Internet Appendix
"Conviction and volume: Measuring the information content of hedge fund trading"

C Maximum likelihood estimation

C.1 Building the likelihoood function

Let $\Theta = (\Theta_1; \Theta_2) = (\sigma_\epsilon^2, \sigma_\eta^2, \sigma_u^2, \pi;\ \beta_1, \lambda_1, \beta_2, \lambda_2, \phi)$ be the vector of parameters. Let $X = (x_{1,s,t}, x_{2,s,t}, r_{1,s,t}, r_{2,s,t})$ be the vector of observables, with $r_2 = p_2 - p_1$ and $r_1 = p_1$. Let $\mathbf{1}(x)$ be an indicator variable equal to 1 if the event x has occured, and 0 otherwise. Let $g(x)$ be the standard normal PDF and $G(x)$ the standard normal CDF.

Solve for the probability of X given Θ. s indexes stocks, t indexes information episodes (rather than quarters), and the subscript of 1 (subscript of 2) denotes the first (second) quarter in each episode. The likelihood function for observing X given Θ, with x_1 and x_2 censored below at x_c and information publicly released after quarter 1 (as opposed to after quarter 2) if $x_1 > x_c$ and $x_2 \leq x_c$, is:

$$likelihood(X|\theta) = \prod_{t=1}^{T}\prod_{s=1}^{S} \mathbf{1}(x_{1,s,t} > x_c\, and\, x_{2,s,t} \leq x_c) * Pr(X_1 = x_{1,s,t}\, and\, R_1 = r_{1,s,t}) +$$
$$\mathbf{1}(x_{1,s,t} > x_c\, and\, x_{2,s,t} > x_c) * Pr(X_1 = x_{1,s,t}\, and\, R_1 = r_{1,s,t}\, and\, X_2 = x_{2,s,t}\, and\, R_2 = r_{2,s,t})$$
$$+ \mathbf{1}(x_{1,s,t} \leq x_c) * Pr(X_1 \leq x_c) \quad (11)$$

Now solve and plug in for the vector of unobserved random variables ($i_{s,t}$,$u_{1,s,t}$, $u_{2,s,t}$) as a function of X and Θ. Combined with taking the log of the likelihood function, this produces:

$$log(likelihood(X|\Theta)) = \sum_{t=1}^{T}\sum_{s=1}^{S} \mathbf{1}(x_{1,s,t} > x_c\, and\, x_{2,s,t} \leq x_c) * \Big\{ log(g(\frac{r_{1,s,t} - \lambda_1 x_{1,s,t}}{\lambda_1 \sigma_u}))$$
$$+ log(g(\frac{x_{1,s,t}}{\beta_1 \phi \sqrt{\sigma_\epsilon^2 + \sigma_\eta^2}})) \Big\} + \mathbf{1}(x_{1,s,t} > x_c\, and\, x_{2,s,t} > x_c) * \Big\{ log(g(\frac{r_{1,s,t} - \lambda_1 x_{1,s,t}}{\lambda_1 \sigma_u})) + log(g(\frac{r_{2,s,t} - \lambda_2 x_{2,s,t}}{\lambda_2 \sigma_u}))$$
$$+ log(g(\frac{x_{1,s,t}}{2\beta_1 \phi \sqrt{\sigma_\epsilon^2 + \sigma_\eta^2}} + \frac{x_{2,s,t} + \beta_2 r_{1,s,t}}{2\beta_2 \phi \sqrt{\sigma_\epsilon^2 + \sigma_\eta^2}})) \Big\} + \mathbf{1}(x_{1,s,t} \leq x_c) * G(\frac{x_c}{\beta_1 \phi \sqrt{\sigma_\epsilon^2 + \sigma_\eta^2}}) \quad (12)$$

The model is also subject to constraints, equations (6)-(10), which implicitly define Θ_2 by

$f_1(\Theta_1;\,\Theta_2) = 0$. Equation (10) is trivial. Equations (6)-(9) can be reduced to single implicit equation for $f_2(\Theta_1; \beta_1) = 0$:

$$a_3 * (2\sqrt{a_1 a_2} - \pi\sigma_u\sqrt{a_3}) = a_2 * (\sqrt{a_1 a_2} - \pi\sigma_u\sqrt{a_3}) \tag{13}$$

with $a_1 = \sigma_u^2 + \beta_1^2\phi^2\sigma_\eta^2$, $a_2 = \sigma_u^2 + \beta_1^2\phi\sigma_\epsilon^2$, and $a_3 = \beta_1^2\phi\sigma_\epsilon^2$. Thus given Θ_1, one can numerically solve for β_1. Once one has solved for β_1, one can explicitly obtain $\lambda_1 = f_3(\beta_1)$, $\lambda_2 = f_4(\beta_1, \lambda_1)$, and $\beta_2 = f_5(\lambda_2)$ in turn.

A maximum likelihood approach is implemented by maximizing equation (12) over Θ_1, subject to the constraints on Θ_2 (equation (13)), given data X.

I use maximum likelihood because it allows me to both utilize the Kyle model's explicit structure on error terms (i.e., u_1 and u_2 are normal and the market maker infers information from order flow based on that parametrization) and to model censoring (by integrating those errors terms over the relevant range for censored data). In the likelihood function, in the case that both $x_{1,s,t}$ and $x_{2,s,t}$ are observed above the point of censoring I combine information from equations (1) and (2) to solve for $i_{s,t}$. In order do so, I utilize a modeling technique to address the fact that the model imposes a restriction on the *data* – not the parameters – that does not hold exactly. The issue is that both equations (1) and (2) can be solved for ϕi. To proceed, I rewrite (1) and (2) with a noise term added to each, $\xi_{1,s,t}$ and $\xi_{2,s,t}$, respectively. I assume both of these variables are i.i.d. $N(0, \sigma_\xi^2)$ and independent of all other random variables. I take $\sigma_\xi^2 \ll \beta_i^2(\sigma_\epsilon^2 + \sigma_\eta^2)$, $i = 1, 2$, so that the likelihood function ignores this noise (since it is so small relative to other sources of variation). The intuition I hope to gain from the structural model is based on *imposing* that hedge fund trades are based on information. Rather than introducing a free parameter for noise in hedge fund trading, I introduce a minimal amount of noise to informed trading to implement the model. The result of this technique is that if both $x_{1,s,t}$ and $x_{2,s,t}$ are observed above the point of censoring, the model simply averages the information they contain for $i_{s,t}$ (from equations (1) and (2)).

C.2 CARA informed trader and public information shocks

I solve the model with risk aversion and new information shocks.

