

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

D Negative market-clearing price for momentum portfolio

In the case of $k_M > \lambda_E$, the demand of momentum investors (RHS Eq. (11)) increases with a positive f faster than the LHS supply can keep up with, implying an increasingly large buying imbalance as f rises (depicted in the third panel of Figure 1). This again suggests that momentum investors buy up to their capacity, leading to a subsequent momentum crash.

However, when $k_M > \lambda_E$ 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 be lucky enough to induce momentum investors into selling (buying) so much winner (loser) stock that their bizarre trade is preferred.

E Joint pdf for the market clearing price

Let

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

which characterizes market clearing as in Eq. (7). We map the primitive random variables into

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

Next, we need

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

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

$$\begin{aligned} p_4(\delta, k_M, f) &= g(\delta)\, h\left(k_M, F^{-1}\right) | \, \mathrm{J} \, |, \\ &= g(\delta)\, h\left(k_M, \frac{1}{\delta}\left(fD - \frac{\delta^E}{1 + \frac{\delta^V}{\sigma_d^2}} k_M\right)\right) \frac{D}{\delta}. \end{aligned}$$

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.

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.

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.

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

M. Brunnermeier and L. 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.

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. *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. Martin-Utrera, F. J. Nogales, and R. Uppal. A portfolio perspective on the multitude of firm characteristics. Working Paper SSRN, 2017.

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.

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

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

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, 2001.

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. *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.

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

P. Kondor and A. Zawadowski. Learning in crowded markets. Working Paper Central European University, 2015.

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.

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

J. Peress. Wealth, information acquisition, and portfolio choice. *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.

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.

R. Sias, H. Turtle, and B. Zykaj. Hedge Fund Return Dependence: Model Misspecification or Liquidity Spirals? *Journal of Financial and Quantitative Analysis*, 52(5):2157–2181, 2017.

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.

