

Crowding and the Moments of Momentum

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

May 5, 2017

Abstract

We develop a model of crowding as in Stein (2009), applied to momentum. The model demonstrates how anticipated and unanticipated crowding translate into the first three moments of momentum returns. We test the model's predictions using proxies for momentum crowding developed from 13F data and find that unanticipated crowding predicts negatively momentum returns, particularly as it relates to the number of institutions following a momentum strategy. By contrast, the intensity of momentum trade is less of a factor. Crowding is also negatively related to volatility of momentum returns, consistent with forward looking investors anticipating, and avoiding, strategy risk. Extending Stein (2009) to incorporate non-linear beliefs, we also show theoretically that crowding does not necessarily cause momentum crashes, and our empirical results confirm that unanticipated crowding does not explain the fat left tail of momentum profits.

*Barroso and Karehnke are at the University of New South Wales, Edelen is at UC Davis. Emails: p.barroso@unsw.edu.au, rmedelen@ucdavis.edu, p.karehnke@unsw.edu.au. We thank seminar participants at UC Riverside, University of Arizona, University of New South Wales and University of Technology Sydney for their helpful comments and suggestions. Errors are the authors' responsibility.

1 Introduction

If informed investors are capital constrained and risk averse, they can be expected to leave some of their information on the table, pushing winner (loser) stock prices up (down) only part way toward what they perceive to be fundamental value. Momentum investing can be regarded as an effort to rationally infer this residual information and act on it by extrapolating past differential returns on winner and loser stocks into a signal of as-yet-unpriced deviations in fundamental value. While this strategy further drives prices toward fundamental value, a momentum factor premium will remain if momentum investors are themselves capital constrained and risk averse.

We use this logic to develop a model of momentum factor returns based on the distribution of capital across three investor types: informed investors, momentum investors, and counterparty investors. Informed investors observe heterogeneous, noisy signals of fundamental value for a multitude of stocks. Momentum investors observe no private signals, but they maintain rational expectations regarding the price formation process. Counterparty investors take the other side of both informed and momentum investors' trades, by construction. Our model is closely related to the Stein (2009) model of crowding applied to under-reacting newswatchers. However, the focus here is how capital uncertainty affects return moments rather than equilibrium pricing and efficiency considerations.

The existence of counterparty investors is a concern (e.g., Milgrom and Stokey, 1982) that we, like most studies, do not directly address. We simply assume a random quantity of counterparty capital that willingly accepts investment exposure in proportion to the price concession it receives relative to a public-information valuation. We take the required price concession to be an exogenous parameter meant to reflect both partial aversion to adverse selection and the option to not participate. At extreme values of this parameter, the no-trade theorem obtains.¹

The momentum cycle begins with informed investors privately observing heterogeneous signals of fundamental value for a number of stocks. They then enter a call auction for each stock on the rank date, and seek to trade in proportion to the strength of their private signals. Momentum and

¹Many studies address the irrationality of counterparty trading by assuming it satisfies some unstated exogenous motive that dominates adverse selection concerns (i.e. noise trading). In the resulting equilibrium, informed traders find willing counterparties according to a price-impact schedule parameterized by the variance of noise trading demands. Our approach is essentially a re-parameterization that sidesteps the complexity of a noisy rational expectations framework. The cost is explicit incorporation of irrationality. The benefit is a greatly streamlined model that gets to the same place.

counterparty investors join that auction, and all investors condition on the market-clearing price. When this rank-date market clears, the momentum portfolio is identified, as is its ranking-period return. An evaluation period then begins, and ends with all investors observing a public release of fundamental value. At that point information is again symmetric and the momentum cycle is completed.

A greater quantity of momentum capital applied in the rank-date market should result in a lower momentum premium, since the residual information of informed capital is then diluted across more momentum capital. That is a straightforward implication of supply and demand considerations. As in Stein (2009), the more important result pertains to the role of uncertainty in momentum capital. Our development emphasizes the potential nonlinear nature of unanticipated momentum capital's impact on the return characteristics of momentum, particular the third moment (negative-tail returns).

In a framing similar to Stein (2009), we assume the momentum investors cannot distinguish between abnormally large ranking-period returns that are driven by informed investors' private signals, and abnormally large ranking-period returns that are driven by unanticipated crowding by other momentum investors. They condition on a belief about the relative contribution from each source (information versus crowding), and form demands accordingly. Like Stein, we begin with a linear assumption. We then explore non linearity in beliefs. We find that assuming linear or non-linear beliefs leads to very different predictions for tail risk.

As a result of the ambiguity in the source of ranking-period price moves, unanticipated crowding induces positive feedback trading. With linear beliefs this feedback effect can result in large negative-tail returns. However, we show that momentum traders can protect themselves from these tail events by applying suitable concavity in their mapping from ranking period returns to beliefs about differential fundamental value. Doing so removes the destabilizing effect of crowding, making it an empirical matter of whether crowding predicts crashes.

Two factors combine to determine the aggregate quantity of capital committed to momentum: the number of investors employing the strategy and the intensity of their trade. We conjecture that it is difficult to forecast the decision to adopt - or terminate - the strategy; hence the number of investors who follow it is relatively opaque. However, since all momentum investors extract the same signal from ranking period returns, and (presumably) employ a similar signal extraction technique

(i.e., beliefs mapping), it should be relatively easy to forecast the intensity of trading conditional on having adopted the strategy. That is, trade intensity reflects a relatively homogeneous reaction to a common setting, as formulated in the model. Under this conjecture the number of momentum investors acting at a given time should be the primary source of crowding uncertainty. Alternatively, if assets under management (AUM) is a strong determinant of trading intensity as would be the case under CRRA preferences, it may be that the intensity of capital allocation is also an important source of uncertainty as AUM can differ substantially across investors. Hence, this also is an empirical matter.

Our empirical specification mirrors this decomposition of uncertainty. Using 13 F data we form proxies for momentum capital by characterizing each institution according to whether they seem to follow a momentum strategy or not. Thus, the basic variable is an indicator function, which is then summed across institutions to proxy for the count of momentum participants. Investors know their own type at the time of trade, but not this aggregate count. We generically refer to this as the *count* proxy.

Aggregate momentum capital is the count of momentum investors times the magnitude of a representative momentum trade. Our second empirical proxy, which we here generically refer to as the *capital* proxy, is essentially the *count* proxy times this representative trade intensity. As previously observed, this intensity multiplier reflects an optimal response to current conditions (risk tolerance, information environment, etc), implying different predictions for the *capital* proxy regarding the moments of momentum returns. First, because the representative intensity is presumed known to each optimizing momentum investors, it does not generate the same feedback effects identified by Stein (2009). Second, trading intensity is generally related positively to the momentum return residual to market clearing on the rank date, reflecting optimal anticipatory demands. Thus, the *capital* proxy may have different implications for predicting moments of momentum return than the *count* proxy, and it is the latter that should be most associated with feedback effects from crowding.

To capture anticipated versus unanticipated crowding, we use both a levels and a changes specification in the *count* and *capital* proxies. We also consider a third empirical specification to proxy the uncertainty in unanticipated crowding, using the expected volatility from a GARCH(1, 1) specification of each of the *count* and *capital* proxies. On the returns side, we consider raw momentum

returns as well as the abnormal return on momentum conditioning on the Fama-French three factors and on a dynamic version of the Fama-French model. Finally, to analyze the impact of crowding on the various moments of momentum we consider the level of momentum returns, the volatility, and the probability of a left-tail (extreme negative return) event, as well as skewness and kurtosis in time-series sorts.

We find that unanticipated crowding predicts momentum returns, as hypothesized in the model and proxied with the change in crowding. First, consistent with *count* being the primary source of unanticipated momentum capital, we find that changes in the *count* proxy best predict negatively momentum returns. Generally, we do not find the same negative predictive relation for momentum returns result with the level of *count*, or using changes (or levels) in the *capital* proxy.

Second, and in line with our extension of the model to allow for non-linear beliefs, probit regressions and time-series sorts suggest that crowding fails to explain momentum's largest crashes. While there is a statistically significant positive relation between changes in the *count* proxy and the probability of a large negative momentum return, bivariate probit regressions suggest that this is a shift in mean-effect rather than an asymmetric increase in tail risk. In addition, momentum's tail risk tends to be lower after high crowding in the time-series, inconsistent with crowding generating large asymmetric feedback effects, and pointing to the empirical relevance of our extension of Stein to allow investors to form non-linear beliefs.

Third, momentum investors appear to anticipate – and shun – risk in the momentum strategy. We find that changes in *count* predict negatively momentum return volatility. That is, the number of investors adopting the momentum strategy appears to forecast high risk in that strategy. We also find that the estimated uncertainty (i.e., Garch volatility) in the crowding measures forecasts momentum return volatility.² Finally, consistent with the notion that momentum investors' aversion to crowding risks derives from an anticipated impact on the risk of the strategy, we find that volatility in crowding positively predicts momentum returns indicating compensation for this risk.

The study is outlined as follows. Section 2 provides a brief summary of the existing literature to put the study in context. Section 3 provide the development and results of the model and section 4 explores its predictions. Section 5 concludes the study.

²A control for lagged volatility in momentum returns is included in this and all regressions.

2 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 (Barroso and Santa-Clara, 2015; 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 aim is to investigate the extent to which uncertainty regarding institutional participation in the momentum strategy is a plausible source of these return characteristics.

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 the gradual diffusion of information among investors (see, e.g., Hong and Stein, 1999; Andrei and Cujean, 2017).

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) finds momentum is more pronounced in firms with more media coverage also supporting an over-reaction interpretation. As a whole this evidence suggests investors crowding in momentum stocks can be partly causing the phenomenon, thus momentum becoming stronger when overconfidence exacerbates such crowding.

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.

Our model builds on an under-reaction setting but differs in its delineation of a critical role for momentum capital - particularly unexpected momentum capital. In this sense it is closest to Stein (2009), which develops the theoretical foundation for a destabilizing influence from crowding. In his model arbitrageurs try to exploit the under-reaction of naïve investors but face uncertainty regarding the total amount of arbitrage capital. Stein shows how the inability of arbitrageurs to know in real time how many others are following the same strategy creates a coordination problem. As a result, rather than providing pricing improvement, it can be that arbitrage capital pushes prices further away from fundamentals. As discussed in the introduction, our model builds off of this insight from Stein (2009), with a development focused more specifically on momentum and a framing of empirical analyses. In addition, we also investigate the role of non-linear concave beliefs, which seem to be empirically important.

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. Abreu and Brunnermeier (2002) argue that arbitrage may be limited and delayed due to synchronization risk, i.e., the uncertainty about when the other arbitrageurs will trade.

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 higher in times of lower momentum capital. Our evidence generally supports their finding, with a different approach and insights in proxying momentum capital. 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 4.7 that momentum gap's predictive power for crash risk is unrelated to various institutional measures of momentum crowding. This corroborating our finding that momentum's largest 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. (2016) 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 the subset of momentum investors shrinks after periods of high volatility.

3 Model

Section 3.1 lays out the assumptions and setting of the model. Section 3.2 develops the demands of each investor type, and section 3.3 develops the equilibrium on the rank date. The predictions from this equilibrium are then developed in section 3.4.

3.1 Setting

Stocks are indexed by j . Each pays a discrete dividend $X_{j,t}$. Dividends evolve according to

$$\log(X_{j,t+1}/X_{j,t}) = \chi_{t+1} + \iota_{j,t+1} \frac{\delta_{t+1}}{2}, \quad (1)$$

where χ_{t+1} is a random zero-mean innovation common to all stocks with variance σ_χ^2 that generates the market return; the indicator $\iota_{j,t+1}$ selects the momentum portfolio, taking on the value 1 or -1 for 10% of all stocks (in each leg); and δ_{t+1} generates the differential return on the two groups of stocks, with variance σ_δ^2 and mean $d_{t+1} - \sigma_\delta^2/2$. We do not model an idiosyncratic component to δ_{t+1} , as diversification in implementing a momentum strategy would eliminate its relevance. All investors know σ_δ^2 and σ_χ^2 . We refer to stocks with $\iota_{j,t+1} = 1$ (-1) as long leg (short leg) stocks.