First, I assume the informed trader has CARA utility. Second, I assume that public "new information" arrives during each period that is independent of the original information draw. This information is public in the sense that it moves prices *without* trading as soon as it is generated. Note that a risk neutral trader would simply ignore such information (if it is mean zero) in her optimization.

Specifically, the informed trader has utility $U = -e^{-aW}$, where W is her wealth at the end of the current two-quarter information event. The informed trader's maximization now must take into account the risk generated by noise trading and the public new information events. The informed trader maximizes $E[profits] - \frac{a}{2}\, variance(profits)$.

Denote the public new information shock each period as $ni_t \sim N(0, \sigma_{ni}^2)$. Profits from the second period of trading are $x_2(\epsilon - p_1 - \lambda_2 u_2 + ni_2) - x_2^2\lambda_2$. Trading during period 2 is assumed to take place before the price shock ni_2, and hence new information in period 2 affects profits on the quantity traded during period 2 (x_2). Choosing x_2 to maximize the mean minus ($\frac{a}{2}*$) the variance of this quantity produces $x_2 = \beta_2(\phi i - p_1)$, with $\beta_2 = 1/D$, and $D = \{2\lambda_2 + a\lambda_2^2\sigma_u^2 + a\sigma_{ni}^2 + a(1-\phi)\sigma_\epsilon^2\}$. Note from this equation that, strictly speaking, with risk aversion the optimal amount of informed trading is no longer linearly related to the mispricing divided by the expected magnitude of noise trading. The former enters linearly, but the latter does not. As a result, my empirical proxy for x_t may not be as effective in this extension.

Stepping back to the first period, the informed trader maximizes the mean minus ($\frac{a}{2}*$) the variance of $x_1(\epsilon - \lambda_1 u_1 + ni_2 + ni_1) - x_1^2\lambda_1 + \pi * (profits\, from\, second\, period)$ over x_1. The quantity traded during period 1 is assumed to be subject to both price shocks ni_1 and ni_2.

This produces $x_1 = \beta_1(\phi i)$, with β_1 satisfying the following equation:

$$\begin{aligned}\beta_1 * \{-(-2\lambda_1 D^4 + 2\lambda_1^2\pi D^3 - 2\pi\lambda_1^2\lambda_2 D^2) + a\{(1-\phi)\sigma_\epsilon^2(D^2 - \pi\lambda_1 D)^2 + \\ \sigma_u^2(\pi^2\lambda_1^2\lambda_2^2 D^2 + (-\lambda_1 D^2 + 2\pi\lambda_1^2 D - 2\pi\lambda_1^2\lambda_2)^2 + \sigma_{ni}^2(D^2 - \pi\lambda_1 D)^2 + \sigma_{ni}^2 D^4\}\} \\ = \{D^4 - 2\pi\lambda_1 D^3 + 2\pi\lambda_1\lambda_2 D^2 - a\{(1-\phi)\sigma_\epsilon^2\pi(D^3 - \pi\lambda_1 D^2) + \\ \sigma_u^2(-\pi^2\lambda_1\lambda_2^2 D^2 + (-\lambda_1 D^2 + 2\pi\lambda_1^2 D - 2\pi\lambda_1^2\lambda_2)(-\pi\lambda_1 D + 2\pi\lambda_1\lambda_2)) + \sigma_{ni}^2\pi(D^3 - \pi\lambda_1 D^2)\}\}\end{aligned} \tag{14}$$

The market maker proceeds as before, given that the informed trader will trade an amount proportional to β_t times the remaining mispricing.

The likelihood function is the same as above. The constraints, however, can no longer be reduced to a single constraint. Instead, I numerically solve the revised constraint for β_1 jointly with equations (8) and (9). Given β_1, λ_1, and λ_2, I calculate D, which in turn gives β_2.

C.3 Converting short-horizon price impact estimates to a quarterly horizon

As a point of comparison, I linearly aggregate three existing short-horizon estimates of *total* price impact (temporary plus permanent) across a calendar quarter. Two of these estimates are from the academic literature, while one is an industry estimate. Reassuringly, my estimate of the permanent price impact component is less than these estimates of total price impact.[57]

Frazzini, Israel, and Moskowitz (2012) estimate that trading 1% of daily volume in a U.S. equity generates 1.30 bps of market impact (their Table 5, column 8, the coefficient that describes price impact that is linear in the fraction of daily volume). Aggregating this figure suggests that trading 1% of volume for an entire quarter generates: 63 trading days *1.30 bps = 0.82% total price impact.

Collin-Dufresne and Fos (2015) find that on average, when 13D filers trade prior to their public filing date, they take up 26.1% of daily volume (their Table 1 row 10). On those same days, the excess return averages 34 bps (their Table 6 column 2). Thus trading 1% of daily volume generates an estimated market impact of (34 bps / 26.1% =) 1.30 bps, the same figure as in Frazzini, Israel, and Moskowitz (2012).

Brennan and Subrahmanyam (1996) find that the the average price impact generated by purchasing 1% of the shares outstanding of a stock in the middle quintile of size (market cap) and the middle quintile of illiquidity is 1.7% (their Table 1, panel B, estimates of C_n). In my sample, on average quarterly volume is roughly 50% of the market cap of a stock (my Table 1). Brennan and Subrahmanyam's estimate thus implies that trading 1% of volume for an entire quarter generates: 1.7% *50% = 0.85% of total price impact.

Investment Technology Group estimates a price impact of approximately 85 bps for consuming 5% of the volume in a $1.4 billion market cap stock over 30 days.[58] This estimate is based on the average execution price of an order (the weighted average of shares traded and the price of each transaction), so it represents a lower bound on the total price impact (final price minus initial price). Early trades will presumably be executed before prices have moved substantially. Nevertheless, aggregating ITG's estimate suggests that trading 1% of volume for an entire quarter generates at least: 85 bps /5% *3 months = 0.51% of total price impact. At the extreme, if one assumes that all price impact is permanent and that component trades are made in infinitesimally small amounts, then the price impact after all trades have been executed will simply be twice this amount (1.02%). Thus ITG's comparable estimate of price impact most likely falls somewhere in the range of 0.51% to 1.02%, the midpoint of which is 0.77%.

[57]These calculations do not account for the fact that I measure trades relative to lagged volume, while the authors of these estimates measure trades relative to contemporaneous volume. In my sample, contemporaneous volume tends to increase relative to lagged volume for high volume consumed positions. Using contemporaneous volume, my estimates of permanent price impact would decline slightly.

[58]Hanson, Samuel G. "The FLV Capital Trading Desk (A)." Harvard Business School Teaching Note 215-053, January 2015.