K. Q. Wang and J. Xu. Market volatility and momentum. *Journal of Empirical Finance*, 30:79–91, 2015.

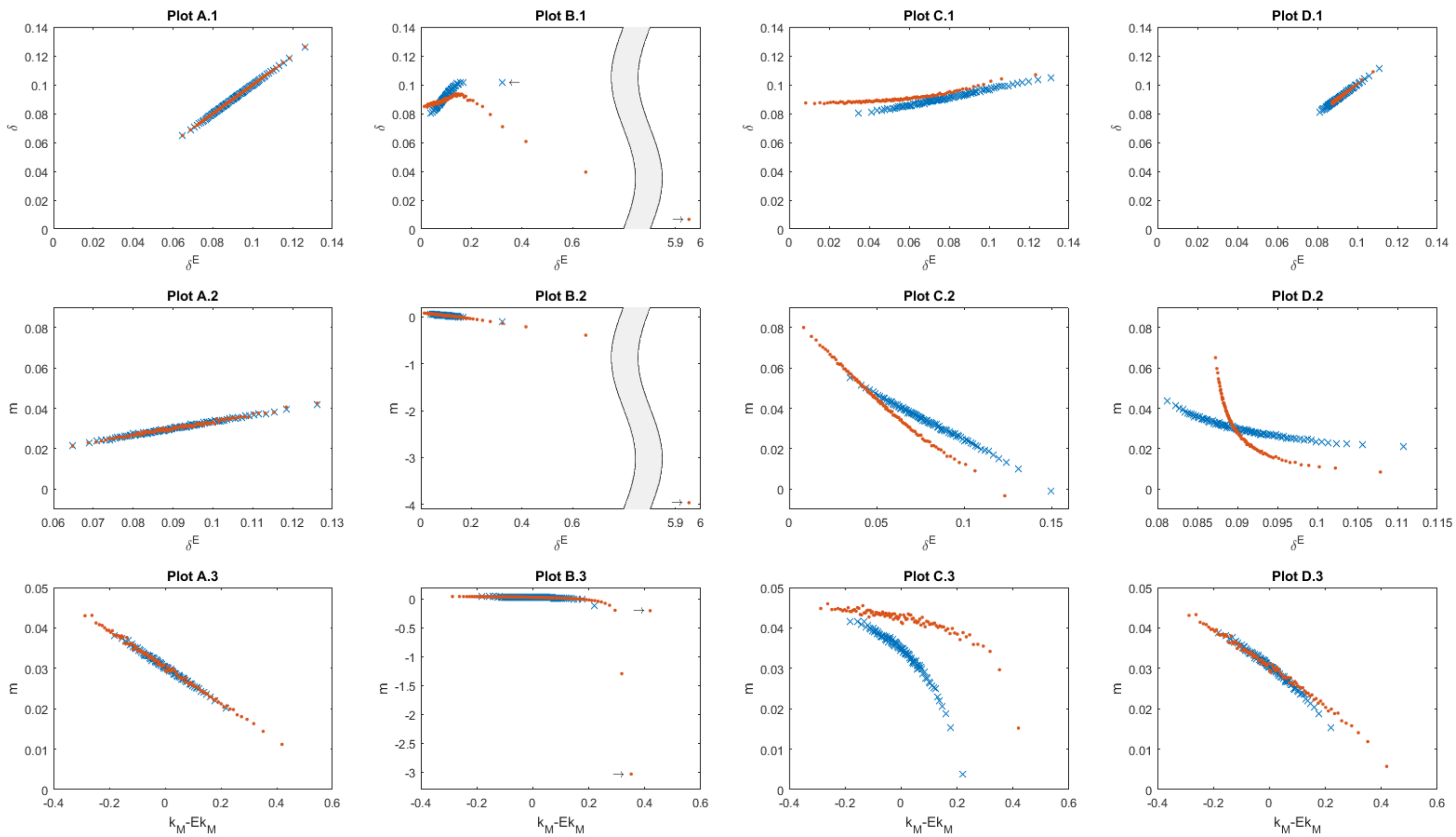

Figure 1: Beliefs and expected momentum returns in simulations

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 $\alpha_i = 12$ (blue Xs) and 3 (red dots), and $Ek_I = Ek_M = 1/3$. δ follows a lognormal 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.1-3, myopic beliefs in Plots B.1-3, optimal linear beliefs in Plots C.1-3, and rational beliefs in Plots D.1-3. The expected momentum return is then $m = \delta - f$. The triplets $\{\delta, m, k_M\}$ 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.

(a) one quarter

(b) four quarters

Figure 2: Measures of crowding

Panel (a) and (b) report the _1qtr and _4qtr crowding measures, respectively, constructed with 13F holdings data in the period from 03/1980 to 09/2015. The shaded areas denote NBER recessions.

Table 1: Momentum returns in simulations

The table reports unconditional return statistics for the simulations described in the caption of Figure 1. Mean, stdev, skew, kurt, min, and max refer to average, standard deviation, skewness, kurtosis, minimum, and maximum, respectively. Panel A reports the results for the case of known crowding, Panel B is myopic beliefs, Panel C is optimal linear beliefs, and Panel D is rational beliefs. We thereby consider the low var(k) case in which the Dirichlet distribution has the concentration parameters $\alpha_i = 12$, and the high var(k) case in which $\alpha_i = 3$. The slopes of the optimal linear beliefs are chosen to maximize the utility of a CRRA investor with $\gamma = 2$, and they are reported in the row labelled λ^{-1}. Profits are the expected portfolio returns of $\gamma = 2$ investor, and certainty equivalents 'cer(γ)' are calculated for $\gamma = 2, 4, 10$, with portfolio weights calculated as in Eq. (6). Momentum returns are given by $d = m + \epsilon$ where ϵ is randomly drawn from a zero-mean normal distribution with standard deviation 0.125. Cer(γ) is an arithmetic return; all other statistics are based on log returns.

	Panel A		Panel B		Panel C		Panel D	
var(k)	low	high	low	high	low	high	low	high
λ^{-1}			1.50	1.50	1.38	1.12		
Expected momentum returns m								
mean	3.0%	3.0%	2.7%	-2.4%	3.4%	4.2%	3.0%	3.0%
stdev	0.8%	1.4%	19.5%	174.2%	1.4%	2.0%	1.1%	1.6%
skew	0.5	0.6	-300.5	-151.3	-0.3	-0.3	0.3	0.4
kurt	3.3	3.1	92991.2	29218.7	4.7	10.8	3.2	3.0
min	0.74%	0.05%	-6046.58%	-38957.17%	-15.64%	-53.10%	-1.38%	-2.55%
max	7.57%	10.26%	10.38%	13.16%	10.78%	13.28%	9.36%	11.53%
Realized momentum returns d								
profit	3.17%	3.65%	-58.05%	-4863.08%	2.18%	0.65%	2.98%	3.44%
cer(2)	2.37%	2.62%	-100.00%	-100.00%	1.94%	0.74%	2.28%	2.53%
cer(4)	1.18%	1.30%	-100.00%	-100.00%	0.96%	0.37%	1.13%	1.25%
cer(10)	0.47%	0.52%	-100.00%	-100.00%	0.38%	0.15%	0.45%	0.50%

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 refer to institutions classified as a momentum trader by one of our measures for at least 2/3 of the available quarters. Crowd refers to GTW_4qtr, BEK_4qtr, and BEKcap_4qtr; likewise the _1qtr versions (as defined in Section 4.1). $\hat{\sigma}_{\text{Crowd}}$ is the estimate of volatility from a GARCH(1,1) model. Panel C contains factor exposures of quarterly momentum factor returns on the Fama-French three factor model and on a dynamic extension including the three factors interacted with dummies for positive past annual factor returns. Alphas are monthly and t-statistics use White standard errors.

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	33.6	21.0	31.9	38.3	29.0	31.8
#Qtrs missing	3.9	1.0	9.9	3.8	0.0	9.7	3.9	1.0	9.9
#Stocks held	143.2	62.9	275.5	204.9	88.6	365.5	125.3	55.4	241.8
Assets mgd.	15.2	2.0	102.4	22.7	2.1	157.3	12.8	1.9	78.7
Turnover	0.21	0.16	0.17	0.26	0.20	0.19	0.20	0.15	0.16
#Institutions	6,360			1,435			5,033		