At the beginning of the momentum cycle (time t) information is symmetric, hence each investor rebalances their holdings to the market portfolio. This results in a public-information valuation vector $\mathbf{P}_t = \mathbf{X}_t/r$ where r denotes the required return on the market portfolio. We do not model r . At some intermediate time that we refer to as the rank date (indexed ' $t + Rnk$ ' < $t + 1$), a subset of 'informed' investors observes d_{t+1} and ι_{t+1} and trades at market-clearing prices $\mathbf{P}_{t+Rnk}$. This generates the ranking period return $\mathbf{r}_{t \rightarrow t+Rnk}$ where $r_{j,t \rightarrow t+Rnk} = \log(P_{j,t+Rnk}/P_{j,t})$. Information symmetry is regained at time $t + 1$ when $\mathbf{X}_{t+1}$ is publicly revealed and all investors rebalance to the market portfolio. Because of this reset, we drop the t subscripts and focus on just one momentum cycle, using the subscript $_0$ to denote time t values and the subscript $_1$ for time $t + 1$ values.

$r_{j,0 \rightarrow Rnk} = 0$ for neutral stocks ($\iota_{j,1} = 0$) because all investors maintain belief $E[X_{j,1}] = X_{j,0}$ for such stocks. Because each informed investor has homogeneous expectations of dividend growth for long-leg stocks, $E[X_{j,1}]/X_{j,0} = e^{d_1/2}$, and each uninformed (i.e., momentum and counterparty) investor likewise has homogeneous expectations for long-leg (indeed, all) stocks, $E[X_{j,1}]/X_{j,0} = 1$, each long leg stock experiences the same ranking period return. Denote this return $r_{+,0 \rightarrow Rnk}$. Short-leg stocks likewise have a homogeneous expected return, of the same magnitude and opposite sign. Denote this return $r_{-,0 \rightarrow Rnk}$. Thus, we can compress the returns of individual stocks to a return on the momentum portfolio. Let m denote this return. In the ranking period

$$m_{0 \rightarrow Rnk} = \sum_{\iota_j=+1} \frac{X_{j,0}}{X_{\pm,0}} r_{+,0 \rightarrow Rnk} - \sum_{\iota_j=-1} \frac{X_{j,0}}{X_{\pm,0}} r_{-,0 \rightarrow Rnk} = r_{+,0 \rightarrow Rnk} - r_{-,0 \rightarrow Rnk}, \quad (2)$$

where stocks are weighted by their beginning of ranking-period valuation. Evaluation-period returns on the momentum factor are

$$m_{Rnk \rightarrow 1} = d_1 - m_{0 \rightarrow Rnk}, \quad (3)$$

which follows from

$$\begin{aligned}
\text{long leg:} \quad e^{r_{+,0 \rightarrow Rnk}} e^{m_{+,Rnk \rightarrow 1}} &= e^{d_1/2} \Rightarrow m_{+,Rnk \rightarrow 1} = \frac{d_1}{2} - r_{+,0 \rightarrow Rnk}, \\
\text{short leg:} \quad e^{r_{-,0 \rightarrow Rnk}} e^{m_{-,Rnk \rightarrow 1}} &= e^{-d_1/2} \Rightarrow m_{-,Rnk \rightarrow 1} = -\frac{d_1}{2} - r_{-,0 \rightarrow Rnk}.
\end{aligned}$$

Thus, we seek an expression for $m_{Rnk \rightarrow 1}$ by way of $m_{0 \rightarrow Rnk}$.

Ranking-period returns are determined by equating the demands of three investor groups. Two are proactive, responding to information signals with active trading. The third is reactive, supplying the positions demanded by proactive investors.

Informed investors control beginning-of-cycle capital K_I . They privately observe ι_{t+1} and d_1 and initiate the momentum portfolio on the rank date by trading on that information. Each forms perfect-information rank-date expectations for the momentum factor return

$$E_I \left[m_{Rnk \rightarrow 1} \mid d_1, m_{0 \rightarrow Rnk} \right] = d_1 - m_{0 \rightarrow Rnk}. \quad (4)$$

Momentum investors control beginning-of-cycle capital K_M . They do not observe d_1 or ι_1 , but condition their demands on $m_{0 \rightarrow Rnk}$

$$E_M \left[m_{Rnk \rightarrow 1} \mid m_{0 \rightarrow Rnk}, \cdot \right] = E_M d_1 - m_{0 \rightarrow Rnk}. \quad (5)$$

Counterparty investors trade counter to informed and momentum investors and are therefore adverse selected. A fraction $(1 - L)$ recognize this, exercising their no-trade option by maintaining $E_C m_{Rnk \rightarrow 1} = 0$. The remaining fraction L face an exogenous motive for trade and willingly accept adverse selection as a cost of satisfying that motive. The de facto expectations of a counter-party investor are

$$E_C m_{Rnk \rightarrow 1} = (1 - L) \cdot 0 + L \cdot (-m_{0 \rightarrow Rnk}) = -L \cdot m_{0 \rightarrow Rnk}. \quad (6)$$

Counterparty investors control beginning-of-cycle capital K_C .

All investors hold power utility preferences with relative risk aversion γ , maximizing ³

$$\log E[u(K_1)] = \log E \left[\frac{K_1^{1-\gamma}}{1-\gamma} \right] \quad (7)$$

on the rank date, where K_1 denotes the investor's capital at the end of the momentum cycle. We use the methodology of Campbell and Viceira (2002, appendix) to derive the following expression for demands:

$$Demand = \frac{Em_{Rnk \rightarrow 1}}{\gamma\sigma_\delta^2} K_0, \quad (8)$$

where K_0 denotes capital at the beginning of the momentum cycle. Details are in the Appendix A.1.

3.2 Rank-date pricing of the momentum portfolio

Using each investor type's expectations in (8), the market clearing condition is

$$L \frac{m_{0 \rightarrow Rnk}}{\gamma\sigma_\delta^2} \cdot K_c = \frac{\tilde{d}_1 - m_{0 \rightarrow Rnk}}{\gamma\sigma_\delta^2} \cdot K_I + \frac{E_M \tilde{d}_1 - m_{0 \rightarrow Rnk}}{\gamma\sigma_\delta^2} \cdot K_M. \quad (9)$$

Using $\tilde{k}_{type}$ to denote the *fraction* of capital from the subscripted type, market-clearing implies

$$\tilde{m}_{0 \rightarrow Rnk} = \tilde{d}_1 \cdot \tilde{k}_I + E_M \tilde{d}_1 \cdot \tilde{k}_M. \quad (10)$$

Stein (2009) proposes a linear specification for the beliefs of momentum investors. Let us suppose that they follow this conjecture, i.e., that ranking period returns follow

$$\tilde{m}_{0 \rightarrow Rnk} = \beta \cdot \tilde{d}_1. \quad (11)$$

More specifically, suppose that momentum investors conjecture

$$\beta \equiv E_M (\tilde{k}_I + \tilde{k}_M). \quad (12)$$

That is, the fraction of informed traders' beliefs that gets incorporated into ranking period returns

³We assume that $\gamma > 1$ throughout the paper.

is equal to the fraction of capital that trades with a rational information motive. Let $\tilde{\beta}$ denote the coefficient that results from this conjecture. Using (11) and (12) in (10) and a bit of algebra,

$$\tilde{\beta} = \beta \frac{1 + \tilde{v}_I}{1 - \tilde{v}_M}, \quad (13)$$

where

$$\tilde{v}_I = \frac{\tilde{k}_I - E\tilde{k}_I}{E\tilde{k}_I} \quad (14)$$

$$\tilde{v}_M = \frac{\tilde{k}_M - E\tilde{k}_M}{E\tilde{k}_I} \quad (15)$$

denotes unanticipated informed and momentum capital, both normalized by the expected informed capital. To focus on the effects of crowding uncertainty, we simplify by presuming $\tilde{k}_I \equiv E\tilde{k}_I$ or $\tilde{v}_I = 0$. From (11) and (13) with this simplification

$$E_M \tilde{d}_1 | \tilde{m}_{0 \rightarrow Rnk} = E_M \tilde{\beta}^{-1} \tilde{m}_{0 \rightarrow Rnk} = \beta^{-1} E_M (1 - \tilde{v}_M) \tilde{m}_{0 \rightarrow Rnk} = \beta^{-1} \tilde{m}_{0 \rightarrow Rnk}, \quad (16)$$

as conjectured (thus confirming rational expectations given linear beliefs). Thus,

$$\tilde{m}_{0 \rightarrow Rnk} = \tilde{d}_1 \cdot \beta \frac{1}{1 - \tilde{v}_M}. \quad (17)$$

Momentum factor returns from the rank date to the end of the momentum cycle are

$$\tilde{m}_{Rnk \rightarrow 1} = \tilde{d}_1 \left(1 - \beta \frac{1}{1 - \tilde{v}_M} \right). \quad (18)$$

As in Stein (2009) this analysis presumes a linear setting, i.e. the slope in momentum investors beliefs given observation of ranking period returns, should not vary with $\tilde{m}_{0 \rightarrow Rnk}$. In Figure 1 Panel (a) we have plotted the results of simulations using randomly generated values for $\tilde{d}_1$ and $\tilde{k}_M$ from the distributions detailed in Appendix B. In particular, 5000 independent draws are applied to (10) using (16) as the expression of momentum investors beliefs. In each case the market clearing $\tilde{m}_{0 \rightarrow Rnk}$ is solved by iteration and multiplied by β^{-1} to yield $E_M \tilde{d}_1 | \tilde{m}_{0 \rightarrow Rnk}$. We then sort the resulting 5000 pairs $(\tilde{d}_1, E_M \tilde{d}_1 | \tilde{m}_{0 \rightarrow Rnk})$ by $E_M \tilde{d}_1 | \tilde{m}_{0 \rightarrow Rnk}$ and form 100 ranked subsets. We compute the average $\tilde{d}_1$ within each subset and plot it on the vertical axis against the corresponding average

$E_M \tilde{d}_1 | \tilde{m}_{0 \rightarrow Rnk}$ on the horizontal axis. The vertical axis is a proxy for the true conditional expectation for $\tilde{d}_1$ at the indicated level of $\tilde{m}_{0 \rightarrow Rnk}$, whereas the horizontal axis is momentum investors' beliefs under an optimal linear strategy. Thus, the figure provides an indication of how $E_M \tilde{d}_1 | \tilde{m}_{0 \rightarrow Rnk}$ performs as a conditional expectation of $\tilde{d}_1$ given $\tilde{m}_{0 \rightarrow Rnk}$. If $E_M \tilde{d}_1 | \tilde{m}_{0 \rightarrow Rnk}$ was a valid conditional expectation it should yield an identity relation – linear with a slope of 1.

[Insert Figure 1 near here]

Figure 1 Panel (a) shows that linear beliefs lead to a systematic over-extrapolation of ranking period returns, $\tilde{m}_{0 \rightarrow Rnk}$ —increasingly so at higher values of $\tilde{m}_{0 \rightarrow Rnk}$. For example, in cases where the market beliefs regarding $\tilde{d}_1$ are 4, the actual $\tilde{d}_1$ averages 2.3. The source of this exuberance is unanticipated crowding, $\tilde{v}_M$. High unanticipated momentum demand $\tilde{v}_M$ impart additional price pressure on the rank-date valuation of the momentum portfolio. Momentum investors interpret this as an incrementally favorable private signal of $\tilde{d}_1$, and trade even more aggressively. When momentum investors employ linear beliefs, this feedback trading generates a strong nonlinear $(1 - \tilde{v}_M)$ denominator term in the market clearing ranking-period return (17). Note, from the first graph in Figure 1 Panel (a) that in the most extreme 2% cases of the simulations the beliefs regarding d_1 under the linear specification approaches infinity, when in fact d_1 plateaus at around 2.3. The result is arbitrarily large negative returns in the evaluation period (18). This can be seen in the second graph of Figure 1 Panel (a). Thus, under a linear specification of beliefs, unbounded crashes in momentum returns are predicted.

The problem is, with linear beliefs higher ranking period returns forever translate into higher expected dividends when in fact, higher ranking period returns increasingly imply a greater influence of unanticipated crowding (not signals of fundamental value). If momentum investors incorporate the concavity revealed in Figure 1 into their inferences of $\tilde{d}_1 | \tilde{m}_{0 \rightarrow Rnk}$, they will temper their inference of a favorable private signal of $\tilde{d}_1$ at high values of $\tilde{m}_{0 \rightarrow Rnk}$. This moderates the feedback that unanticipated crowding has on rank-date pricing. Indeed, if momentum investors apply just the right amount of concavity to their inferences, the plot in Figure 1 should approximate the identity function (linear with a slope of 1). In that case, the conditional expectation $E_M \tilde{d}_1 | \tilde{m}_{0 \rightarrow Rnk}$ that momentum investors employ in forming their demands yields a market clearing price that consistently forecast $\tilde{d}_1$, at all levels of $m_{0 \rightarrow Rnk}$.