D Data appendix

D.1 Standardized earnings surprises (SUE)

In the data, to ensure that earnings reflect firm performance over the same period that hedge funds are trading, I include only companies with calendar quarter-end fiscal periods to match the 13F effective dates. For these companies, hedge fund trading over the course of quarter t can be mapped to an earnings release that reflects company performance over that same quarter.

The earnings return in quarter $t+1$ is measured as the return over the three trading-day window centered around the first Compustat earnings announcement date in quarter $t+1$ for stock s, using characteristic-adjusted daily returns. The return on "other days" in the quarter is the average daily characteristic-adjusted return outside of the earnings window, multiplied by three for comparability.

Standardized unexpected earnings (SUE), $SUE_{s,t}$, is measured as $\frac{earnings_{s,t+1} - median\ analyst\ forecast_{s,t}}{P_{s,t}}$, as in Baker, Litov, Wachter, and Wurgler (2010). To form the median forecast, I take the median across the last earnings forecast made by each analyst who published an earnings forecast during quarter t. I use only analyst forecasts made during quarter t to ensure that forecasts are made during the same time interval over which I measure hedge fund trades.

To faciliate interpretation, I standardize SUE in the cross section by quarter.

D.2 Constructing mutual fund flows

I identify funds subject to extreme fund-flows as in Coval and Stafford (2007, CS).

First, I link CRSP mutual fund returns and assets to the Thompson Reuters mutual fund holdings data, using the MFLINKS dataset provided by WRDS. As in CS, I remove funds with an IOC code (Thompson Reuters) of international, municipal bonds, bonds and preferred, or metals (1, 5, 6, or 8). I also eliminate funds with fewer than 5 holdings or with less than \$5 million in assets. I aggregate multiple share classes in CRSP (which are all linked to a single Thompson Reuters fund-quarter holdings entry), summing assets and forming returns as the asset-weighted average return of the underlying share classes. I then use the CRSP data to measure fund flows for fund f during quarter t: $FLOW^{crsp}_{f,t} = assets^{crsp}_{f,t} - assets^{crsp}_{f,t-1} * (1 + ret^{crsp}_{f,t}) - mergers^{crsp}_{f,t}$, where $ret^{crsp}_{f,t}$ is the return of fund f from the end of quarter $t-1$ until the end of quarter t, $assets^{crsp}_{f,t}$ is the total net assets of fund f at the end of quarter t, and $mergers^{crsp}_{f,t}$ represents the assets that fund f gained from mutual fund mergers during quarter t. I denote these variables as "CRSP" variables to explicitly denote that they are taken from CRSP, as opposed to returns and flows calculated using the holdings data (13Fs or mutual fund holdings). I then translate this into "relative" flows at the fund level: $flow^{crsp}_{f,t} = \frac{FLOW^{crsp}_{f,t}}{assets^{crsp}_{f,t-1}}$. I sort mutual funds into deciles at the end of each quarter t based on their $flow^{crsp}_{f,t}$.

Funds in the top decile of flows ($flow^{crsp}_{f,t}$) are "extreme inflow" funds, while funds in the bottom decile are "extreme outflow" funds.

D.3 Constructing funds' return gaps

Kacperczyk, Sialm, and Zheng (2008, KSZ) construct a measure of the differential between a fund's returns and the returns of its underlying holdings (assuming that trades are made costlessly at period ends), dubbed the fund's "return gap." KSZ find that funds with the highest return gaps (where the fund returns are much greater than the holding returns) generate the highest overall fund-level returns.

KSZ's Appendix A lists a comprehensive explanation of their sample selection. I follow the same process to identify mutual funds to include in the sample. KSZ filter by the Thompson Reuters (IOC) and CRSP (ICDI, Strategic Insight, Weisenberger, Policy) mutual fund objective codes. They also eliminate funds that hold less than 80% or above 105% in stocks, on average. KSZ eliminate funds with fewer than 10 holdings or with less than \$5 million in assets. They aggregate share classes in CRSP by forming the asset-weighted return of different shareclasses before matching with the Thompson Reuters holding data. I follow all of these procedures.

The monthly return gap is the differential between a fund's gross returns reported to CRSP (formed by taking the net return each month and adding back the expense ratio divided by 12) and the returns of the fund's most recently reported asset holdings during that month. With m indexing months, the net fund return is reported to CRSP, $ret^{crsp}_{f,m}$, as described above in Appendix D.2. $grossret^{CRSP}_{f,m} = ret^{crsp}_{f,m} + \frac{expense\,ratio_{f,m-1}}{12}$, where $expense\,ratio_{f,m-1}$ is the fund's most recently reported annual expense ratio as of the previous month end. $holdret_{f,m} = \frac{\sum_s ret_{s,m} * shares_{s,f,m-1} * P_{s,m-1}}{\sum_s shares_{s,f,m-1} * P_{s,m-1}}$, where $shares_{s,f,m-1}$ are the most recent shareholdings reported by manager f in stock s as of the end of the previous month $(m-1)$, $P_{s,m-1}$ is the price of stock s as of the most recent month end $(m-1)$, and $ret_{s,m}$ is the total return of stock s during month m. I include fund holdings that are up to six months old when calculating holding period returns.

At each calendar quarter end, I rank funds into deciles based on $\sum_{k=1}^{12} grossret^{CRSP}_{f,m-k+1} - holdret_{f,m-k+1}$, where m indexes months. The top decile includes funds with the highest return gap. In my analysis in Section 5.2, because I analyze fund trades I only include funds that file consecutive quarter end holdings reports.

E Additional results

E.1 Additional summary tables

Table E.1: Summary statistics

This table displays additional summary statistics of volume consumed (aggregation method 1) portfolios by decile. Calculations are based on 13F filings from 12/31/1989-9/30/2012. Statistics are calculated as the time-series average across 13F filings. The value at each quarter t is calculated as the equal-weighted average across all stocks s in the corresponding decile portfolio at t. For manager statistics, before averaging across stocks a data point is calculated for each stock s as the equal-weighted average across all funds f who purchased stock s during quarter t. For stock-characteristic quintile averages, the value of a given characteristic for stock s is calculated as of the end of quarter $t-1$, to distinguish stock characteristics from the potential price impact of trades during quarter t. For quintiles, a value of 5 represents a higher measure of the underlying statistic, i.e., the largest market cap quintile, the highest book-to-market quintile, or the highest trailing 12-month performance (excluding the most recent month) quintile. Stocks below the 20th percentile of NYSE market cap have been removed.