Panel B. Crowding variables

	GTW_				BEK_				BEKcap_			
	mean	med.	stdev.	ar(1)	mean	med.	stdev.	ar(1)	mean	med.	stdev.	ar(1)
_1qtr measure:												
ΔCrowd	0.001	0.002	0.071	-0.375	0.000	0.004	0.079	-0.582	0.000	0.000	0.034	-0.095
Crowd	0.463	0.461	0.059	0.253	0.495	0.501	0.059	0.090	0.050	0.042	0.051	0.779
$\hat{\sigma}_{\text{Crowd}}$	0.056	0.054	0.008	0.745	0.054	0.051	0.013	0.618	0.028	0.022	0.018	0.695
_4qtr measure:												
ΔCrowd	0.000	-0.001	0.017	0.109	0.000	0.000	0.019	-0.055	0.000	0.000	0.012	-0.020
Crowd	0.108	0.108	0.029	0.837	0.125	0.124	0.033	0.844	0.020	0.016	0.020	0.818
$\hat{\sigma}_{\text{Crowd}}$	0.015	0.016	0.002	0.979	0.018	0.017	0.002	0.898	0.011	0.008	0.007	0.715

Panel C. Returns

	Mom on FF3					Mom on dynamic FF3							
	alpha	mkt	SMB	HML	R2	alpha	mkt	SMB	HML	Dmkt	DSMB	DHML	R2
coefficient	0.016	-0.35	-0.48	-0.59	12%	0.014	-0.85	-0.68	-0.95	0.81	0.48	1.02	25%
t-statistic	(4.4)	(-2.0)	(-2.1)	(-2.1)		(4.3)	(-2.9)	(-2.7)	(-2.7)	(2.8)	(1.2)	(2.1)	

Table 3: Transition frequencies

Panel A tabulates the probability of remaining in the classification indicated in the row heading at the time indicated in the column header (q indexes quarters), conditional on the later period not containing a missing observation. The likelihood is here the conditional relative to the unconditional probability of classification. Panel B contains the probability of being a net buyer of momentum stocks (i.e., being classified as a momentum investor according to BEK_1qtr) in the quarter indicated in the column heading conditional on the classification indicated in the row heading in q. The likelihood is now taken relative to the unconditional probability of being a net buyer of momentum stocks. Panel C tabulates the transition probabilities for the stocks' membership in the winner, loser, or middle 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
GTW_1qtr	0.54	0.54	0.45	1.20	1.19
GTW_4qtr	0.71	0.34	0.10	7.05	3.32
BEK/BEKcap_1qtr	0.57	0.56	0.49	1.17	1.16
BEK/BEKcap_4qtr	0.71	0.31	0.12	5.99	2.62

Panel B. Institutions' trading

	probabilities			likelihood		
	q	q+1	q+4	q	q+1	q+4
GTW_1qtr	0.68	0.57	0.56	1.39	1.16	1.15
GTW_4qtr	0.78	0.70	0.69	1.60	1.44	1.42
BEK/BEKcap_1qtr	1.00	0.57	0.56	2.06	1.17	1.16
BEK/BEKcap_4qtr	1.00	0.71	0.68	2.06	1.46	1.41

Panel B. Stock returns

	q+1			q+4			All
	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 4: Momentum factor returns on crowding measures

Each column represents a predictive regression of quarterly momentum factor returns (1981–2015) on crowding. Each panel presents three specification: (1) without controlling for risk-factors; (2) controlling for the Fama and French three factor model; and (3) controlling for the dynamic factor model with the three factors interacted with dummies for positive past annual factor returns. Each set considers the three indicated crowding measures. The regressor 'Crowd' refers to the level of the crowding measure at the end of quarter q-1; ΔCrowd_q refers to the change over quarter q; and $\hat{\sigma}_{\text{Crowd}}$ is the estimate of volatility from a GARCH(1,1) model. 'Realized vol. of Mom rets.' is a control variable equal to the lagged realized volatility of daily momentum returns over the previous quarter (intercepts not tabulated). The t-statistics are computed with White standard errors.

Panel A. Crowding measures constructed using four-quarter trading histories									
Model:	cumulative returns			FF3			dynamic FF3		
Measure:	GTW	BEK	BEKcap	GTW	BEK	BEKcap	GTW	BEK	BEKcap
ΔCrowd_q	-0.29	-0.41	-0.27	-0.37	-0.47	-0.31	-0.33	-0.44	-0.22
	(-1.4)	(-2.1)	(-0.9)	(-1.8)	(-2.7)	(-1.1)	(-1.8)	(-2.4)	(-0.6)
$\text{Crowd}_{q\text{-}1}$	-0.50	-0.15	0.28	-0.53	-0.13	0.26	-0.58	-0.12	0.33
	(-3.4)	(-1.1)	(1.0)	(-3.7)	(-1.3)	(1.2)	(-4.3)	(-1.3)	(1.6)
$\hat{\sigma}_{\text{Crowd}}$	4.61	1.83	0.14	5.52	2.51	0.25	6.61	1.60	-0.13
	(2.3)	(0.8)	(0.2)	(2.8)	(1.2)	(0.4)	(3.7)	(0.8)	(-0.2)
Realized vol.	-0.29	-0.34	-0.32	-0.31	-0.36	-0.33	-0.25	-0.30	-0.27
of Mom rets.	(-1.6)	(-1.8)	(-1.7)	(-2.2)	(-2.6)	(-2.4)	(-2.2)	(-2.5)	(-2.3)
Adj-rsquare	12.1%	10.1%	9.3%	25.8%	24.3%	22.6%	37.7%	33.3%	32.3%
Panel B. Crowding measures constructed using one-quarter trading histories									
Model:	cumulative returns			FF3			dynamic FF3		
Measure:	GTW	BEK	BEK	GTW	BEK	BEKcap	GTW	BEK	BEKcap
ΔCrowd_q	-0.03	-0.03	-0.04	-0.02	-0.05	-0.10	-0.03	-0.11	-0.11
	(-0.4)	(-0.5)	(-0.5)	(-0.4)	(-0.8)	(-1.3)	(-0.4)	(-1.3)	(-1.1)
$\text{Crowd}_{q\text{-}1}$	-0.07	-0.02	0.14	-0.06	-0.02	0.11	-0.02	0.00	0.16
	(-1.0)	(-0.2)	(1.3)	(-0.9)	(-0.2)	(1.3)	(-0.2)	(0.0)	(2.1)
$\hat{\sigma}_{\text{Crowd}}$	0.78	0.24	0.05	0.76	0.33	0.17	0.83	0.30	0.03
	(1.8)	(0.7)	(0.2)	(1.9)	(1.2)	(0.8)	(1.6)	(0.9)	(0.1)
Realized vol.	-0.31	-0.30	-0.32	-0.32	-0.32	-0.35	-0.27	-0.25	-0.29
of Mom rets.	(-1.7)	(-1.6)	(-1.6)	(-2.3)	(-2.2)	(-2.4)	(-2.4)	(-1.9)	(-2.4)
Adj-rsquare	9.1%	6.7%	9.6%	21.6%	20.2%	23.5%	31.8%	33.2%	35.1%