We reverse this logic and presume that momentum investors apply a power function to $\tilde{m}_{0 \rightarrow Rnk}$ in forming expectations of $\tilde{d}_1$. Thus, momentum investors form beliefs

$$E_\rho \tilde{d}_1 \equiv b^{-1} \left(\tilde{m}_{0 \rightarrow Rnk}^{1-\rho} / (1-\rho) \right), \quad (19)$$

where the ρ subscript indicates the degree of concavity presumed. With $\rho = 0$ beliefs are linear, and as ρ goes to 1 beliefs become logarithmic. Let ρ^* denote the equilibrium concavity, i.e., where E_{ρ^*} equates to the true conditional expectation of $\tilde{d}_1$. using ρ^* , beliefs yield a Figure 1 plot of the average d as a function of beliefs that is a line at 45 degrees. Figure 1 Panel (b) depicts an example of such beliefs using a concavity parameter ρ chosen to temper inferences of d_1 at higher values of $m_{0 \rightarrow Rnk}$. Note that when momentum investors follow these concave beliefs, there is no explosive reaction of beliefs to crowding, nor are there precipitous crashes in value when fundamental value is realized.

Beliefs E_{ρ^*} eliminate the destabilizing influence of unanticipated momentum crowding. To see this, note that ranking period returns $\tilde{m}_{0 \rightarrow Rnk}$ can derive from either private signals of $\tilde{d}$ or unanticipated crowding. When beliefs are formed according to E_{ρ^*} , investors overreact to $\tilde{m}_{0 \rightarrow Rnk}$ when it derives from unanticipated crowding to exactly the same degree that they underreact to $\tilde{m}_{0 \rightarrow Rnk}$ when it derives from private information about $\tilde{d}$. Thus, unanticipated crowding does not destabilize, it only introduces noise. However, when momentum investors form beliefs according to $\rho < \rho^*$, momentum investors react too strongly to large values of $\tilde{m}_{0 \rightarrow Rnk}$, so that unanticipated crowding generates a disproportionate feedback response.

Result 1 *If momentum investors apply insufficient concavity to their expectations of d given m , their demands will take the rank-date pricing of the momentum portfolio past the point where the evaluation-period returns on that portfolio fairly compensate for the risk in the momentum strategy. Indeed, at sufficiently large values of ranking period returns this endogenous source of pricing error from arbitrage activity can lead to an arbitrarily large over (under) valuation of winner (loser) stocks, resulting in a precipitous negative return on the momentum portfolio.*

In particular, linear beliefs generate momentum crashes.

Conversely, if momentum investors apply the correct amount of concavity to their expectations of d given m , the conditional expectation of d given m that results from their demands aligns with

their expectations and there are no momentum crashes.

3.3 Model results

Let us first focus on the β term in (18), holding $\tilde{d}$ constant and for the moment ignoring unanticipated momentum capital (i.e., assume $\tilde{v}_M = 0$). β is the fraction of the information of informed traders that is incorporated into ranking period return. A higher value of expected momentum capital, $E\tilde{k}_M$, implies a greater degree to which $\tilde{d}$ is incorporated into ranking period returns, which in turn implies a lower residual remaining in the form of a momentum return. We summarize this as

Result 2 *The diffusion of information into ranking period returns is increasing in the expected momentum capital, $E\tilde{k}_M$, implying a negative relation between $E\tilde{k}_M$ and momentum returns, where*

$$\left. \frac{\partial \tilde{m}_{Rnk \rightarrow 1}}{\partial E\tilde{k}_M} \right|_{\tilde{d}_1 \text{ fixed}, \tilde{v}_M=0} = \tilde{d}_1. \quad (20)$$

That is, an increase in anticipated momentum capital causes a constant proportional decrease in momentum returns.

Now consider the effects of $\tilde{v}_M$ on momentum returns. To explore this term, first recall that momentum investors form beliefs of end-of-cycle dividends by extrapolating ranking period returns under the conjecture (11). Using expression (17) for the $\tilde{m}_{0 \rightarrow Rnk}$ that results from this conjecture, momentum investors' beliefs are

$$E_M \tilde{d}_1 = \beta^{-1} \tilde{m}_{0 \rightarrow Rnk} = \tilde{d}_1 \cdot \frac{1}{1 - \tilde{v}_M}. \quad (21)$$

That is, the $\tilde{d}_1$ that momentum investors infer from ranking period returns is distorted by unanticipated crowding, $\tilde{v}_M$. The larger $\tilde{v}_M$, the greater the distortion. Potentially, this distortion is unbounded as $\tilde{v}_M$ goes to 1. Now observe that, from the demand expression (8), the aggressiveness with which each dollar of momentum capital is allocated to the momentum trade is determined by the intensity of beliefs. Putting the two observations together, we see that unanticipated momentum capital $\tilde{v}_M$ causes *each dollar* of momentum capital to be allocated more aggressively. This generates a positive feedback effect that amplifies the assimilation of $\tilde{d}_1$ into ranking period returns and attenuates the residual momentum return.

Result 3 *Unanticipated momentum capital $\tilde{v}_M$ is negatively related to momentum returns by way of the distorting effect that $\tilde{v}_M$ has on inferences of $\tilde{d}_1$ from ranking period returns:*

$$\left. \frac{\partial \tilde{m}_{0 \rightarrow Rnk}}{\partial \tilde{v}_M} \right|_{\tilde{d}_1, \beta} = \frac{\tilde{d}_1 \beta}{(1 - \tilde{v}_M)^2}. \quad (22)$$

Because this becomes increasingly large at high values of $\tilde{v}_M$, the negative effect of $\tilde{v}_M$ on momentum returns is potentially unbounded implying an elevated probability of large negative momentum returns (fat left tails). In particular, when $\tilde{v}_M$ exceeds the anticipated supply of counterparty capital

$$\tilde{v}_M > 1 - E_M(\tilde{k}_I + \tilde{k}_M) = L \cdot E_M \tilde{k}_C \quad (23)$$

momentum investors extrapolate ranking period returns past fundamental value. At that point, momentum trading becomes feedback trading resonating with itself. By contrast, at low (negative) values of $\tilde{v}_M$, the distortion effect attenuates so the right tail in the distribution of momentum returns inherits the distribution of $\tilde{d}_1$.

Result 3 outlines three hypotheses: (1) unanticipated momentum capital predicts negatively momentum returns; (2) the effect of unanticipated momentum capital is potentially stronger than that of anticipated momentum capital (which is constant and bounded); and (3) unanticipated momentum capital predicts a fat left tail in the distribution of momentum returns. In the next section we explore these predictions as well characterize the relation between momentum capital and returns more generally.

4 Empirical section

We base our empirical analysis on the trading of institutions. Because institutions dominate equities trading, we seek empirical proxies that retain their meaning even if institutions trade only with themselves. Thus, the empirical analog to the quantity of momentum capital in quarter q , $\tilde{k}_M$, should refer to an ex ante measures of capital committed to momentum during that quarter, rather than to the aggregate realized institutional purchase/sale of momentum stocks. In the model we have framed $\tilde{k}_M$ as a random number of institutions engaging in momentum, each of whom

trades with a deterministic demand curve. We set up our empirical proxies likewise, considering crowding measures based on both the count of institutions following a momentum strategy (which we think of as the most plausible source of unanticipated crowding), and on the amount of capital backing up that strategy (which we think of as a more deterministic overlay in intensity of demands related to, for example, risk tolerance). Of course, it is certainly possible that this intensity factor is the primary source of uncertainty, as in the Stein (2009) model. Our specification facilitates inference on the source.

4.1 Momentum behaviour, crowding, and returns

We use quarterly holdings from the Thomson Reuters Institutional (13F) database starting at the end of the first quarter of 1980 until the end of the third quarter of 2015, and stock data from CRSP in the same period. All prices and shares held by institutions are adjusted with the CRSP adjustment factors. To construct momentum trading variables we use indicator variables for the standard winner and loser portfolios defined using NYSE decile breakpoints. We consider only common stocks (CRSP share code 10 and 11) listed on AMEX, NYSE and Nasdaq, and construct our measures using flow-adjusted net purchases of momentum stocks.

Daily and monthly momentum returns are obtained from Kenneth French’s online data library, for March 1980 through December 2015. The momentum return at time t is defined as the return of the winners (those in the top 10% of the distribution according to returns from months $t - 12$ to $t - 2$) minus the return of the losers (those in the bottom 10% of the same distribution).⁴ The decile cut-off points are determined using only NYSE listed stocks to avoid undue influence of micro-cap stocks. The returns of the winners and the losers portfolios are value-weighted within each decile.

At the end of quarter q , 13F institution i ($i = 1, \dots, N_q$) has capital under management

$$K_{i,q} \equiv \sum_{j=1}^J P_{j,q} w_{i,j,q},$$

where $j = 1, \dots, J$ indexes stocks and $w_{i,j,q}$ is the adjusted number of shares held by institution i in

⁴The last month is skipped to avoid confounding with the short-term reversal effect of Jegadeesh (1990).

stock j , and $P_{j,q}$ is the adjusted price of stock j . The net flow of institution i is

$$Flow_{i,q} = K_{i,q} - K_{i,q-1}(1 + r_{i,p,q}),$$

where $r_{i,p,q} \equiv \sum_{j=1}^J \frac{P_{j,q-1}w_{i,j,q-1}}{K_{i,q-1}}r_{j,q}$ is the return on the portfolio held at the end of quarter $q-1$ and $r_{j,q}$ is the return on stock j during quarter q . We allocate flows to the beginning of period portfolio to measure buying net of flow-induced trading. The expected number of shares held at quarter end q given flow is $1 + Flow_{i,q}/K_{i,q-1} = K_{i,q}/K_{i,q-1} - r_{i,p,q}$, and the flow-adjusted net purchase of momentum stocks is

$$Buy_{i,q} = \sum_{j=1}^J P_{j,q} \left(w_{i,j,q} - w_{i,j,q-1} \left(\frac{K_{i,q}}{K_{i,q-1}} - r_{i,p,q} \right) \right) \iota_{j,q}, \quad (24)$$

where $\iota_{j,q}$ is an indicator variable that takes the value 1 if the stock is in the winner decile on the portfolio formation date q , -1 if it is in the loser in the decile, and zero otherwise. $Buy_{i,q}$ measures trading up to the rank date. We also consider a measure of slower-moving momentum capital $BuyPl_{i,q}$ which measures buying 'plus one' quarter after the rank date:

$$BuyPl_{i,q} = \sum_{j=1}^J P_{j,q+1} \left(w_{i,j,q+1} - w_{i,j,q} \left(\frac{K_{i,q+1}}{K_{i,q}} - r_{i,p,q+1} \right) \right) \iota_{j,q}. \quad (25)$$

The variables $Buy_{i,q}$ and $BuyPl_{i,q}$ do not directly proxy momentum capital. They are preliminary variables that we use to categorize institutions as momentum traders with the indicator function $\mathbb{1}_{[\cdot]}$. For example, $\mathbb{1}_{Buy_{i,q}} = 1$ if $Buy_{i,q} > 0$ and zero otherwise.

The third preliminary variable used in constructing our crowding measures is a measure of capital intensity

$$Cap_{i,q} = \sum_{j=1}^J P_{j,q} w_{i,j,q} \iota_{j,q}.$$

$Cap_{i,q}$ is meant to reflect the optimal trading of a representative momentum investor as set up in the model. Because all other momentum investors likewise follow this representative demand intensity, $Cap_{i,q}$ does not represent a source of randomness in the model; crowding uncertainty stems from the number of peer institutions that adopt a momentum strategy, $\mathbb{1}_{[\cdot]}$. Of course, this sourcing of crowding uncertainty need not be the case; either count or intensity could drive ambiguity in

investors' anticipation of concurrent crowding. Our proxies are designed to shed light on this empirical matter.