Decile of volume consumed (t)	Avg mgr # positions	Avg mgr assets ($ MM)	Avg mgr age (quarters)	Avg stock size quintile	Avg stock book quintile	Avg stock momentum quintile
1	88.4	$478	21.4	3.61	2.71	3.07
2	86.2	$559	20.9	3.69	2.69	3.13
3	84.1	$602	20.8	3.66	2.69	3.14
4	83.0	$652	21.1	3.59	2.73	3.16
5	81.8	$708	21.2	3.50	2.75	3.19
6	81.0	$780	21.5	3.40	2.76	3.22
7	79.4	$821	21.4	3.28	2.78	3.23
8	78.1	$898	21.5	3.14	2.79	3.17
9	76.0	$1,085	21.9	2.97	2.82	3.16
10	72.5	$1,116	21.9	2.73	2.85	3.14

E.2 Contemporaneous performance portfolios

Table E.2: Contemporaneous performance

This table displays the contemporaneous market-adjusted and characteristic-adjusted monthly performance of calendar-time portfolios sorted into deciles based on volume consumed in quarter t by aggregation methods 1 (columns 1-2), 2 (column 3), and 3 (column 4). For comparison, I also display the monthly performance during quarter $t+1$ of portfolios of all stocks sorted by quarter t characteristic-adjusted performance (method †) and the monthly performance during quarter t of portfolios sorted by $valoftrade^{open}_{s,t} = shares\,traded_{s,t} * P_{s,t-1}$ (method ‡). Calculations are based on 13F filings from 12/31/1989-9/30/2012. Positions are weighted equally. T-statistics are displayed in brackets. ** and * denote significance at the 5% and 10% levels, respectively.

Column:	(1)	(2)	(3)	(4)	(5)	(6)
Agg. method:	(1)	(1)	(2)	(3)	(†)	(‡)
Decile of volume consumed (t)	Mkt.-adj ret (t)	Char.-adj ret (t)	Char.-adj ret (t)	Char.-adj ret (t)	Char.-adj ret (t+1)	Char.-adj ret (t)
1	0.25%	0.14%	0.44%	0.29%	0.03%	0.72%
	[1.43]	[1.47]	[6.29]**	[2.38]**	[0.11]	[5.14]**
2	0.43%	0.27%	0.51%	0.03%	0.16%	0.95%
	[2.42]**	[2.93]**	[6.95]**	[0.34]	[1.32]	[8.71]**
3	0.49%	0.37%	0.51%	-0.15%	0.18%	0.99%
	[2.95]**	[4.25]**	[6.64]**	[-1.79]*	[2.37]**	[9.95]**
4	0.60%	0.46%	0.50%	-0.09%	0.16%	0.82%
	[3.39]**	[4.85]**	[6.57]**	[-0.99]	[2.78]**	[7.91]**
5	0.65%	0.49%	0.64%	-0.09%	0.04%	0.78%
	[3.76]**	[5.67]**	[8.46]**	[-0.93]	[0.68]	[8.32]**
6	0.93%	0.69%	0.73%	0.33%	0.09%	0.80%
	[4.95]**	[7.87]**	[9.11]**	[3.51]**	[1.59]	[9.31]**
7	0.95%	0.73%	0.90%	0.55%	-0.01%	0.65%
	[5.01]**	[7.63]**	[12.38]**	[6.69]**	[-0.20]	[6.97]**
8	0.93%	0.73%	1.03%	0.76%	-0.12%	0.52%
	[4.68]**	[7.15]**	[12.36]**	[7.80]**	[-1.90]*	[5.70]**
9	1.38%	1.18%	1.15%	0.92%	-0.07%	0.54%
	[6.44]**	[10.13]**	[13.74]**	[9.29]**	[-0.81]	[6.06]**
10	2.28%	2.07%	1.63%	1.75%	0.03%	0.21%
	[9.90]**	[13.69]**	[14.61]**	[14.63]**	[0.16]	[2.04]**
L/S (10-1)	2.04%	1.94%	1.18%	1.47%	0.00%	-0.52%
	[9.54]**	[10.44]**	[8.44]**	[13.05]**	[0.01]	[-3.00]**

E.3 Future trading portfolios

Table E.3: Future trading

This table displays the volume consumed (% of quarterly volume) during quarter t+1 of calendar-time portfolios sorted into deciles based on volume consumed in t by aggregation methods 1 (columns 1-2), 2 (column 3), and 3 (column 4). For each portfolio, volume consumed in quarter $t+1$ is calculated using the same aggregation method used to calculate volume consumed during quarter t. Calculations are based on 13F filings from 12/31/1989-9/30/2012. Positions are weighted equally. T-statistics of the long-short portfolios are displayed in brackets. ** and * denote significance at the 5% and 10% levels, respectively. Volume consumed has been winsorized at the 1%/99% levels.

Column:	(1)	(2)	(3)	(4)
Agg. method:	(1)	(1)	(2)	(3)
Decile of volume consumed (t)	Volume consumed (t)	Volume consumed (t+1)	Volume consumed (t+1)	Volume consumed (t+1)
1	0.05%	1.01%	0.01%	-1.21%
2	0.18%	1.22%	0.03%	-0.33%
3	0.37%	1.37%	0.05%	-0.05%
4	0.63%	1.52%	0.08%	0.12%
5	0.99%	1.81%	0.12%	0.33%
6	1.52%	2.06%	0.17%	0.42%
7	2.33%	2.42%	0.27%	0.46%
8	3.72%	2.94%	0.43%	0.48%
9	6.62%	3.75%	0.82%	0.63%
10	17.63%	5.11%	2.90%	0.98%
L/S (10-1)	17.59%	4.09%	2.89%	2.19%
	[34.02]**	[22.17]**	[23.28]**	[17.45]**

E.4 Additional tables for earnings regressions

Table E.4: SUE and earnings returns

This table displays additional results involving earnings announcement returns and standardized earnings surprises. The characteristic-adjusted earnings return measures the return of stock s during the three trading-day window centered around its first earnings announcement during quarter t+1. SUE is the standardized earnings surprise for stock s, normalized to have a cross-sectional standard deviation of one each quarter. $ME_{s,t}$, $V_{s,t-1}^{-1}$, $IOR_{s,t}$, and $BEME_{s,t}$ are the log of market cap, the log of the inverse of dollar volume, the level of institutional ownership, and the log of the book-to-market ratio of stock s at the end of quarter t (t-1 for volume), respectively. All variables are winsorized at the 1%/99% levels. Calculations are based on 13F filings from 12/31/1989-9/30/2012. T-statistics are displayed in brackets. ** and * denote significance at the 5% and 10% levels, respectively. Panel A shows the coefficient on SUE by three groups of volume consumed (none or bottom quintile, the middle three quintiles, and the top quintile) in a Fama-MacBeth regression of the earnings return on SUE using observations with positive SUE. Panel B repeats the analysis of panel A using using observations with negative SUE.