Table 5: 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 10% (Panel A) or 5% (Panel B) of the full-sample (1981–2015) distribution, where returns are either raw or dynamic Fama and French residuals. Each set considers three crowding proxies, in all cases using 4qtr. The regressor 'Crowd' refers to the level at the end of quarter q-1; ΔCrowd_q refers to the change over quarter q; and $\hat{\sigma}_{\text{Crowd}}$ is the estimate of volatility from a GARCH(1,1) model. 'Realized vol. of Mom rets.' is a control variable equal to the lagged realized volatility of daily momentum returns over the previous quarter (intercepts 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 regressors effect on the left tail probability is equal in magnitude and opposite in sign to the (untabulated) effect on the right tail probability.

	Panel A. Predicting the 10% left tail						Panel B. Predicting the 5% left tail					
Dependent variable:	cumulative returns			dynamic FF3 residuals			cumulative returns			dynamic FF3 return residuals		
4qtr Crowding measure:	GTW	BEK	BEKcap	GTW	BEK	BEKcap	GTW	BEK	BEKcap	GTW	BEK	BEKcap
ΔCrowd_q	12.9	20.2	14.6	16.5	20.7	12.3	5.8	31.2	47.4	19.5	1.0	26.9
	(1.1)	(2.1)	(1.1)	(1.3)	(2.0)	(0.8)	(0.4)	(2.1)	(2.1)	(1.2)	(0.1)	(1.2)
	[0.45]	[0.12]	[0.72]	[0.92]	[0.58]	[0.99]	[0.70]	[0.13]	[0.07]	[0.60]	[0.45]	[0.38]
$\text{Crowd}_{q\text{-}1}$	16.1	10.9	-2.6	21.6	10.4	-4.5	10.8	19.7	-15.3	24.9	7.9	-19.5
	(2.0)	(2.0)	(-0.3)	(2.2)	(1.9)	(-0.5)	(1.0)	(1.9)	(-1.0)	(1.8)	(1.2)	(-1.3)
	[0.44]	[0.13]	[0.51]	[0.91]	[0.10]	[0.34]	[0.96]	[0.14]	[0.40]	[0.99]	[0.47]	[0.55]
$\hat{\sigma}_{\text{Crowd}}$	57.1	48.2	5.7	-186.7	86.0	28.1	-77.4	116.8	-7.5	-219.5	106.3	21.5
	(0.4)	(0.8)	(0.2)	(-1.3)	(1.4)	(1.2)	(-0.4)	(1.4)	(-0.2)	(-1.1)	(1.6)	(0.8)
	[0.28]	[0.17]	[0.56]	[0.90]	[0.07]	[0.71]	[0.75]	[0.06]	[0.77]	[0.47]	[0.02]	[0.22]
Realized vol.	14.8	12.8	11.4	11.7	11.9	9.9	12.2	16.6	16.3	10.8	10.5	11.5
of Mom rets.	(3.9)	(4.1)	(3.8)	(3.8)	(3.8)	(3.3)	(3.4)	(3.7)	(3.4)	(3.1)	(3.1)	(3.1)
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]

Table 6: 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. In all cases the crowing measure is constructed using a four-quarter trading history. 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 for the difference between T3 and T1 obtained with the delta method.