We construct two sets of proxies based on the construction of $\mathbb{1}_{[\cdot]}$, generically referenced as $_1\text{qtr}$ and $_4\text{qtr}$ measures. The $_1\text{qtr}$ set are defined as

$$\text{Cnt_1qtr}_q = \frac{1}{N_q} \sum_{i=1}^{N_q} \mathbb{1}_{\text{Buy}_{i,q}}, \quad (26)$$

$$\text{CntP1_1qtr}_q = \frac{1}{N_q} \sum_{i=1}^{N_q} \mathbb{1}_{\text{BuyP1}_{i,q}}, \quad (27)$$

$$\text{Cap_1qtr}_q = \frac{\sum_{i=1}^{N_q} \text{Cap}_{i,q} \mathbb{1}_{\text{Buy}_{i,q}}}{\sum_{i=1}^{N_q} K_{i,q}}, \quad (28)$$

these measures classify institutions as momentum investors based only on trading in the most recent quarter. The $_4\text{qtr}$ set of measures are defined similarly

$$\text{Cnt_4qtr}_q = \frac{1}{N_q} \sum_{i=1}^{N_q} \mathbb{1}_{\sum_{l=0}^3 \mathbb{1}_{\text{Buy}_{i,q-l}}=4}, \quad (29)$$

$$\text{CntP1_4qtr}_q = \frac{1}{N_q} \sum_{i=1}^{N_q} \mathbb{1}_{\sum_{l=0}^3 \mathbb{1}_{\text{BuyP1}_{i,q-l}}=4}, \quad (30)$$

$$\text{Cap_4qtr}_q = \frac{\sum_{i=1}^{N_q} \text{Cap}_{i,q} \mathbb{1}_{\sum_{l=0}^3 \mathbb{1}_{\text{Buy}_{i,q-l}}=4}}{\sum_{i=1}^{N_q} K_{i,q}}, \quad (31)$$

except that they employ a more stable and persistent (and less timely) classification of institutions, requiring consistent momentum trading for four consecutive quarters. Note that both the $_1\text{qtr}$ and the $_4\text{qtr}$ versions of the Cap_ measure observe trade intensity only in the most recent quarter; i.e., the measures differ only with respect to the classification of institutions.⁵ In short, Cnt_ measures seek to identify the fraction of 13F institutions supplying momentum capital to the market, whereas Cap_ measures add an overlay to identify the representative intensity with which that capital is applied. In the model the former is the source of ambiguity in crowding; the latter is a deterministic function of the setting.

In addition to these various measures for crowding we consider three specifications, as per the

⁵On occasion we generically refer to these as Cnt_ and Cap_ measures.

model. Thus, Crowd_{q-1} is a generic reference to the level of the measure, and is our proxy for anticipated crowding. ΔCrowd_q is a generic reference to the change in the measure, and is our proxy for unanticipated crowding. $\hat{\sigma}_{\text{Crowd}}$ is a generic reference to the expected volatility of the measure using a GARCH(1, 1) specification, and is our proxy for crowding uncertainty. ΔCrowd_q is generally our primary variable of interest, as it proxies the key dimension of crowding as goes the moments of momentum factor returns.

Table 1 provides summary statistics for the 13F data (Panel A); proxies for momentum investing (Panel B); and returns of the strategy (Panel C). In Panel A, an institution is considered a momentum investor if they are classified as such by one of our measures for at least 2/3 of the available quarters. By this determination only 22% (1414/6360) of institutions consistently follow a momentum strategy. By contrast, Grinblatt et al. (1995) find that 77% of mutual funds are momentum investors. The difference is likely attributable to a difference in definition and the fact that 13F data is at the institution, rather than portfolio, level.⁶

[Insert Table 1 near here]

Also from Table 1, Panel A, momentum institutions differ materially from other institutions. Momentum institutions have a higher turnover (24% compared to 21%), manage more assets (2.46 billion versus 1.23 billion), and hold a more diversified collection of portfolios (213 stocks on average versus 123).

Table 1, Panel B provides descriptive statistics of the crowding variables. Using the mean of either Cnt_- measure with the Crowd_{q-1} specification at a 1qtr horizon, we infer that approximately 50% of institutional investors are classified as momentum investors in a given quarter. This suggests the importance of identifying *consistent* momentum investors rather than just aggregating institutional trading in momentum stocks—in a given quarter the aggregate is roughly in balance. Nevertheless, the identity of institutions trading with momentum is persistent, as can be seen with the 4qtr measures. These average 11.7% for Cnt_- and 10.2% for CntP1_- , compared with the $0.5^4 = 6.25\%$ value implied under a null hypothesis of no momentum-trading institutions (each trades randomly with and against momentum in a given quarter). Also from Panel B, most crowding variables show strong persistence as captured by their coefficients in an AR(1) regression.

⁶Grinblatt et al. (1995) define momentum investors each quarter if they buy the winners and sell the losers as defined by the returns over that same quarter.

Given this evidence of persistence, we estimate the volatility of the crowding measures using the residuals from the respective AR(1) regressions. We adopt the usual GARCH(1,1) specification of Bollerslev (1986), to estimate the conditional volatility of the residuals.

Table 1, Panel C provides the output of a regression of momentum returns using two risk models: the Fama-French 3 factor model (abbreviated FF3) and a dynamic version of the same model (DFF3). We consider these specifications, as well as raw momentum returns, in parallel throughout our analyses.⁷ Momentum has a negative exposure to the market, size, and value factors, hence risk-adjusted momentum returns are even larger than raw returns (monthly alpha 1.4 - 1.6%, as in Asness et al. (2013) and Asness et al. (2014)).

Grundy and Martin (2001) show that the momentum portfolio has strongly time-varying risk exposures due to its rapidly changing composition. In the DFF3 specification we include regressors with an interaction dummy variable that takes the value 1 if the factor has a positive return in the previous year and zero otherwise (D preface in the variable names). Consistent with Grundy and Martin (2001), we find that the betas of momentum respond significantly to lagged returns on the market (t-stat of 2.8) and the value factor (t-stat of 2.1). The dynamic model better captures the three-factor risks of momentum, as the R-squared of the regression increases from 12% to 25%. However, as pointed out in Barroso (2014), time-varying risk exposure does not explain the alpha of the strategy.

In unreported results we also find that momentum has substantial crash risk in our sample, combining high excess kurtosis with pronounced left-skewness. This is unsurprising as our sample period includes the eventful momentum crash of March-May 2009.

Table 2 Panel A considers the persistence of our momentum measures using the transition probability for the classification of an individual institution, e.g., $\mathbb{1}_{Buy_{i,q}}$ and $\mathbb{1}_{\sum_{l=0}^3 \mathbb{1}_{BuyPI_{i,q-l}}=4}$. For all measures the probability of maintaining the current momentum classification in the following quarter is more than 50%, and it is 69% (64%) according to _4qtr measures. This implies persistence. Transition probabilities for momentum versus non-momentum institutions are more comparable at a four-quarter horizon when using _4qtr measures. Using the case of Cnt_, the probability of transitioning to a momentum-trading institution four quarters ahead is 29% for an institution currently

⁷We often make reference to FF3 or DFF3 models or residuals. In fact, in the case of returns regressions the dependent variable is the raw momentum return and the FF3 or DFF3 regressors are included as controls, and for the crash and volatility analysis the residuals from the FF3 or DFF3 models are used.

classified as momentum, versus an unconditional probability of 11% and a probability of 9% for a current non-momentum trader. The results in the case of CntP1_ are similar. Thus, the _4qtr measures provide a meaningful identification of momentum-trading institutions.

[Insert Table 2 near here]

One possible concern is that investors' preference for certain sectors or investment styles make them trade persistently in the direction of (or against) momentum for several quarters in a row. This possibility is challenged by the rapidly changing composition of the momentum portfolio itself. To illustrate, Table 2 Panel B shows the persistence of stocks' membership in either leg of the momentum portfolio. Winners have a 55% 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 64% of loser stocks retaining that classification the following quarter and 31% retaining it after four quarters.

All in all, Tables 1 and 2 suggest that our proxies for crowding – particular _4qtr measures – exhibit a level of persistence at the individual institution level that suggests successful identification of institutions maintaining a purposeful allocation of capital towards an ever-changing momentum portfolio. The question then is whether these proxies can be used to identify the crowding effects detailed in the model.

4.2 Crowding and conditional expected returns on the momentum factor

Table 3 shows the results of predictive regressions of momentum returns on the various crowding measures. All momentum trading measures are appropriately 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 Cnt_1qtr_q to predict momentum returns in the quarter $q + 1$, and the change in CntP1_1qtr_q to predict momentum returns in $q + 2$. As a control we include the lagged realized volatility of momentum computed from the squared daily returns of the WML (winner minus loser) portfolio in the previous quarter. Barroso and Santa-Clara (2015) show that this strongly predicts (negatively) momentum returns.⁸ We consider three

⁸In 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.

specifications for the dependent variable – the return on the momentum factor (raw), and its risk-adjusted returns using both the static (FF3) and dynamic (DFF3) Fama-French 3 factor models. Finally, because computing the regressors requires up to six quarters of data, the regression sample begins in September 1981 and ends in December 2015.

[Insert Table 3 near here]

The model predicts that both anticipated crowding (proxied with $Crowd_{q-1}$) and unanticipated crowding (proxied with $\Delta Crowd_q$) negatively relate to momentum returns, as momentum capital pushes rank-date valuations closer to fundamental value. However, the effect for unanticipated crowding should be larger for three reasons: (1) it cannot be countered ex ante with an optimal demand response, (2) it induces a feedback effect as pointed out in Stein (2009), and (3) it potentially triggers a reversal of positions as momentum investors subsequently learn that their collective investment in the strategy is too high. We see statistically significant evidence of this at the 1% level⁹ in Table 3 with each of the three specifications of the dependent variable—with two categorical exceptions. First, the statement applies in Panel A with the $_4qtr$ identification measures, but not in Panel B where statistical reliability is lost with the noisier $_1qtr$ identification measures. Second, the statement applies to $Cnt_$ and $CntPl_$ measures but not the $Cap_$ measure. This difference is consistent with the model set up in which momentum investors find it relatively difficult to condition on the number of peers engaging in the strategy, but relatively easy to condition on how intensely each will choose to trade under current circumstances.

The regressions in Table 3 do not present an unambiguous relation for the levels specification, $Crowd_{q-1}$. While the estimates are generally consistent with a negative relation, in the case of $Cap_$ the estimate is significantly positive. If $Crowd_{q-1}$ indeed relates to the anticipated level of momentum capital, and $Cap_$ indeed relates to the anticipated intensity of deployment, then this result potentially fits the model. To wit, anticipated capital does not generate feedback, attenuating the negative influence of demands. If that attenuation is sufficient, it is possible that the positive relation between returns and the optimal response of the representative demand curve generates the positive coefficient on $Cap_$. This may be pushing the interpretation of the proxies beyond

⁹Unless 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.

their capacities. Nevertheless, this intriguing possibility is further confirmed when we estimate the coefficients on Cnt_ and Cap_ measures jointly in Table 8.

Finally, from Table 3 the anticipated volatility of crowding from the GARCH (1, 1) forecast, $\hat{\sigma}_{\text{Crowd}}$, is positively related to the expected return of momentum at the 1% level for one specification (CntP1_ and DFF3 returns), but the relation is generally insignificantly positive. A positive relation potentially indicates that uncertainty about the number of competing momentum investors inhibits participation in the strategy, but that inference is at best cloudy. In our later consideration of volatility in momentum returns (Table 6) we find evidence consistent with this story, but again the evidence is cloudy.

4.3 Crowding and negative tail events in momentum returns

Of particular interest is the relation between unanticipated crowding in momentum and the pronounced left tail in the distribution of momentum returns. This left tail is well-known empirically, and shown in the model to be an implication of the perverse feedback effect of crowding identified by Stein (2009). To assess this link between the left tail of momentum returns and crowding, we run probit regressions using indicators for a return below the 10th and below the 5th percentile of the distribution, controlling for lagged volatility of momentum returns. We also test whether probit right tail results of the crowding variables are distinct from the expected return results in Table 3, which would shift the overall distribution to the left. This is tested by estimating a bi-variate probit specification including the corresponding right tail and constructing a Wald test for whether the coefficients in the left and right tail regression sum to zero. Here, a downward shift in mean corresponds to a significant positive regressor in the left-tail regression, and a failure to reject the null in the Wald test.

[Insert Table 4 near here]

Table 4 presents the results. To save space, Table 4 only considers the _4qtr measures and the raw and DFF3 return specifications. Consistent with the argument that unanticipated peers are the primary source of surprises in crowding, we find a statistically reliably positive relation between ΔCrowd_q and subsequent left-tail events using Cnt_ and CntP1_ measures for the 10% left tail, but weaker evidence for the 5% tail. The relation between the change in Cap_ measures and left-tail

events is also positive, but not statistically reliable. This weaker link is again consistent with the conjecture that capital intensity is endogenous to the optimized demands of momentum traders, who each follow a deterministic demand curve; whereas the number of peers competing for the trade is exogenous to the traders' calculations.

Neither the level (Crowd_{q-1}) nor the volatility of crowding ($\hat{\sigma}_{\text{Crowd}}$) reliably predicts left-tail events. This is to be expected if Crowd_{q-1} indeed proxies for anticipated crowding, as anticipation prevents a skewed impact on momentum returns. Likewise, prediction regarding the volatility of crowding is not nearly so sharp as predictions regarding the direct measure of unanticipated crowding, ΔCrowd_q .