Panel A: Regression of characteristic-adjusted earnings returns (t+1) on positive SUE (t+1) by volume consumed groups

Volume consumed (t)	Coefficient on SUE (t+1)	[t-stat]
None or bottom quintile	2.52%	[7.93]**
Middle quintiles	2.08%	[8.73]**
Top quintile	1.50%	[3.62]**
Volume consumed (t)	**Constant**	**[t-stat]**
None or bottom quintile	omitted	
Middle quintiles	0.16%	[1.65]*
Top quintile	0.54%	[3.37]**
Controls	**Coefficient**	**[t-stat]**
$ME_{s,t}$	-0.31%	[-4.01]**
$V_{s,t-1}^{-1}$	-0.06%	[-0.83]
$IOR_{s,t}$	0.75%	[4.23]**
$BEME_{s,t}$	-0.24%	[-5.49]**
Test	**F-stat**	**p-value**
SUE coefficient: top-bottom?	6.35	0.014**
Fama-MacBeth		Y
Observations		80,362
R-squared		0.047

Panel B: Regression of characteristic-adjusted earnings returns (t+1) on negative SUE (t+1) by volume consumed groups

Volume consumed (t)	Coefficient on SUE (t+1)	[t-stat]
None or bottom quintile	0.42%	[8.05]**
Middle quintiles	0.30%	[4.55]**
Top quintile	0.49%	[3.69]**

Volume consumed (t)	Constant	[t-stat]
None or bottom quintile	omitted	
Middle quintiles	0.21%	[2.49]**
Top quintile	0.89%	[5.53]**

Controls	Coefficient	[t-stat]
$ME_{s,t}$	0.35%	[3.87]**
$V_{s,t-1}^{-1}$	0.16%	[2.08]**
$IOR_{s,t}$	-1.10%	[-4.17]**
$BEME_{s,t}$	0.50%	[6.56]**

Test	F-stat	p-value
SUE coefficient: top-bottom?	0.24	0.627
Fama-MacBeth		Y
Observations		40,387
R-squared		0.047

E.5 Mutual fund portfolios

Table E.5: Mutual fund trades, volume consumed, and performance

This table displays the volume consumed (% of quarterly volume) and monthly performance of mutual fund trades. Stocks are sorted into deciles based on volume consumed (aggregation method 1) during quarter t. Calculations are based on mutual fund holdings from 12/31/1989-9/30/2012 (except for active share results, which end at 12/31/2009). Positions are weighted equally. T-statistics are displayed in brackets. ** and * denote significance at the 5% and 10% levels, respectively. In calculating future performance, stocks below the 20th percentile of NYSE market cap have been removed. Volume consumed has been winsorized at the 1%/99% levels. Panel A displays volume consumed, contemporaneous performance, and future performance of mutual fund trades. Panel B displays future performance of the trades of subsets of mutual funds: funds in the top/bottom quintile of return gap or funds with above/below median active share.

Panel A: Mutual fund volume consumed – contemporaneous performance and future performance

Column:	(1)	(2)	(3)	(4)	(5)
Decile of volume consumed (t)	Volume consumed (t)	Mkt.-adj ret (t)	Char.-adj ret (t)	Mkt.-adj ret (t+1)	Char.-adj ret (t+1)
1	0.10%	0.22%	-0.01%	0.15%	0.03%
		[1.35]	[0.10]	[0.84]	[0.32]
2	0.31%	0.40%	0.22%	0.19%	0.10%
		[2.61]*	[2.31]**	[1.29]	[1.10]
3	0.56%	0.57%	0.36%	0.25%	0.05%
		[3.73]**	[4.39]**	[1.76]*	[0.55]
4	0.83%	0.67%	0.45%	0.11%	0.06%
		[4.53]**	[5.64]**	[0.83]	[0.72]
5	1.14%	0.81%	0.59%	0.14%	0.05%
		[5.19]**	[6.61]**	[1.13]	[0.66]
6	1.54%	0.85%	0.70%	0.20%	0.10%
		[5.65]**	[8.24]**	[1.56]	[1.36]
7	2.06%	0.91%	0.68%	0.21%	0.11%
		[5.68]**	[7.96]**	[1.66]*	[1.42]
8	2.82%	1.02%	0.79%	0.20%	0.07%
		[5.77]**	[7.96]**	[1.64]	[0.83]
9	4.20%	0.96%	0.74%	0.11%	-0.05%
		[5.11]**	[6.76]**	[0.93]	[-0.67]
10	7.48%	1.06%	0.79%	0.22%	0.17%
		[4.38]**	[5.28]**	[1.79]*	[1.85]*
L/S (10-1)	7.38%	1.05%	0.80%	0.07%	0.14%
		[5.45]**	[5.22]**	[0.46]	[1.08]

Panel B: Mutual fund subsets – future performance

Column:	(1)	(2)	(3)	(4)
	Return gap, top quintile	Return gap, bottom quintile	Active share > median	Active share < median
Decile of volume consumed (t)	Char.-adj ret (t+1)	Char.-adj ret (t+1)	Char.-adj ret (t+1)	Char.-adj ret (t+1)
1	0.20%	0.13%	0.03%	0.10%
	[1.75]*	[1.16]	[0.29]	[0.79]
2	0.03%	0.01%	0.10%	0.00%
	[0.28]	[0.13]	[0.93]	[-0.02]
3	-0.07%	0.12%	0.00%	0.05%
	[-0.71]	[1.20]	[-0.02]	[0.48]
4	0.01%	0.05%	0.21%	0.05%
	[0.09]	[0.61]	[2.04]**	[0.54]
5	0.06%	0.06%	-0.09%	0.03%
	[0.59]	[0.58]	[-1.08]	[0.36]
6	0.07%	0.03%	0.09%	-0.02%
	[0.73]	[0.41]	[0.91]	[-0.21]
7	0.25%	0.19%	-0.02%	0.14%
	[2.50]**	[2.16]**	[-0.20]	[1.89]*
8	0.26%	0.10%	0.39%	0.05%
	[2.49]**	[1.08]	[4.27]**	[0.56]
9	0.21%	0.23%	0.26%	0.09%
	[1.90]*	[2.32]**	[2.69]**	[1.09]
10	0.37%	0.07%	0.29%	-0.02%
	[3.15]**	[0.58]	[2.78]**	[-0.21]
L/S (10-1)	0.18%	-0.06%	0.26%	-0.12%
	[1.03]	[-0.42]	[1.62]	[-0.82]

E.6 Wurgler and Zhuravskaya (2002) and "best ideas"

Wurgler and Zhuravskaya (2002, WZ) also motivate weighted idiosyncratic risk as a trade-level limit, similar to Cohen, Polk, and Silli (2010). WZ model an arbitrageur that has exponential utility with constant absolute risk aversion (who is thus a mean-variance optimizer). When the arbitrageur is aware of a mispriced stock, she buys (or sells) that stock and attempts to hedge the position with a substitute portfolio. In this framework, idiosyncratic risk captures the risk of the trade after hedging. WZ show that variation in idiosyncratic risk helps explain cross-sectional variation in stock returns around index additions.