	ΔCrowd			Crowd			Realized vol.
	GTW	BEK	BEKcap	GTW	BEK	BEKcap	of Mom rets.
Volatility							
T1	25.7	27.9	32.0	32.6	33.5	21.8	15.3
T2	26.3	27.0	18.5	26.5	19.3	26.3	17.3
T3	25.9	22.9	25.7	16.5	23.2	29.4	38.7
	(0.0)	(-1.0)	(-1.2)	(-3.5)	(-2.2)	(1.8)	(5.7)
Skewness							
T1	-1.8	-2.5	-1.2	-1.7	-2.0	-0.4	-0.3
T2	-1.2	-1.3	-0.5	-1.1	0.0	-2.4	-0.3
T3	-1.5	0.2	-2.1	-0.6	0.0	-1.2	-1.2
	(0.2)	(4.1)	(-0.8)	(1.8)	(3.3)	(-0.9)	(-2.0)
Kurtosis							
T1	15.4	15.4	8.5	10.5	10.5	4.7	4.0
T2	9.0	8.5	4.2	8.2	3.8	15.3	4.1
T3	10.5	5.4	14.1	4.7	5.6	9.7	6.5
	(-1.0)	(-3.6)	(1.2)	(-2.8)	(-2.4)	(1.7)	(2.1)

Table 7: 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 three sets of dependent variables using daily: (1) raw returns on the momentum portfolio; (2) residual returns using the Fama and French three factor (FF3) model; and (3) residual on FF3 using dynamic weights. Each set considers the three indicated crowding measures. The regressor 'Crowd' refers to the level of the crowding measure at the end of quarter q-1; $\Delta\text{Crowd}_\text{q}$ refers to the change over quarter q; and $\hat{\sigma}_{\text{Crowd}}$ is the estimate of volatility from a GARCH(1,1) model. 'Realized vol. of Mom rets.' is a control variable equal to the lagged realized volatility of daily momentum returns/residuals over the previous quarter (intercepts not tabulated). The t-statistics are computed with Newey-West standard errors with 3 lags.

Panel A. Crowding measures constructed using four-quarter trading histories						
Dependent variable:	vol of returns			vol of dynamic FF3 residuals		
Crowding measure:	GTW	BEK	BEKcap	GTW	BEK	BEKcap
$\Delta\text{Crowd}_\text{q}$	-0.06	-0.16	-0.01	-0.05	-0.11	-0.10
	(-0.4)	(-0.9)	(-0.0)	(-0.5)	(-0.6)	(-0.4)
$\text{Crowd}_\text{q-1}$	-0.05	-0.10	0.02	-0.10	-0.09	0.00
	(-0.5)	(-1.8)	(0.2)	(-1.2)	(-2.0)	(0.0)
$\hat{\sigma}_{\text{Crowd}}$	-1.69	0.74	0.71	-0.75	0.51	0.56
	(-0.9)	(0.6)	(1.5)	(-0.6)	(0.6)	(1.9)
Realized vol.	0.77	0.77	0.76	0.74	0.74	0.73
of Mom rets.	(9.1)	(7.3)	(7.8)	(9.1)	(6.7)	(7.5)
Adj-rsquare	63.5%	63.4%	63.8%	59.5%	59.2%	59.8%

Panel B. Crowding measures constructed using one-quarter trading histories						
Dependent variable:	vol of returns			vol of dynamic FF3 residuals		
Crowding measure:	GTW	BEK	BEKcap	GTW	BEK	BEKcap
$\Delta\text{Crowd}_\text{q}$	-0.08	-0.09	-0.03	-0.07	-0.09	-0.06
	(-2.2)	(-2.2)	(-0.3)	(-2.6)	(-2.9)	(-0.7)
$\text{Crowd}_\text{q-1}$	-0.05	-0.02	0.05	-0.02	-0.03	0.05
	(-1.3)	(-0.5)	(1.3)	(-0.6)	(-0.7)	(1.5)
$\hat{\sigma}_{\text{Crowd}}$	0.47	0.44	0.04	0.48	0.32	0.06
	(1.6)	(2.1)	(0.2)	(1.6)	(1.6)	(0.5)
Realized vol.	0.80	0.77	0.79	0.76	0.74	0.75
of Mom rets.	(8.3)	(7.1)	(7.4)	(7.9)	(6.4)	(7.2)
Adj-rsquare	64.6%	65.9%	63.2%	61.3%	63.1%	59.8%

Table 8: Momentum factor returns as a determinant of crowding

Each column represents a predictive regression of a different crowding measure on lag momentum returns and lag momentum realized volatility. Panel A shows the results for four-quarter measures and Panel B for one-quarter measures. Volatility is computed using daily momentum returns (intercepts not tabulated). The t-statistics are computed with Newey-West standard errors with 3 lags.

	Panel A			Panel B		
Crowding horizon:	4qtr			1qtr		
Crowding measure:	GTW	BEK	BEKcap	GTW	BEK	BEKcap
1yr return$_{q-1}$	0.39	0.28	0.28	1.03	0.53	0.74
	(2.6)	(1.1)	(2.3)	(3.5)	(2.0)	(2.2)
1yr return$_{q-5}$	0.53	0.49	0.12	0.53	0.57	0.30
	(3.0)	(2.2)	(1.1)	(2.2)	(1.9)	(0.9)
1yr volatility$_{q-1}$	-0.38	-0.38	-0.03	-0.25	-0.39	-0.09
	(-4.4)	(-2.9)	(-0.6)	(-1.7)	(-2.3)	(-0.7)
1yr volatility$_{q-5}$	0.19	0.08	-0.09	0.31	0.17	-0.22
	(2.4)	(0.6)	(-1.2)	(2.5)	(1.1)	(-1.1)
Adj-rsquare	18.9%	16.3%	18.0%	10.4%	4.7%	18.8%

Table 9: Regressions of momentum return moments on crowd count and crowd capital jointly estimated

Each column presents a regression of the indicated momentum return metric as the dependent variable and the indicated horizon for estimating the crowding measure (1qtr or 4qtr). In the case of 'left-tail' the regression is Probit as in Table 5. Return refers to the dynamic Fama-French 3 factor residual for the Probit and Volatility panels, and to a regression with the dynamic FF3 factors as controls in the Returns panel. All regressors in the row headings are included in each regression. Thus, the regressions in the columns labelled 'Returns' correspond to Table 4 using BEK and BEKcap and dynamic FF3 (but estimated jointly). Likewise, the regressions in the columns labelled 'left-tail' and 'Volatility' correspond to jointly estimated versions of Tables 5 and 7, respectively, for the case of BEK and dynamic FF3 residuals. T-statistics are reported in parenthesis, and the p-values of a Wald test that the effect of a regressor on the left and corresponding right tail (not tabulated) sum to zero are reported for the probits in brackets.