The effect of unanticipated crowding, however, is in most cases not statistically distinguishable from a shift in mean effect, casting doubt on whether ΔCrowd_q really predicts a fat left tail in momentum returns in line with Stein's (2009) argument regarding the perverse effects of unanticipated crowding, or rather a downward shift in mean. This lends support to our extension of the model to concave ranking period beliefs reducing the destabilizing impact of unanticipated crowding. While crowding does seem to matter for asset pricing, the probit regressions and Wald tests suggest that there is only limited evidence for the large crowding-induced crashes predicted by linear beliefs.

[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 behaviour 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. So the momentum strategy in equities does not seem to be particularly crowded around the event of its major crash in our sample period.

For a direct examination of the relation between momentum crashes and (un)anticipated crowding, we sort momentum returns in the time-series according to the lagged level and changes in crowding. Momentum's tail risk takes the form of pronounced left skewness combined with excess kurtosis. We also estimate conditional volatility to enable an interpretation of the standardized

measures skewness and kurtosis in terms of total risk.¹⁰ A thorough analysis of the predictive relation between crowding and volatility is conducted in the next section.

[Insert Table 5 near here]

Table 5 shows the volatility, skewness, and kurtosis of the strategy conditional on lagged crowding variables and momentum's own lagged realized volatility (as computed from daily momentum returns in the previous quarter). Periods after high changes in crowding do not have significantly different volatility from months after low changes. There is also no evidence supporting higher crash risk after high unanticipated crowding. None of the differences for skewness is significant and for kurtosis the only one that is significantly different from zero is for the Cnt_ measure and it goes in the opposite direction - suggesting unanticipated crowding reduces fat tails. High unanticipated crowding, as captured by Cnt_ or CntP1_, predicts less left skewness and smaller kurtosis. The evidence in Table 3 is consistent with a coordination problem for momentum investors as in Stein (2009)'s model. The analysis with sorts suggests though that this coordination problem does not play a central role in explaining momentum crashes.

Columns 4 to 6 of Table 5 show the results for the levels of crowding. Here, the results for Cap_ are not significant for any of the moments of momentum. For Cnt_ and CntP1_ the results are statistically significant and all go against the hypothesis of crowding causing (or at least predicting) higher risk in the strategy. After months with high crowding, as measured by the percentage of investors following momentum strategies, volatility is lower, skewness is higher and kurtosis is smaller.

Barroso and Santa-Clara (2015) shows that a volatility-managed momentum strategy has much smaller crash risk than original momentum and suggest that this is due to volatility clustering generating excess kurtosis by construction. They do not explicitly examine the link between volatility and crash risk of original momentum though. We present results for such an exercise in the last column of Table 5.

Returns of months after a quarter with high volatility have an annualized volatility of 38.7%, versus 15.3% after low volatility. The difference between the two terciles has a t-statistic of 5.7. On top of that we find that skewness is smaller (more negative) after high volatility and kurtosis

¹⁰A 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.

is higher (differences between terciles for both of these results are significant at the 5% level). So high volatility forecasts positively both volatility and higher-order risk. Volatility clustering alone is not sufficient to explain this result.

It is tempting to interpret the informational content of volatility in momentum as indirect evidence of crowding in the strategy. The contrast with the results we obtain from direct measures of crowding with institutional trading data questions that interpretation though. The next section focuses on the relation between crowding and volatility.

4.4 Crowding and the volatility of momentum returns

In this section we present predictive regressions for realized volatility of momentum returns, computed from raw and risk adjusted (both static and dynamic) daily returns over the quarter. In all cases we include lag realized volatility as a control.

[Insert Table 6 near here]

Table 6 presents the results. First, note that lagged realized volatility has very strong predictive power for subsequent volatility, with t-statistics between 6.7 and 8.5 across all regressions. This confirms the results of Barroso and Santa-Clara (2015) who show that momentum has highly predictable volatility. Table 6 also shows that uncertainty about crowding, as proxied with $\hat{\sigma}_{\text{Crowd}}$, also predicts positively the subsequent volatility of momentum returns. The coefficients have the predicted positive sign for all proxies and 7 out of 18 show significance at the 5% level, with a couple more at the 10% level. This shows that crowding is a determinant of risk in momentum returns.

The observation that $\hat{\sigma}_{\text{Crowd}}$ predicts positively risk in the momentum strategy supports the earlier finding in Table 3 suggesting (weakly) that $\hat{\sigma}_{\text{Crowd}}$ predicts positively momentum returns. If momentum investors pull back from the strategy when $\hat{\sigma}_{\text{Crowd}}$ is high in the belief that $\hat{\sigma}_{\text{Crowd}}$ predicts risk, a positive relation in Table 3 follows. The evidence in Table 6 supports that belief. One final observation is that the forecasting power of $\hat{\sigma}_{\text{Crowd}}$ in Table 6 is strongest with the _1qtr measures. Presumably stability and persistence in the _4qtr measures – an advantage in the ΔCrowd_q and Crowd_{q-1} specifications – is a disadvantage when it comes to proxying uncertainty.

Table 6 provides a fairly robust inference that crowding predicts negatively volatility in momentum returns. The coefficient estimate on both the ΔCrowd_q and the Crowd_{q-1} specifications (applied to Cnt_measures) is statistically significant at the 5% level in 13 of 24 cases across Panels A and B, and is negative in all cases. One possible explanation for this is a reversal of causality, with expectations of risk in the momentum strategy affecting institutions' willingness to participate in the strategy (i.e., influencing the Cnt_measures). This would require that forward-looking institutions observe a wider information set than just lagged volatility in momentum returns, but that is surely plausible. We present evidence relating to this conjecture in the next section.

Finally, note that the explanatory power of crowding measures 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, it is estimated at a daily frequency whereas crowding measures are based on holdings data observed at a quarterly frequency. 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 insight beyond persistence in setting. We believe that this fact makes up for the shortcoming in predictive power.

4.5 Determinants of crowding

Barroso and Santa-Clara (2015) show that volatility strongly predicts negatively the performance of the momentum strategy. As such, rational momentum investors should reduce their exposure to the strategy whenever recent volatility is high. The previous section alludes to the possibility that this concern for risk indeed attenuates crowding. Here we explore the matter directly using lagged realized volatility as a predictor of crowding.

We also consider the lagged first moment of momentum returns, motivated by several studies. 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. If such periods also relate to crowding, particularly unanticipated crowding, the model implications may be involved. Piazzesi and Schneider (2009) also use survey data to study the presence of momentum investors in the US housing market. They find there is evidence of a time-varying subset of momentum investors that doubles in size towards the end of the boom in the housing market. We examine the

matter in equities.

Table 7 shows the results of regressions of crowding on one-year returns and one-year volatility for the momentum factor, computed using daily observations and lagged at the indicated horizon. Because 13F observations of institutional trading occur over a quarter, we lag the returns and volatility by a quarter. Also, to ensure predetermined values for the `_4qtr` measures, we add one-year returns and volatility lagged five quarters. We use the level of `Cnt_` and `Cap_` measures as dependent variables.

[Insert Table 7 near here]

We find that one-year returns predict positively crowding in momentum. The coefficients on lagged returns from Table 7 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. However, in the case of returns lagged five quarters, there is a reliable positive relation with `Cnt_` measures, but not `Cap_` measure. Results are more comparable at different horizons in the case of `_1qtr` measures. The evidence here confirms that lagged returns positively influence crowding.

A noteworthy pattern in the relation between lagged returns and crowding is the shift in strength at various horizons: `Cnt_` measures react stronger at the five-quarter lag whereas `Cap_` reacts strongest at the one quarter lag. This effect is not attributable to overlap concerns, as it is seen with `_1qtr` measures as well. This suggests that the intensity of the representative demand curve of existing momentum investors' responds relatively quickly to past returns, but that institutions' adoption and termination of the strategy occurs with a more delayed response. This makes sense as the choice of investment strategy is surely a weightier decision than the parameterization of an existing strategy.

Regarding one-year volatility, from Table 7 we also find predictability for the crowding variables. In the case of `_4qtr` measures the dependence of crowding on momentum-return volatility occurs only when the observation periods overlap, obfuscating the sequencing of events. Indeed, from Table 6 we have already seen that `Cnt_` measures of crowding predict negatively future risk in the momentum strategy. From Table 7, crowding also *reacts* negatively to past risk in the mo-

momentum strategy, but only at a relatively high frequency (one quarter). This is seen most clearly in the case of the $_1qtr$ measures where there is no overlap and there is a statistically reliable (at the 1% level) reduction in crowding. Combining inferences from the anticipation and reaction analyses (i.e., Tables 6 and 7), we conclude that risk is a primary determinant of crowding; that its impact is largely anticipatory; and that the response is fast—materially faster than the reaction to past momentum returns.

Interestingly, this observation applies to entry and exit to the strategy (Cnt_ measures) rather than to the intensity of trading, where the relation is not statistically reliable. Note that if an institution reverses its momentum holdings, the input to the indicator function $\mathbb{1}_{Buy_{i,q}}$ turns negative; the indicator turns to zero; and Cnt_ measures fall. The evidence in Table 7 suggests, therefore, that risk in the strategy doesn't merely reduce the intensity of momentum demands (as proxied with Cap_ measures), it induces wholesale exit from the strategy. Barroso and Santa-Clara (2015) suggest that momentum investors should scale their position according to recent volatility.¹¹ These results suggest that institutions, in aggregate, react in a rather more dramatic fashion.

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 particularly over-confident 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 controlling 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. The volatility results suggest that forward-looking rational momentum investors successfully time risk in the strategy. That may be partly causal as exodus from the strategy potentially generates self-fulfilling risk. On the other hand, chasing momentum returns is harder to rationalize since returns to the strategy do not show time-series autocorrelation, indeed recent high returns seem to increase the crash risk for the strategy (Chabot et al., 2014). However, if the response to past returns is more euphoric than rational, our evidence on the role of crowding in predicting subsequent momentum performance offers a possible explanation for the Chabot et al. result.

¹¹They examine the benefits of risk management with different windows to compute volatility, all shorter than one year.

4.6 Capital versus crowd

We have argued that Cnt_t measures track the number of momentum investors and Cap_t measures track the intensity of their trade. In our final analysis we consider the two dimensions of crowding in a multivariate setting. While the results largely confirm the preceding analyses, this analysis serves to highlight the differing nature of the two sets of proxies. As the intent is to summarize, we consider all three moments of momentum collectively in a single table: returns; left-tail events; and volatility. We only present one case, that of residual returns from the dynamic Fama-French model using the Cnt_t (rather than CntP1_t) identification of the number of momentum-trading institutions.

[Insert Table 8 near here]

Results for returns are presented in the first two columns of Table 8, for the _1qtr and _4qtr measure, respectively. In comparing the columns, it is clear that the more stable and persistent _4qtr measure does a better job of predicting momentum returns. This highlights the importance of identifying momentum-trading institutions, rather than just aggregating institutional trading in momentum stocks, in studying the effects of crowding. From the _4qtr column it is also clear that the *number* of momentum investors, rather than the *intensity* of their trading, is most relevant to the crowding story developed in Stein (2009). This is further supported by the strength of the relation for ΔCrowd_q , which we argue proxies for unanticipated crowding. Note also that the positive relation of $\hat{\sigma}_{\text{Crowd}}$ in predicting momentum returns is here statistically significant at the 1% level.

In contrast with the negative predictive relation of Cnt_t measures for returns, the Cap_t measure is significantly positively related to future returns using the Crowd_{q-1} specification. Because the Table 8 regressions estimate marginal effects and the marginal difference in the Cap_t measure is the intensity of the representative momentum trade, this evidence suggests that the important dimension of a feedback effect from crowding is the number of institutions, rather than the manner in which they trade. The positive coefficient estimate supports the view that Crowd_{q-1} , using the $\text{Cap}_{\text{_4qtr}}$ measure, captures optimal intensity in the momentum trade. According to the model, that optimum is high when (or rather anticipating that) the momentum on the stocks in question is also high.

The left-tail results in Table 8 confirm the results in Table 4: unanticipated crowding increases the probability of moderate left-tails events consistent with a shift in mean effect, and the effect is attributable to uncertainty in the number of momentum-trading institutions. Volatility results likewise confirm the earlier analysis (Table 6).

4.7 Robustness: comparison with the momentum gap

Barroso and Santa-Clara (2015) find that momentum has a strongly time-varying volatility that is essentially specific to the strategy. Barroso and Maio (2016) find a similar result with the betting-against-beta strategy of Frazzini and Pedersen (2014). One plausible interpretation of these results is that time-varying volatility could be gauging investor crowding in each strategy. Recent closely related work proposes return-based estimates of arbitrage activity in momentum and reinforces that interpretation establishing a link between those estimates and institutional trading in the strategy.