WZ use two empirical proxies for idiosyncratic risk. The first proxy is the variance of the simple market-adjusted return of a stock. The second proxy is the variance of a stock's return relative to the return of a characteristic-matched portfolio. The matching portfolio is constructed by finding three stocks in the same industry with similar market capitalizations and book-to-market ratios to the stock in question.

CAPM idiosyncratic variance – which I employ in Section 7.2 to identify funds' "best ideas" – closely corresponds to WZ's first proxy. It captures the risk remaining in a stock after the arbitrageur hedges that stock using the (beta-weighted) market portfolio. In unreported results, I also employ the variance of stocks' characteristic-adjusted returns to proxy for idiosyncratic risk when identifying funds' best ideas. This proxy is similar in spirit to WZ's second proxy. This measure of risk implicitly supposes the arbitrageur hedges her position in a stock with its characteristic-matched (DGTW) portfolio. Using characteristic-adjusted idiosyncratic risk produces similar results to my results using CAPM idiosyncratic risk. Best ideas remains uninformative.

Robustness to this variation is consistent with WZ. WZ find that the correlation between their two measures of idiosyncratic risk is 0.98. WZ find that idiosyncratic risk is difficult to hedge in general, as it is hard to find close substitutes for individual stocks.

E.7 Best ideas extended results

Table E.6: Volume consumed and best ideas

This table displays results comparing volume consumed and best ideas. It shows the characteristic-adjusted monthly performance during quarter $t+1$ of calendar-time portfolios sorted independently along measures of volume consumed and best ideas in quarter t. Positive volume consumed (aggregation method 2) positions are sorted into quintiles, with all positions with zero or negative values placed into a separate bin. Positions are independently sorted by their intra-manager best ideas ranking (relative to other stocks s in fund f's portfolio at quarter t). Calculations are based on 13F filings from 12/31/1989-9/30/2012. Positions are weighted equally. Stocks below the 20th percentile of NYSE market cap have been removed. T-statistics are displayed in brackets. ** and * denote significance at the 5% and 10% levels, respectively. The proportion of total positions within each bin is displayed in italics.

Char.-adj ret (t+1) / [t-stat] / *proportion of total positions*)

Best ideas position rank (t; 1 = highest best ideas)

		21+	11-20	6-10	4-5	2-3	1
		0.08%	0.08%	0.05%	0.06%	0.05%	0.06%
	Sale or hold	[2.95]**	[1.81]*	[0.82]	[0.64]	[0.51]	[0.38]
		32.4%	*7.3%*	*3.8%*	*1.5%*	*1.5%*	*0.7%*
		0.03%	0.27%	0.02%	0.31%	0.11%	0.36%
	1	[0.78]	[2.96]**	[0.19]	[1.62]	[0.49]	[1.19]
		8.6%	*1.4%*	*0.7%*	*0.2%*	*0.2%*	*0.1%*
		0.04%	0.06%	-0.08%	0.02%	0.08%	0.04%
Volume consumed quintile (t)	2	[0.78]	[0.64]	-[0.55]	[0.08]	[0.38]	[0.16]
		7.5%	*1.7%*	*0.9%*	*0.3%*	*0.3%*	*0.1%*
		0.17%	0.07%	0.02%	-0.01%	-0.01%	0.28%
	3	[3.44]**	[0.80]	[0.17]	-[0.07]	-[0.03]	[1.01]
		6.5%	*2.1%*	*1.1%*	*0.4%*	*0.4%*	*0.2%*
		0.27%	0.19%	0.13%	0.09%	0.20%	-0.01%
	4	[5.40]**	[2.32]**	[1.24]	[0.59]	[1.27]	-[0.04]
		5.4%	*2.3%*	*1.3%*	*0.5%*	*0.5%*	*0.2%*
		0.30%	0.33%	0.57%	0.30%	0.30%	0.41%
	5	[4.91]**	[4.86]**	[6.27]**	[2.40]**	[2.20]**	[2.02]**
		3.9%	*2.4%*	*1.6%*	*0.7%*	*0.7%*	*0.4%*

F Competition

F.1 Competition

The Kyle model assumes the insider is an information monopolist. The model can be extended to the case of multiple informed agents.

In my empirical results, I primarily focus on aggregating purchases at the stock-quarter level because price impact should aggregate. Disaggregated purchases at the stock-fund-quarter level also strongly predict future stock performance (Table 3). In this appendix, I examine variation in how multiple funds simultaneously trade a single stock.

Holden and Subrahmanyam (1992; HS) show that in a Kyle model with multiple identically informed agents and a large number of periods, informed traders aggressively compete, rapidly pushing prices towards fair value. Foster and Viswanathan (1996; FV) and Back, Cao, and Willard (2000) show that in contrast, prices gradually move towards fair value over time – as in the single agent case – if the informed agents' private signals are sufficiently heterogeneous.[59] These theoretical studies take the level of competition as exogenously fixed.

Applying these multi-agent versions of the Kyle model to my data poses several challenges. First, competition is not exogenously fixed. Competition varies based on how funds assign their limited attention. To a first approximation, competition may be randomly assigned. I model this extension below. In reality, competition is endogenous. Skilled funds may be adept at deciding what stocks to learn more about: more mispriced stocks may attract more competition. Second, in order to model competition, one must take a stance on the information structure underlying not only asset prices and what funds know about asset prices, but also what funds know about what other funds know about asset prices. Funds act based on their expectations of competitors' behavior. Third, the models assume that agents act independently. Some funds may coordinate their actions, as many hedge fund managers share common employment and educational backgrounds.[60] Fourth, the number of time periods has major implications for some competitive effects. My assumption that trade occurs once a quarter – the frequency of my data – is more stark in such an environment.

I explicitly elaborate on the first point. In Appendix F.2, I construct a one-period Kyle model that features a random level of competition. Each of two traders are randomly active or inactive. I assume the econometrician can only observe informed purchases, an aspect of my data. This model makes a key point: observing more insiders purchasing an asset

[59] Koudijs (2014) also notes this distinction when applying the Kyle model to his data.