Dependent variable:	Returns		10% left-tail		5% left-tail		Volatility	
Crowding horizon:	1qtr	4qtr	1qtr	4qtr	1qtr	4qtr	1qtr	4qtr
Crowd = BEK								
ΔCrowd_q	-0.11	-0.43	4.4	20.4	9.5	-2.4	-0.10	-0.08
	(-1.7)	(-2.7)	(1.3)	(1.9)	(1.8)	(-0.2)	(-3.0)	(-0.6)
			[0.68]	[0.44]	[0.36]	[0.45]		
$\text{Crowd}_{q\text{-}1}$	-0.04	-0.28	6.1	12.2	15.6	11.6	-0.04	-0.13
	(-0.6)	(-2.8)	(1.3)	(1.9)	(2.2)	(1.4)	(-1.0)	(-2.8)
			[0.20]	[0.70]	[0.05]	[0.52]		
$\hat{\sigma}_{\text{Crowd}}$	0.11	3.41	-5.0	56.9	32.3	37.7	0.37	0.34
	(0.3)	(1.7)	(-0.3)	(0.7)	(1.2)	(0.4)	(2.0)	(0.4)
			[0.77]	[0.07]	[0.74]	[0.09]		
Crowd = BEKcap								
ΔCrowd_q	-0.01	0.12	1.3	4.7	12.9	18.9	0.07	-0.01
	(-0.1)	(0.4)	(0.2)	(0.3)	(1.5)	(0.8)	(1.1)	(-0.1)
			[0.98]	[0.93]	[0.17]	[0.28]		
$\text{Crowd}_{q\text{-}1}$	0.16	0.70	-4.3	-6.9	-9.7	-21.3	0.06	0.12
	(2.3)	(3.1)	(-1.2)	(-0.5)	(-1.4)	(-1.2)	(2.1)	(1.6)
			[0.84]	[0.18]	[0.40]	[0.35]		
$\hat{\sigma}_{\text{Crowd}}$	0.05	-0.84	14.1	26.9	-1.4	30.1	-0.04	0.41
	(0.2)	(-1.1)	(1.3)	(0.9)	(-0.1)	(0.8)	(-0.4)	(1.5)
			[0.24]	[0.53]	[0.44]	[0.78]		
Control:								
Realized vol.	-0.28	-0.35	11.2	12.3	13.3	13.4	0.74	0.69
of Mom rets.	(-2.2)	(-3.1)	(3.4)	(3.5)	(3.4)	(3.2)	(6.2)	(6.1)
			[0.00]	[0.00]	[0.00]	[0.03]		

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 'Mom Gap', and the momentum gap variable orthogonal to the variables shown for the other columns. 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.

	Mom Gap	orthogonal to					
		ΔMom Inst	Win Inst	Crowd			Realized vol. of Mom rets.
				GTW	BEK	BEKcap	
Volatility							
T1	12.8	12.2	13.5	12.7	13.2	13.3	21.0
T2	19.2	19.9	18.7	19.4	19.0	19.2	17.8
T3	38.6	38.5	38.6	38.6	38.6	38.4	35.6
	(6.4)	(6.7)	(6.2)	(6.6)	(6.4)	(6.4)	(3.2)
Skewness							
T1	-0.3	-0.4	-0.3	-0.3	-0.2	-0.5	-0.7
T2	0.0	0.0	0.0	-0.1	0.0	0.3	-0.2
T3	-1.3	-1.3	-1.3	-1.3	-1.3	-1.3	-1.5
	(-2.4)	(-2.2)	(-2.6)	(-2.5)	(-2.7)	(-2.0)	(-1.4)
Kurtosis							
T1	3.3	3.4	3.2	3.4	3.4	3.4	6.4
T2	3.8	3.7	4.2	3.8	3.9	4.8	4.8
T3	6.6	6.6	6.6	6.6	6.6	6.5	8.4
	(2.9)	(2.8)	(3.1)	(2.9)	(3.0)	(2.8)	(1.1)

Internet Appendix to accompany the paper

"Crowded trades and crash risk: The case of momentum"

(Not for publication)

This Internet Appendix contains two robustness checks for the simulation analysis in Section 3 of the paper. Our benchmark is the simulation with Dirichlet concentration parameters $\alpha_i = 12$, i.e., the low $var(k)$ case, and we investigate the impact of changing the distributional assumptions for δ and higher concentration parameters.

First, we ask whether our results are robust to using a uniform distribution for δ instead of a lognormal distribution. In particular, we let δ follow a uniform distribution on $[0.06, 0.12]$, and leave the setting otherwise identical to the one in the paper. The results are reported in Figure IA.1 and Table IA.1. In summary, the results are very similar to those in Section 3's low $var(k)$ case. 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.