Lou and Polk (2013) propose comomentum, a measure of abnormal co-movement of stocks in the momentum portfolio. They show it is positively related to aggregate institutional ownership of the winners portfolio. Huang (2015) compute the momentum gap, the difference in cumulative return of the winners versus the losers in the formation period of the portfolio. This is defined as the 75th percentile of the return of stocks in the formation period minus the 25th percentile.¹² Huang (2015) shows the momentum gap is related to the difference of short interest by institutions between the losers and the winners portfolios, a measure of trading in momentum ($\Delta\text{Mom Inst}$). Both these studies find that indirect proxies of crowding in momentum predict crash risk in the strategy. Our empirical results with trading-based measures of crowding in momentum do not support that interpretation and this motivates a comparison.

We focus on the relation between institutional trading and momentum gap. We choose this variable for comparison due to its simplicity and because it is a very strong predictor of risk and return for momentum.

[Insert Table 9 near here]

Table 9 shows the conditional moments of momentum. In each column, we rank all months in our sample in terciles according to the value of a sorting variable in the last available quarter. In

¹²See details of the construction of this measure in Huang (2015).

the first column the ranking variable is the momentum gap. A high momentum gap, interpreted as an indication that the strategy is crowded, forecasts higher volatility, lower skewness, and higher excess kurtosis, all with statistical significance at the 1% level. These results largely confirm Huang (2015).

Next, we assert that the momentum gap can be a reflection of crowded trades or it could be instead determined by fundamental shocks driving prices far apart in the formation period of momentum (or some other reason unrelated to crowding in momentum per se). To examine the extent to which other reasons besides crowding could be plausible candidates for the strong informational content of momentum gap, we orthogonalise it with respect to each estimate of crowding obtained from institutional holdings or trading (we call this $\text{gap}^\perp$).

Columns 2 to 6 show $\text{gap}^\perp$ preserves substantial predictive power. The difference in volatility between months with high versus low $\text{gap}^\perp$ is close to the one obtained sorting with momentum gap alone. High $\text{gap}^\perp$ predicts lower skewness in all specifications with crowding variables and the reduction is statistically significant at the 5% level for $\Delta\text{Mom Inst}$, Win Inst , and Cap_- . High $\text{gap}^\perp$ also predicts higher kurtosis for the strategy. Four of the five specifications show statistical significant differences in kurtosis at least at the 5% level.¹³ Taken together 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 necessarily come from crowding in the strategy.

One pertinent question is the relation of momentum gap with the volatility measure we use as a control in many of our empirical exercises. We note that the momentum gap is defined as the inter-quartile range of the return distribution for stocks in a given formation period. This is naturally a measure of (cross-sectional) dispersion in returns that should by construction be closely related to volatility. We indeed find that momentum gap has a high correlation with the volatility of momentum of 0.73 in our sample period. In the last column of table 9 we show the results using momentum gap orthogonalised with respect to realized volatility in momentum. We find that $\text{gap}^\perp$ is still a robust predictor of volatility. In months after high $\text{gap}^\perp$ volatility is 35.6% versus 21.0% after low $\text{gap}^\perp$, the t-statistic for the difference is 3.2, significant at the 1% level. This suggests momentum gap captures information not necessarily contained in volatility. The evidence

¹³In unreported results we also examined momentum gap orthogonal to all variables simultaneously with similar results.

for higher order risk is weaker though. Months after high gap⁺ have more negative skewness and higher kurtosis but none of the differences is statistically significant. Hence there is likely some overlap in the informational content of momentum gap and volatility for higher order risk. This is not very surprising given their high correlation.

Of course, one alternative interpretation for our results in this section is that, due to estimation error, the residual of momentum gap with respect to crowding could be still driven by unobserved crowding. It is conceivable that return-based measures of crowding can benefit from sampling at a higher frequency and as result capture arbitrage activity better. On the other hand, to the extent that their strong informational content can not be safely attributed to crowding, it is also plausible that measures such as the momentum gap, or volatility have predictive power for other reasons (e.g., fundamental shocks to stock prices).

5 Conclusion

We provide a model based on supply and demand considerations for the momentum factor to study how momentum moments arise due to the trading behavior of investors. Our model delivers intuitive predictions such as that momentum factor returns are decreasing in changes in momentum capital and increasing in the uncertainty about momentum capital, but also novel predictions regarding momentum's second and higher moments. We obtain that uncertainty about momentum capital predicts momentum return volatility and that large unanticipated changes in momentum capital may lead to crashes in momentum strategy returns for linear beliefs. We also show that crashes need not occur when momentum investors conjecture non-linear concave beliefs.

Using quarterly holdings of 13F institutions in the period from 1980 to 2015, we construct several proxies for momentum capital based on aggregate momentum trading to test the predictions of our model. We find that changes in our proxies negatively predict momentum strategy returns. We also find evidence that uncertainty about momentum capital predicts higher momentum volatility, and that our measures of momentum trading are positively related to momentum crashes. Our empirical findings are generally in line with the idea that momentum trading and the uncertainty thereof contribute to momentum's moments. At the same time, crowding - as measured by the trading of institutional investors - is unable to explain momentum's largest crashes.

A Derivations

A.1 Derivation of Eq. (8)

First notice that solving Eq. (7) is equivalent to solving each of the following

$$\max \quad \frac{K_0^{1-\gamma}}{1-\gamma} \cdot E \left[e^{(1-\gamma)g} \right] \quad \Leftrightarrow \quad \min \quad \log E \left[e^{(1-\gamma)g} \right],$$

(presuming $\gamma > 1$). Second, to solve for the fraction of wealth invested in the risky portfolio, we follow Campbell and Viceira (2002, appendix) and approximate the objective function using a second-order Taylor expansion of g around $r_p - f = 0$:

$$\begin{aligned} g &\cong f + \log(1) + \frac{\varsigma e^0}{1 + \varsigma(e^0 - 1)} (r_p - f) + \frac{1}{2} \frac{\varsigma \left[e^0 (1 + \varsigma(e^0 - 1)) - \varsigma e^{2 \cdot 0} \right]}{(1 + \varsigma(e^0 - 1))^2} (r_p - f)^2, \\ &\cong f + \varsigma (r_p - f) + \frac{1}{2} (\varsigma - \varsigma^2) \sigma^2, \end{aligned} \quad (32)$$

where $(r_p - f)^2$ is replaced with its expected value σ^2 . Using (32), we can then rewrite the maximization problem to

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

which has solution

$$\varsigma = \frac{\mu - f}{\gamma \sigma^2}. \quad (33)$$

We now determine μ and σ^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 $\mathbf{w}' = \begin{bmatrix} 1 & w_m \end{bmatrix}$, 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 opti-

mization

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

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

$$\boldsymbol{\mu} = \begin{bmatrix} r - f \\ Em_{Rnk \rightarrow 1} \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_\chi^2 & 0 \\ 0 & \sigma_\delta^2 \end{bmatrix},$$

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

$$w_m = \frac{Em_{Rnk \rightarrow 1} / \sigma_\delta^2}{(r - f) / \sigma_\chi^2}. \quad (34)$$

Using (34), the parameters of the optimal risky portfolio are

$$\mu_p - f = \begin{bmatrix} r - f & E[m_{Rnk \rightarrow 1}] \end{bmatrix} \begin{bmatrix} 1 \\ w_m \end{bmatrix} = r - f + w_m E[m_{Rnk \rightarrow 1}], \quad (35)$$

and

$$\sigma_p^2 = \mathbf{w}'\Sigma\mathbf{w} = \sigma_\chi^2 + w_m^2 \sigma_\delta^2 = \frac{\sigma_\chi^2}{r - f} (r - f + w_m E[m_{Rnk \rightarrow 1}]). \quad (36)$$

Taking the ratio gives

$$\varsigma = \frac{r - f}{\gamma \sigma_\chi^2}. \quad (37)$$

Combining (34) and (37),

$$Demand = w_m \cdot \varsigma \cdot K_0 = \frac{Em_{Rnk \rightarrow 1}}{\gamma \sigma_\delta^2} K_0. \quad (38)$$

B Distributions for the simulations

The realized differential dividend, d_1 , in each simulated market clearing is presumed to follow a log normal distribution and the realized momentum capital, $\tilde{k}_M$, is presumed to follow a beta distribution. The former draws from the evidence in Andersen et al. (2001) and references therein.

Because $\tilde{d}_1$ sets the scale of the market setting, we presume the simplest case for $\tilde{d}_1$ where the exponent in the lognormal is a standard normal distribution. This implies a mean of about 1.6. The beta distribution assumption stems from the observation that $\tilde{k}_m$ is necessarily between 0 and 1. We parameterize this with the following logic. Beta has two parameters, a and b , which can be expressed in terms of the expected value and variance of the distribution. The expected value is a parameter of the market setting presumed to be 1/3 in Figure 1. For the variance, we first solve for the a and b that give maximal variance subject to (1) hit the target $E\tilde{k}_m$ and (2) keep both a and b above 2. Below 2, probability mass clumps at the boundary and we want a single internal local maximum. If each of a and b are greater than 2 then the probability of a boundary outcome (0 or 1) is zero; we insist on that. Once the maximum variance is identified, we use a value for a and b that generates variance half of that maximum.

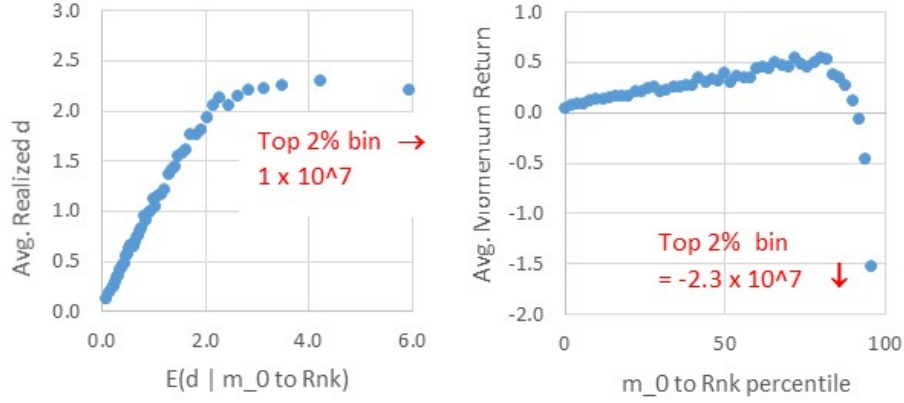
References

- D. Abreu and M. K. Brunnermeier. Synchronization risk and delayed arbitrage. *Journal of Financial Economics*, 66:341 – 360, 2002.
- 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.
- C. S. Asness, A. Frazzini, R. Israel, and T. J. Moskowitz. Fact, fiction and momentum investing. *Journal of Portfolio Management*, Fall, 2014.
- D. Avramov, S. Cheng, and A. Hameed. Time-varying liquidity and momentum profits. *Journal of Financial and Quantitative Analysis*, 51(6):1897–1923, 2016.

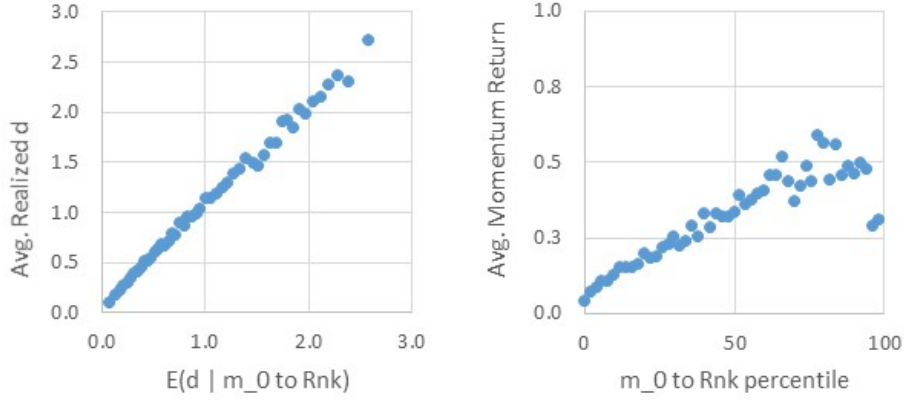
- N. Barberis, A. Shleifer, and R. Vishny. A model of investors sentiment. *Journal of Financial Economics*, 49:307–343, 1998.
- P. Barroso. The bottom-up beta of momentum. Working Paper SSRN, 2014.
- P. Barroso and P. F. Maio. Managing the risk of the 'betting-against-beta' anomaly: Does it pay to bet against beta? Working Paper SSRN, 2016.
- 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.
- 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.
- 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.