[60] In perhaps the best known example, a number of proteges of Julian Robertson manage hedge funds. This group of funds, known as "Tiger Cubs," frequently trade in the same stocks. "There at least 30 'Tiger Cubs'...[and] 40-odd 'Tiger Seeds,' or funds that are backed by Robertson's money....it is believed that many of the managers still share ideas," from http://www.benzinga.com/trading-ideas/long-ideas/12/09/2876699/the-five-stocks-tiger-cubs-love#ixzz3n4IPTUFL, accessed 9/15/2015.

increases estimated price impact but also increases the expected value of the asset. The correlation of the informed traders' signals determines which effect dominates.

If the econometrician observes a single informed trader purchasing a stock, the econometrician may expect that stock will perform particularly well in the future. The informed trader was able to build her position at a lower price because the second trader did not also purchase shares. However, observing two informed traders purchasing a stock increases the estimate of the asset's value. When forming the posterior distribution of the information, two observations receive more weight than one. Furthermore, the second informed trader may have been active but received a negative signal (and therefore gone unobserved).[61]

With perfectly correlated signals, observing a single informed trader leads to greater expected future returns. Signals are identical, so a second purchase would not increase the expected value of the asset. In contrast, with relatively uncorrelated signals, a second purchase increases the expected value of the asset by more than the incremental price impact. The same reasoning applies to observing a single trader purchasing $2x$ shares compared to observing two traders who each purchase $1x$ shares. I illustrate these points with a parametrized example in Appendix F.4 and Figure F.1.

Empirically, when multiple funds simultaneously purchase a stock in my sample, the amount that each fund purchases varies substantially. When at least three funds purchase a stock simultaneously, the mean ratio of the standard deviation of volume consumed divided by average volume consumed is 1.25.[62] This pattern suggests information signals may be weakly correlated.[63]

In Table F.1, I present regressions of future returns on proxies for competition: the number of funds that purchase a stock (positively related to competition) and the average volume consumed in that stock (negatively related to competition). The number of funds that purchase a stock is positively related to the stock's future returns after controlling for its volume-consumed quintile. The average volume consumed is negatively related to future returns. However, the predictive effects of these proxies are insignificant when I limit the sample to the top quintile of volume consumed, where I have the strongest evidence that hedge funds trade based on information. These findings provide some evidence for the multi-agent Kyle model in which funds have relatively uncorrelated signals (as in FV). At the very least, more observable competition for a given total amount of trading does not appear

[61]If hedge funds endogenously allocate their attention, random assignment may understate this effect. Hedge funds may actively avoid competing with each other except in assets that are particularly mispriced.

[62]That is, I calculate $\frac{var(volconsumed_{s,f,t})^{1/2}_{s,t}}{\sum_{f=1}^{F} volconsumed_{s,f,t}/N_{s,t}}$ for stock s at quarter t. $N_{s,t}$ is the number of funds with positive volume consumed in stock s during quarter t, and the volatility calculation includes only positive observations of volume consumed in stock s during quarter t. I then average across stocks and quarters.

[63]This variation could also reflect non-information based motives for trade.

to predict strongly diminished future returns (an implication of HS). The complications outlined above caution against interpreting these results too strongly.

F.2 Kyle model with a random number of informed traders (0, 1, or 2)

I construct a one-period Kyle model with an uncertain number of informed traders. Notation is the same as in Section 2, but without time subscripts. Each of two informed traders has a δ probability of being "active" in the asset, independent of whether the other trader is present. If both traders are active, they draw signals $i = \epsilon + \eta$ and $i' = \epsilon + \eta'$, with η and η' i.i.d. $N(0, \sigma_\eta^2)$. The market maker can not observe traders' presence, and therefore reacts to trades using a probability weighted average of the linear reaction function she would employ in each scenario.

"Future returns" are proxied by $\epsilon - p$. These are the returns realized by asset holders after trading takes place at price p.

The model solution proceeds as it does for the two-period model in Appendix A.

First, optimize from the perspective of an informed trader. She solves $max_x E[x(\epsilon - \lambda u - \lambda x - \delta\lambda\beta(\phi i')|\, i]$, where i is the agent's own information signal and i' is the information signal of the other agent (if that agent is active). The solution, after setting $\beta = \beta'$ (since the agents are identical), gives $\beta = \frac{1}{\lambda(2+\delta)}$.

Note that $E[p] = \lambda\beta\phi(i + i')$ if two traders are present, and $E[p] = \lambda\beta\phi i$ if one trader is present. Thus $\lambda\beta * (\#traders)$ represents the proportion of the informed traders' information that gets into prices in expectation. Suppose that at least a single informed trader is present. As $\delta \to 0$, this reduces to the classic Kyle model solution that $\lambda\beta = \frac{1}{2}$. If the single informed trader knows that the odds of her competing with another informed trader are approximately zero, then she will trade to get half of her information into price. As $\delta \to 1$, on the other hand, $\lambda\beta \to \frac{1}{3}$. Since both traders are active with certainty, that means that $2\beta\lambda = \frac{2}{3}$ of their information gets into price. Thus as more agents compete over the asset, they get more information into prices for a given true amount of information ϵ.

The market maker posts a single linear response coefficient λ (so that $p = \lambda(x + u)$). The market maker probabilistically averages her response function across the scenarios of no active informed traders, one active informed trader, and two active informed traders:
$\lambda = 2(1-\delta)\delta\frac{\beta\phi\sigma_\epsilon^2}{\beta^2\phi^2(\sigma_\epsilon^2+\sigma_\eta^2)+\sigma_u^2} + \delta^2\frac{\beta\phi\sigma_\epsilon^2}{4\beta^2\phi^2\sigma_\epsilon^2+2\beta^2\phi^2\sigma_\eta^2+\sigma_u^2}$.

F.3 Expected returns conditional on observing one vs. two traders when the econometrician only observes purchases

Assume the econometrician only observes informed purchases. In the model, if two informed traders purchase an asset, then the econometrician can infer the information of both traders. However, if the econometrician observes one informed purchase, she can not be sure if there was in fact only a single informed trader active or if instead a second informed trader was active but decided not to purchase (i.e., shorted) the asset.

In order to compare expected returns conditional upon observing informed purchases from one vs. two traders, three quantities are needed: (1) $E[\epsilon - p]$ if there is truly one trader active; (2) $E[\epsilon - p]$ if there are two traders active but one trader decides to short the asset; and (3) $E[\epsilon - p]$ if there are two traders active and both purchase the asset.

For the first quantity, the calculation is simple: $E[\epsilon - p|\, i] = E[\epsilon - \lambda\beta\phi i|\, i] = \phi i(\frac{1+\delta}{2+\delta})$, if only one trader is active and we observe i.