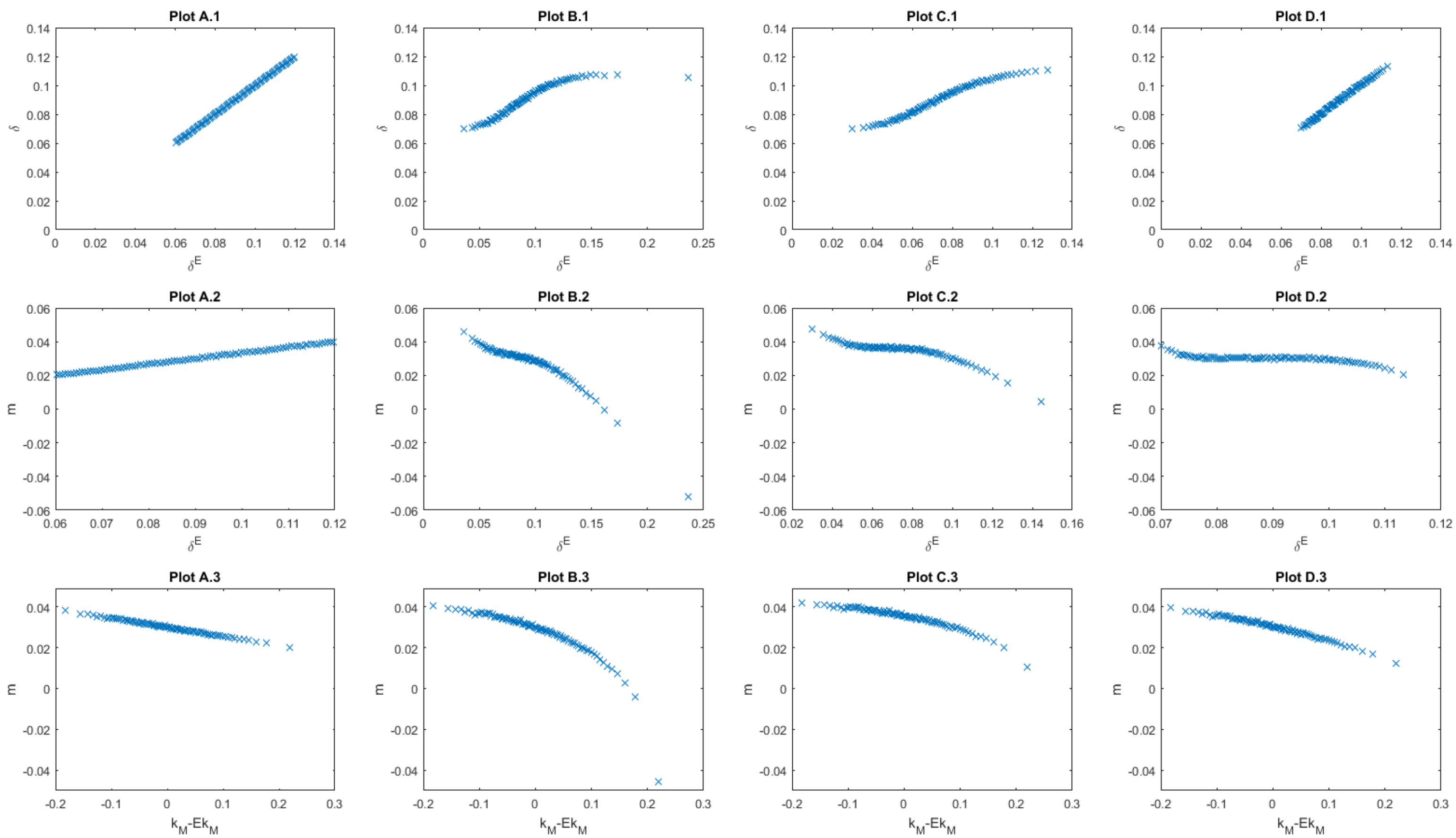

Figure IA.1: Beliefs and expected momentum returns in simulations – uniform distribution

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 $\alpha_i = 12$ (blue Xs), 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.1-3, myopic beliefs in Plots B.1-3, optimal linear beliefs in Plots C.1-3, and rational beliefs in Plots D.1-3. The expected momentum return is then $m = \delta - f$. The triplets $\{\delta, m, k_M\}$ 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.