- A. Frazzini and L. H. Pedersen. Betting against beta. *Journal of Financial Economics*, 111(1): 1–25, 2014.
- M. Grinblatt, S. Titman, and R. Wermers. Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior. *The American Economic Review*, 85(5):1088–1105, 1995.
- K. Grobys, J. Ruotsalainen, and J. J. Äijö. Risk-managed industry momentum and momentum crashes. Working Paper SSRN, 2016.
- 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. Evidence of predictable behavior of security returns. *The Journal of Finance*, 45(3): 881–898, 1990.
- N. Jegadeesh and S. Titman. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, 48(1):65–91, 1993.

- P. Kondor and A. Zawadowski. Learning in crowded markets. 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.
- P. Milgrom and N. Stokey. Information, trade and common knowledge. *Journal of Economic Theory*, 26(1):17–27, 1982.
- R. Novy-Marx. Fundamentally, momentum is fundamental momentum. Working Paper 20984, National Bureau of Economic Research, 2015.
- 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.
- K. G. Rouwenhorst. International momentum strategies. *The Journal of Finance*, 53(1):267–284, 1998.
- 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. 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.



(a) Beliefs linear in $m_{0 \rightarrow Rnk}$



(b) Concave beliefs such that $E(\tilde{d}_1 | m_{0 \rightarrow Rnk}) \approx \text{numerical average } \tilde{d}_1 | m_{0 \rightarrow Rnk}$

Figure 1: Simulations of Beliefs and Momentum Returns with $E\tilde{k}_M = 0.333$, $k_I = 0.333$

The simulations use 5,000 independent random draws of $\{k_M, d_1\}$ where k_M is momentum capital and d_1 is the signal of differential fundamental value for winners minus losers. The market clearing ranking period return $m_{0 \rightarrow Rnk}$ is solved for each $\{k_M, d_1\}$ pair by iteration using a linear (Panel (a)) or isoelastic (Panel (b)) specification of beliefs. The triples $\{d_1, E[\tilde{d}_1 | m_{0 \rightarrow Rnk}], m_{Rnk \rightarrow 1}\}$ are then sorted by $E[\tilde{d}_1 | m_{0 \rightarrow Rnk}]$ into 50 equally populated bins and the average realized d_1 (first graph) and average realized momentum return (second graph) is plotted. The numeraire for momentum returns is the expected value of d_1 . Expectations are linear in Panel A (with the restriction that ranking period returns be $< 10^7$ times the expectation of d). In the case of concave expectations (Panel (b)) an isoelastic function is used with concavity parameters chosen to equate the conditional average d to the corresponding $E[\tilde{d}_1 | m_{0 \rightarrow Rnk}]$ within each bin. Under conditionally rational expectations, beliefs generate an identify mapping in the first plot. In Panel (a), momentum returns have mean: -4.1×10^5 ; stdev: 5.9×10^6 , and skew -15.9 , and in Panel (b) momentum returns have mean 0.32, stdev: 0.40, and skew -0.40.

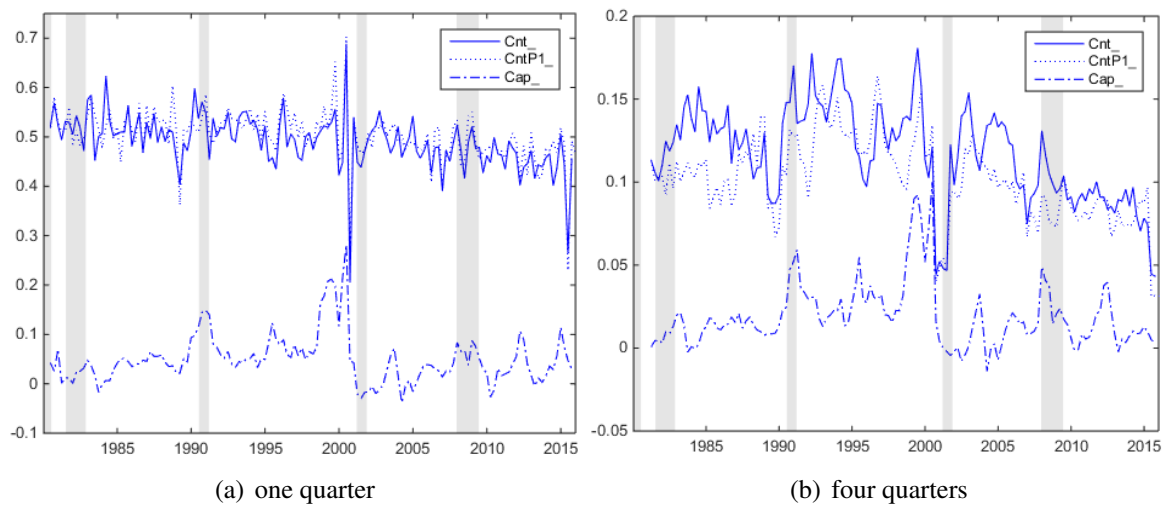


Figure 2: Measures of crowding

Panel (a) and (b) report the _1qrt and _4qrt 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: Descriptive statistics of momentum returns

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 Cnt_1qtr, Cnt_1qtrP1, and Cap_1qtr; likewise the _4qtr extensions (as defined in Section 5.1). $\hat{\sigma}_{\text{Crowd}}$ is the estimate of volatility from a GARCH(1,1) model. Panel C contains factor exposures of quarterly momentum returns on the Fama-French three factor model and a dynamic extension with 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			momentum investors			not momentum investors						
	mean	med.	stdev.	mean	med.	stdev.	mean	med.	stdev.				
#Qtrs of data	34.3	24.0	32.1	34.3	24.0	32.1	34.3	24.0	32.1				
#Qtrs missing	3.6	0.0	9.5	3.6	0.0	9.5	3.6	0.0	9.5				
#Stocks held	143.2	62.9	275.5	213	90.1	372.5	123.1	55.2	236.6				
Assets mgd.	15.2	2.0	102.4	24.6	2.1	161.6	12.3	1.9	76.6				
Turnover	0.21	0.16	0.17	0.24	0.18	0.18	0.21	0.15	0.16				
#Institutions	6,360			1,414			5,059						
Panel B. Crowding variables													
	Cnt_			CntP1_			Cap_						
	mean	med.	stdev.	mean	med.	stdev.	mean	med.	stdev.				
			ar(1)			ar(1)			ar(1)				
_1qtr measure:													
ΔCrowd	0.000	0.004	0.075	-0.579	0.000	0.002	0.077	-0.563	0.000	0.001	0.034	-0.092	
Crowd	0.492	0.495	0.055	0.083	0.504	0.509	0.054	0.001	0.050	0.042	0.051	0.782	
σ̂Crowd	0.052	0.049	0.010	0.367	0.053	0.052	0.005	0.523	0.029	0.023	0.018	0.703	
_4qtr measure:													
ΔCrowd	-0.001	0.000	0.017	-0.153	-0.001	0.000	0.017	-0.181	0.000	0.001	0.012	0.029	
Crowd	0.117	0.120	0.029	0.818	0.102	0.101	0.026	0.770	0.019	0.014	0.020	0.810	
σ̂Crowd	0.017	0.016	0.003	0.881	0.018	0.017	0.003	0.756	0.011	0.008	0.007	0.666	
Panel C. Returns													
	FF3					dynamic FF3							
	alpha	mkt	SMB	HML	R ²	alpha	mkt	SMB	HML	Dmkt	DSMB	DHML	R ²
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 2: Transition frequencies

The table presents the probability of transitioning from the event in the row heading to that in the column heading at the indicated time (q indexes quarters), conditional on the later period not containing a missing observation. Panel A tabulates the transition at the level of individual institutions in terms of the momentum classification used to construct the indicated crowding variable ('1' stands for being a momentum trader and '0' for not being a momentum trader). Panel B tabulates stocks' membership in the winner, loser, or middle deciles of the momentum ranking. 'All' refers to the unconditional probability of classification.

Panel A. Institutions' trading								
Indicator for > 0:		q+1		q+4		All		
		1	0	1	0			
Cnt_/Cap_1qrt	1	0.56	0.44	0.55	0.45	0.48		
	0	0.41	0.59	0.41	0.59	0.52		
Cnt_/Cap_4qrt	1	0.69	0.31	0.29	0.71	0.11		
	0	0.04	0.96	0.09	0.91	0.89		
Cnt_1qrtP1	1	0.55	0.45	0.54	0.46	0.50		
	0	0.44	0.56	0.45	0.55	0.50		
Cnt_4qrtP1	1	0.64	0.36	0.19	0.81	0.10		
	0	0.04	0.96	0.08	0.92	0.90		
Panel B. Stock returns								
		q+1			q+4			All
		Win.	mid	Los.	Win.	mid	Los.	
Winner		0.55	0.42	0.02	0.16	0.60	0.23	0.13
mid		0.08	0.83	0.09	0.12	0.74	0.14	0.68
Loser		0.02	0.34	0.64	0.17	0.52	0.31	0.18

Table 3: 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									
<i>Model:</i> <i>Measure:</i>	cumulative returns			FF3			dynamic FF3		
	Cnt_	CntPI_	Cap_	Cnt_	CntPI_	Cap_	Cnt_	CntPI_	Cap_
ΔCrowd_q	-0.57 (-2.8)	-0.59 (-2.8)	-0.27 (-1.0)	-0.57 (-3.2)	-0.59 (-3.6)	-0.32 (-1.3)	-0.55 (-2.8)	-0.62 (-3.0)	-0.28 (-0.9)
Crowd_{q-1}	-0.17 (-1.1)	-0.04 (-0.3)	0.40 (1.8)	-0.14 (-1.2)	0.10 (0.7)	0.41 (2.1)	-0.11 (-1.0)	0.09 (0.7)	0.48 (2.5)
$\hat{\sigma}_{\text{Crowd}}$	2.39 (1.4)	2.76 (1.6)	-0.20 (-0.3)	2.53 (1.7)	3.26 (2.4)	-0.04 (-0.1)	1.81 (1.3)	2.42 (1.6)	-0.36 (-0.6)
Realized vol. of Mom rets.	-0.36 (-1.9)	-0.32 (-1.8)	-0.31 (-1.6)	-0.36 (-2.6)	-0.29 (-2.1)	-0.32 (-2.4)	-0.30 (-2.6)	-0.23 (-1.9)	-0.26 (-2.3)
Adj-rsquare	13.5%	14.4%	10.5%	26.6%	28.2%	24.4%	35.3%	37.6%	35.0%
Panel B. Crowding measures constructed using one-quarter trading histories									
<i>Model:</i> <i>Measure:</i>	cumulative returns			FF3			dynamic FF3		
	Cnt_	CntPI_	Cap_	Cnt_	CntPI_	Cap_	Cnt_	CntPI_	Cap_
ΔCrowd_q	-0.07 (-1.0)	-0.06 (-0.7)	-0.06 (-0.7)	-0.07 (-1.1)	-0.07 (-0.9)	-0.11 (-1.3)	-0.12 (-1.3)	-0.09 (-1.1)	-0.13 (-1.2)
Crowd_{q-1}	-0.04 (-0.4)	-0.04 (-0.4)	0.14 (1.4)	-0.03 (-0.4)	-0.01 (-0.1)	0.12 (1.5)	0.02 (0.2)	0.00 (-0.0)	0.17 (2.3)
$\hat{\sigma}_{\text{Crowd}}$	0.15 (0.5)	0.68 (0.9)	0.05 (0.2)	0.26 (1.0)	1.01 (1.7)	0.14 (0.7)	0.29 (0.7)	0.89 (1.1)	-0.02 (-0.1)
Realized vol. of Mom rets.	-0.29 (-1.6)	-0.30 (-1.6)	-0.32 (-1.6)	-0.30 (-2.1)	-0.32 (-2.1)	-0.34 (-2.4)	-0.23 (-1.8)	-0.26 (-2.0)	-0.28 (-2.4)
Adj-rsquare	6.9%	7.2%	9.9%	20.2%	21.0%	23.8%	33.8%	32.4%	35.9%

Table 4: Crowding and the left-tail of momentum returns

Each column represents a Probit regression with an indicator for next-quarter momentum returns in the bottom 10% of the full-sample (1981 - 2015) distribution (Panel A) or bottom 5% (Panel B). Each panel presents two sets of dependent variables: (1) the 3 month return on the momentum portfolio; and (2) its residual on the dynamic Fama and French model with the three factors interacted with dummies for positive past annual factor returns. Each set considers the three indicated crowding measures. In all cases the crowding measure is constructed using a four-quarter trading history. 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). 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 in brackets.

<i>Dependent variable: Crowding measure:</i>	Panel A. Predicting the 10% left tail						Panel B. Predicting the 5% left tail					
	cumulative returns			dynamic FF3 residuals			cumulative returns			dynamic FF3 residuals		
	Cnt_	CntPI_	Cap_	Cnt_	CntPI_	Cap_	Cnt_	CntPI_	Cap_	Cnt_	CntPI_	Cap_
ΔCrowd_q	25.6 (2.5) [0.08]	26.9 (2.0) [0.28]	11.7 (0.8) [0.87]	25.1 (2.4) [0.69]	51.3 (3.2) [0.04]	16.4 (1.0) [0.79]	32.1 (2.2) [0.31]	59.9 (2.6) [0.06]	52.9 (2.1) [0.07]	1.6 (0.1) [0.24]	28.6 (1.7) [0.65]	34.9 (1.4) [0.32]
Crowd_{q-1}	11.6 (1.8) [0.21]	10.0 (1.3) [0.18]	-7.0 (-0.8) [0.76]	9.6 (1.5) [0.26]	6.5 (0.7) [0.11]	-7.8 (-0.9) [0.42]	20.2 (1.8) [0.20]	11.7 (0.9) [0.58]	-18.2 (-1.4) [0.32]	7.9 (1.0) [0.62]	10.2 (0.9) [0.27]	-21.6 (-1.5) [0.58]
$\hat{\sigma}_{\text{Crowd}}$	20.3 (0.4) [0.04]	14.8 (0.3) [0.02]	25.1 (1.2) [0.96]	54.4 (1.2) [0.05]	16.5 (0.3) [0.13]	30.4 (1.4) [0.55]	80.3 (1.2) [0.02]	-15.0 (-0.2) [0.67]	3.3 (0.1) [0.65]	76.0 (1.5) [0.02]	67.8 (1.1) [0.20]	31.4 (1.6) [0.00]
Realized vol. of Mom rets.	14.0 (4.3) [0.00]	13.1 (4.0) [0.00]	10.6 (3.0) [0.01]	12.2 (3.8) [0.00]	11.7 (3.2) [0.00]	10.5 (3.4) [0.00]	17.2 (3.7) [0.00]	14.9 (3.2) [0.00]	17.1 (3.4) [0.00]	10.6 (3.0) [0.00]	10.2 (2.8) [0.00]	11.4 (3.0) [0.00]

Table 5: 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 (' Δ Crowd') or volatility in the last available quarter for each month. The G1 stands for the bottom tercile, G2 for the second tercile and G3 for the top tercile. The values in parenthesis are t-statistics for the difference between G3 and G1 obtained with the delta method.

	Δ Crowd			Crowd			Realized vol. of Mom rets.
	Cnt_	CntP1_	Cap_	Cnt_	CntP1_	Cap_	
Volatility							
G1	27.3	25.9	30.0	33.5	33.0	21.8	15.3
G2	27.4	26.9	21.0	19.4	21.3	30.6	17.3
G3	23.2	25.2	26.5	23.2	22.1	24.8	38.7
	(-1.0)	(-0.2)	(-0.7)	(-2.0)	(-2.1)	(1.1)	(5.7)
Skewness							
G1	-1.83	-1.37	-1.34	-2.05	-2.13	-0.48	-0.34
G2	-1.94	-2.75	-0.66	0.34	-0.07	-2.54	-0.31
G3	-0.22	-0.17	-1.97	-0.18	0.16	0.07	-1.18
	(1.8)	(1.3)	(-0.7)	(3.3)	(4.1)	(1.1)	(-2.0)
Kurtosis							
G1	13.97	9.99	10.30	10.44	10.65	4.75	4.03
G2	11.79	16.47	5.81	4.16	4.46	13.65	4.06
G3	4.48	5.59	12.68	5.35	5.80	5.22	6.52
	(-4.3)	(-1.3)	(0.7)	(-2.6)	(-2.4)	(0.4)	(2.1)

Table 6: 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; 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/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									
<i>Dependent variable:</i>	vol of returns			vol of FF3 residuals			vol of dynamic FF3 residuals		
<i>Crowding measure:</i>	Cnt_	CntP1_	Cap_	Cnt_	CntP1_	Cap_	Cnt_	CntP1_	Cap_
ΔCrowd_q	-0.19 (-1.1)	-0.45 (-3.2)	-0.06 (-0.2)	-0.21 (-1.2)	-0.44 (-3.1)	-0.12 (-0.4)	-0.19 (-1.0)	-0.38 (-2.4)	-0.17 (-0.6)
Crowd_{q-1}	-0.14 (-2.3)	-0.17 (-1.9)	0.05 (0.4)	-0.16 (-2.5)	-0.17 (-2.0)	0.00 (0.0)	-0.12 (-2.2)	-0.15 (-2.4)	0.05 (0.5)
$\hat{\sigma}_{\text{Crowd}}$	0.65 (0.8)	1.02 (1.2)	0.68 (1.9)	0.96 (1.1)	1.34 (1.5)	0.86 (2.3)	0.45 (0.8)	0.44 (0.7)	0.51 (2.0)
Realized vol. of Mom rets.	0.76 (7.4)	0.77 (8.1)	0.76 (8.1)	0.72 (7.1)	0.73 (7.9)	0.72 (8.2)	0.73 (6.7)	0.74 (7.5)	0.72 (7.6)
Adj-rsquare	63.7%	66.2%	64.2%	60.7%	63.4%	61.7%	59.7%	62.5%	60.7%
Panel B. Crowding measures constructed using one-quarter trading histories									
<i>Dependent variable:</i>	vol of returns			vol of FF3 residuals			vol of dynamic FF3 residuals		
<i>Crowding measure:</i>	Cnt_	CntP1_	Cap_	Cnt_	CntP1_	Cap_	Cnt_	CntP1_	Cap_
ΔCrowd_q	-0.11 (-2.8)	-0.08 (-1.8)	-0.02 (-0.1)	-0.10 (-2.9)	-0.08 (-2.0)	-0.04 (-0.4)	-0.10 (-3.4)	-0.08 (-2.3)	-0.06 (-0.7)
Crowd_{q-1}	-0.07 (-1.5)	-0.02 (-0.3)	0.04 (1.3)	-0.06 (-1.6)	-0.01 (-0.2)	0.03 (0.8)	-0.05 (-1.6)	-0.03 (-0.6)	0.04 (1.5)
$\hat{\sigma}_{\text{Crowd}}$	0.50 (2.3)	1.38 (2.6)	0.10 (0.6)	0.60 (2.7)	1.55 (2.9)	0.16 (0.9)	0.39 (1.9)	1.02 (2.1)	0.08 (0.7)
Realized vol. of Mom rets.	0.78 (8.3)	0.77 (7.8)	0.78 (7.3)	0.75 (8.5)	0.74 (7.9)	0.74 (7.1)	0.75 (7.6)	0.74 (7.1)	0.75 (6.9)
Adj-rsquare	66.0%	66.0%	63.2%	63.5%	63.6%	60.0%	63.5%	62.8%	60.0%

Table 7: 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.

<i>Crowding horizon:</i> <i>Crowding measure:</i>	Panel A			Panel B		
	4qtr			1qtr		
	Cnt_	CntPI_	Cap_	Cnt_	CntPI_	Cap_
1yr return _{q-1}	0.23 (1.2)	0.06 (0.5)	0.27 (2.3)	0.47 (2.0)	0.40 (2.2)	0.73 (2.3)
1yr return _{q-5}	0.52 (2.9)	0.48 (3.7)	0.16 (1.4)	0.66 (2.6)	0.68 (3.3)	0.31 (1.0)
1yr volatility _{q-1}	-0.41 (-3.3)	-0.31 (-2.9)	-0.05 (-0.9)	-0.46 (-2.9)	-0.37 (-3.0)	-0.10 (-0.7)
1yr volatility _{q-5}	0.11 (0.9)	0.04 (0.4)	-0.07 (-0.9)	0.18 (1.3)	0.07 (0.5)	-0.24 (-1.2)
Adj-rsquare	20.0%	17.5%	16.3%	6.5%	6.7%	18.9%

Table 8: 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. 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 3 using Cnt_ and Cap_ and DFF3 (but estimated jointly). Likewise, the regressions in the columns labelled 'left-tail' and 'Volatility' correspond to jointly estimated versions of Tables 4 and 5, respectively, for the case of Cnt_ and DFF3. 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.

<i>Dependent variable:</i>		Returns		10% left-tail		5% left-tail		Volatility	
<i>Crowding horizon:</i>		1qtr	4qtr	1qtr	4qtr	1qtr	4qtr	1qtr	4qtr
Crowd = _Cnt:									
ΔCrowd_q		-0.11 (-1.7)	-0.51 (-3.1)	6.7 (1.8) [0.33]	26.2 (2.4) [0.44]	10.3 (1.5) [0.38]	-1.0 (-0.1) [0.40]	-0.12 (-3.2)	-0.13 (-1.2)
Crowd_{q-1}		-0.04 (-0.5)	-0.32 (-2.7)	6.5 (1.3) [0.28]	14.5 (1.9) [0.63]	16.7 (2.2) [0.14]	19.9 (1.8) [0.96]	-0.07 (-2.0)	-0.15 (-2.7)
$\hat{\sigma}_{\text{Crowd}}$		-0.03 (-0.1)	3.03 (2.2)	-2.9 (-0.1) [0.94]	7.8 (0.1) [0.19]	43.1 (1.4) [0.60]	-10.2 (-0.1) [0.26]	0.39 (2.3)	0.24 (0.4)
Crowd = _Cap:									
ΔCrowd_q		-0.03 (-0.3)	0.07 (0.3)	3.7 (0.6) [0.73]	9.5 (0.6) [0.75]	21.6 (2.0) [0.07]	24.9 (0.9) [0.29]	0.07 (0.9)	-0.07 (-0.3)
Crowd_{q-1}		0.16 (2.2)	0.74 (3.7)	-3.9 (-1.1) [0.65]	-14.5 (-1.3) [0.46]	-12.5 (-1.7) [0.38]	-29.3 (-1.9) [0.83]	0.05 (1.9)	0.14 (1.6)
$\hat{\sigma}_{\text{Crowd}}$		0.06 (0.2)	-0.96 (-1.4)	10.3 (1.1) [0.42]	36.5 (1.5) [0.92]	-1.2 (-0.1) [0.49]	52.8 (1.9) [0.09]	0.02 (0.2)	0.41 (1.7)
Control:									
Realized vol. of Mom rets.		-0.27 (-2.2)	-0.36 (-3.3)	11.1 (3.3) [0.00]	14.1 (3.6) [0.01]	16.6 (3.4) [0.00]	15.7 (3.2) [0.06]	0.75 (6.9)	0.68 (6.2)

Table 9: 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 ' Δ Mom Inst' which is the percentage difference in aggregate short interest between past losers and winners (see, Huang, 2015), and 'Win Inst' which is the aggregate institutional ownership of the winner decile (see, Lou and Polk, 2013). G1 stands for the bottom tercile, G2 for the second tercile and G3 for the top tercile. The values in parenthesis are t-statistics for the difference between G3 and G1 obtained with the delta method.

Mom Gap		orthogonal to					
		Δ Mom	Win	Crowd			Realized vol.
		Inst	Inst	Cnt_	CntP1_	Cap_	of Mom rets.
Volatility							
G1	12.8	12.2	13.5	14.3	14.5	12.9	21.0
G2	19.2	19.9	18.7	18.1	18.0	19.5	17.8
G3	38.6	38.5	38.6	38.6	38.6	38.4	35.6
	(6.4)	(6.7)	(6.2)	(6.0)	(5.9)	(6.7)	(3.2)
Skewness							
G1	-0.32	-0.37	-0.27	-0.40	-0.56	-0.41	-0.68
G2	-0.01	0.04	0.04	0.07	0.13	0.25	-0.22
G3	-1.30	-1.25	-1.30	-1.30	-1.31	-1.27	-1.51
	(-2.4)	(-2.2)	(-2.6)	(-1.9)	(-1.7)	(-2.1)	(-1.4)
Kurtosis							
G1	3.32	3.41	3.17	4.21	4.11	3.50	6.42
G2	3.84	3.66	4.19	4.07	4.04	4.60	4.83
G3	6.61	6.57	6.62	6.62	6.66	6.51	8.42
	(2.9)	(2.8)	(3.1)	(1.8)	(2.0)	(2.6)	(1.1)