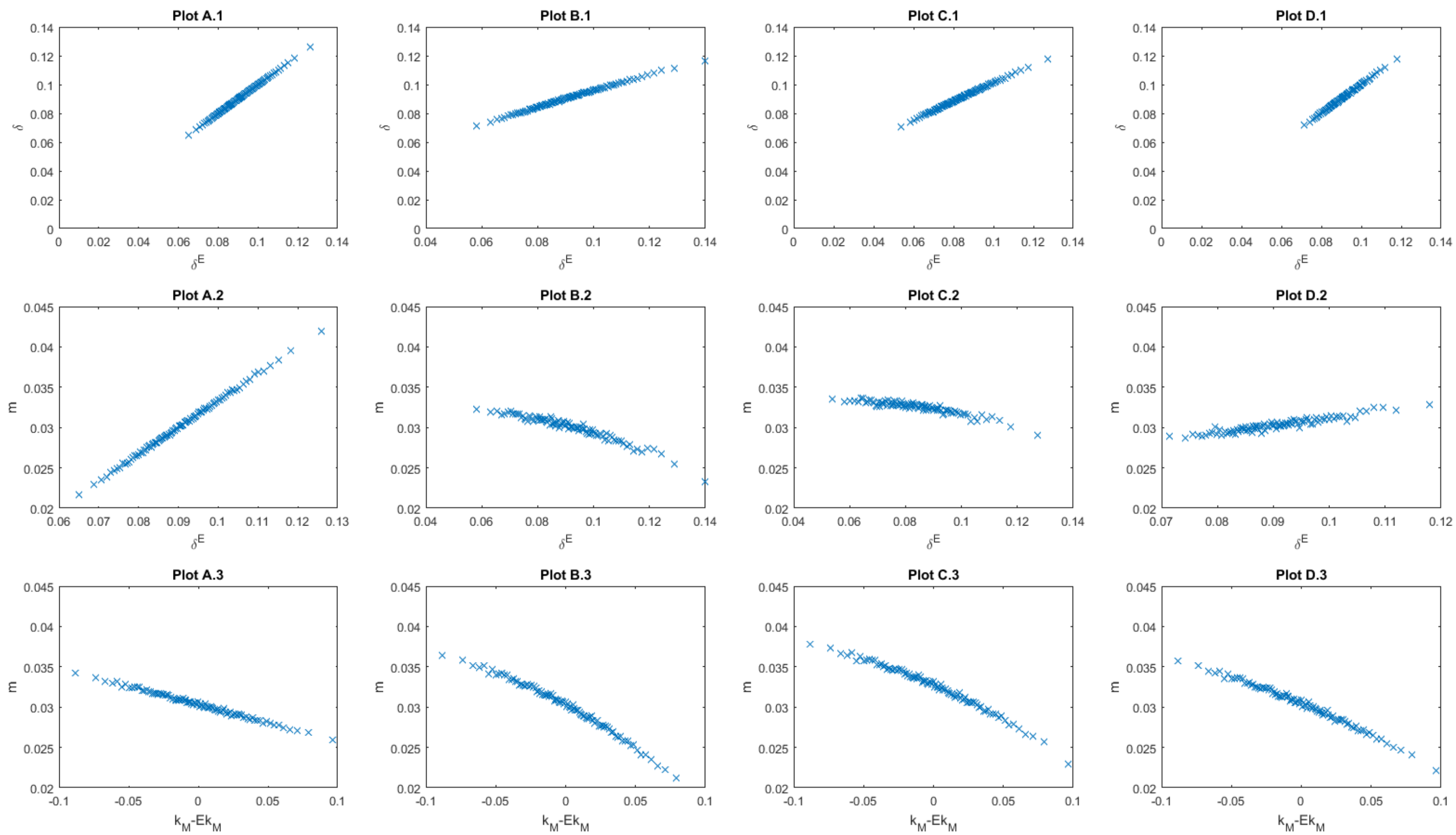

Figure IA.2: Beliefs and expected momentum returns in simulations – very low $var(k)$

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 $\alpha_i = 60$ (blue Xs), which still implies $Ek_I = Ek_M = 1/3$. In addition, as opposed to Figure IA.1 δ follows a lognormal 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.

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. The top part 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. Panel A reports the results for known crowding, Panel B is the case with myopic beliefs, Panel C with optimal linear beliefs, and Panel D with rational beliefs. The Dirichlet distribution has the concentration parameters $\alpha_i = 12$, and δ follows a uniform distribution on $[0.06,\ 0.12]$. The slopes of the 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 $\gamma = 2$ investor, and certainty equivalents 'cer(γ)' are calculated for $\gamma = 2, 4, 10$, with portfolio weights calculated as in (6). Momentum returns are given by $d = m + \epsilon$ 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.

	Panel A	Panel B	Panel C	Panel D
λ^{-1}		1.50	1.34	
Expected momentum returns m				
mean	3.0%	2.8%	3.5%	3.0%
stdev	0.9%	2.7%	1.4%	1.4%
skew	0.5	-92.9	-0.1	0.3
kurt	3.1	17386.3	9.6	2.9
min	0.61%	-534.41%	-35.57%	-2.04%
max	7.51%	9.22%	9.38%	8.56%
Realized momentum returns d				
profit	3.11%	1.75%	1.99%	2.78%
cer(2)	2.30%	-100.00%	1.85%	2.16%
cer(4)	1.14%	-100.00%	0.92%	1.07%
cer(10)	0.45%	-100.00%	0.37%	0.43%

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.

	Panel A	Panel B	Panel C	Panel D
λ^{-1}		1.50	1.44	
Expected momentum returns *m*				
mean	3.0%	3.0%	3.3%	3.0%
stdev	0.5%	0.8%	0.7%	0.7%
skew	0.4	0.2	0.3	0.3
kurt	3.4	3.3	3.3	3.3
min	1.39%	-0.39%	0.41%	0.18%
max	6.13%	7.12%	7.35%	7.35%
Realized momentum returns *d*				
profit	3.02%	2.90%	2.64%	2.91%
cer(2)	2.29%	2.15%	2.24%	2.24%
cer(4)	1.14%	1.07%	1.11%	1.11%
cer(10)	0.45%	0.43%	0.44%	0.44%