For the second quantity, we need the expectation of ϵ conditional on the second agent drawing a negative signal. This calculation utilizes the truncated normal distribution, so there is no analytical solution. Instead, solve for ϵ by maximizing its likelihood: $g(\frac{i-\epsilon}{\sigma_\eta})G(\frac{-\epsilon}{\sigma_\eta})g(\frac{\epsilon}{\sigma_\epsilon})$, with g and G the standard normal PDF and CDF, respectively (the first term represents the probability that the first informed trader draws a signal i, given ϵ; the second term reflects the probability that the second signal i' is negative, given ϵ; and the third term represents the prior probability that ϵ takes the given value). Then calculate the expected signal for the second trader, conditional on it being less than zero, using the moments of a truncated normal distribution (truncated at zero, with mean ϕi and variance $(1-\phi)\sigma_\epsilon^2 + \sigma_\eta^2$). Given the expected signal for the second trader, the expected value of ϵ, and ϕi for the observed trader, calculate $E[\epsilon - p] = E[\epsilon] - \lambda\beta(i + E[i'])$, with all the expectations conditional on observing i and knowing that the second unobserved trader receives a negative signal i'.

The third quantity is calculated similarly. We need the expectation of ϵ conditional on a hypothetical positive draw for the second trader. Solve for the expected value of ϵ by maximizing the likelihood $g(\frac{i-\epsilon}{\sigma_\eta})(1 - G(\frac{-\epsilon}{\sigma_\eta}))g(\frac{\epsilon}{\sigma_\epsilon})$. Then proceed as above.

Finally, calculate the expected return conditional on observing one trader purchasing the asset as the probability weighted average of (1) there being only a single active trader and (2) the possiblity that a second trader was active but decided not to purchase the asset:

$$E[\epsilon - p_1|\, observe\, one\, trader] = \frac{Prob(truly\ one\ trader)}{Prob(truly\ one\ trader) + Prob(unobserved\ second\ trader)} * E[\epsilon - p_1|\, truly\, one\, trader]$$
$$+ \frac{Prob(unobserved\ second\ trader)}{Prob(truly\ one\ trader) + Prob(unobserved\ second\ trader)} * E[\epsilon - p_1|\, unobserved\, second\, trader].$$

Compare that quantity to $E[\epsilon - p_1|\, observe\, two\, traders]$.

An alternative manner of conceptualizing this dynamic is to consider expected returns conditional on observing a single purchase of $2x$ to expected returns conditonal on observing two smaller purchases that sum to $2x$.

The latter expectation is trivial. Assuming that the two traders observe signals of i_{two} and i'_{two}, where $i_{two} + i'_{two} = i_{one}$, with i_{one} the signal of the single large trader and $i_{two}, i'_{two}, i_{one} >$

0, then $E[\epsilon - p] = E[\epsilon] - \lambda\beta\phi(i_{two} + i'_{two})$. $E[\epsilon] = \frac{\frac{2}{\sigma_\eta^2}(i_{two}+i'_{two})/2}{\frac{1}{\sigma_\epsilon^2}+\frac{2}{\sigma_\eta^2}}$, based on forming a normal posterior from a prior (the distribution of ϵ) and data (observations of i and i').

The former expectation is calculated using the same method as above: probabilistically average the expectation if the trader is active on her own and the expectation if a second trader was active but decided not to purchase the asset.

F.4 Parametrized example

To get a quantitative sense of these dynamics, I construct a parametrized example of the model. I assume the econometrician only observes informed purchases.

Figure F.1 illustrates expected returns, $E[\epsilon - p]$, as one varies the noise of the informed traders' signals, σ_η^2. I assume that $\sigma_\epsilon^2 = 1$ and $\delta = 0.5$ (note that I do not need to make an assumption on σ_u^2, since I only need to know $\lambda\beta$, not λ on its own). These results are based on 50,000 simulations of the model for each value of σ_η^2. In each simulation, I randomly draw a positive value of $i = \epsilon + \eta$. I then calculate expected returns conditional on seeing a single purchase based on that signal. I also calculate expected returns if one were to observe a (random) second informed purchase. Finally, I calculate expected returns if instead of seeing a single informed purchase, the econometrician observes two informed purchases that are each half the size of the (larger) single purchase.

Expected returns are higher conditional upon observing a single informed purchase, compared to what one would expect if one observed a second informed purchase, for $\sigma_\eta^2 \leq 1$. At higher values of σ_η^2, the increase in the expected value of ϵ from observing a second purchase outweighs the increase in price impact (expected value of p). In that part of the parameter space, expected returns are higher if the econometrician observes two informed purchases.

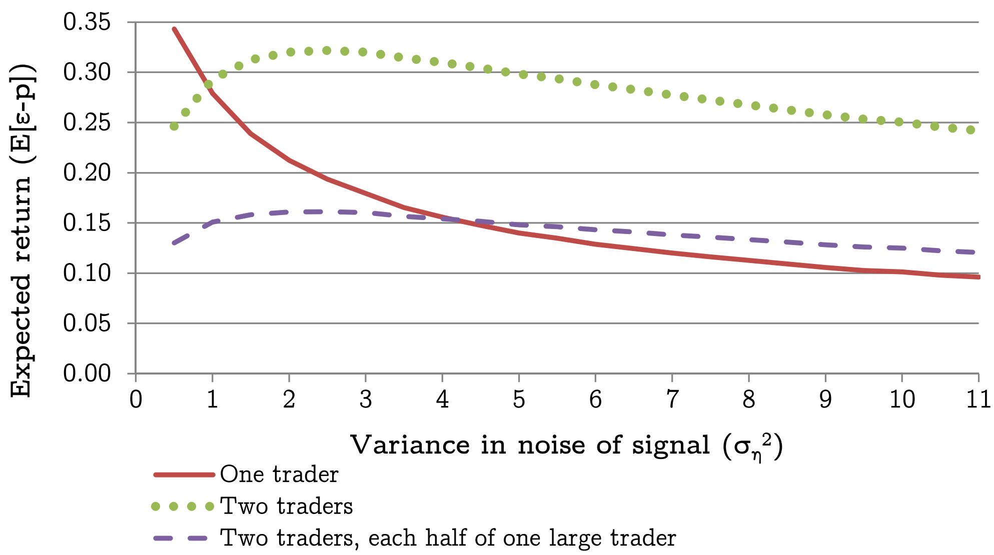
I also compare one informed purchase to two informed purchases that are each half the size of the single purchase.[64] In this scenario, the point of preference shifts to a higher value of σ_η^2. Expected returns are higher for observing a single informed purchase if $\sigma_\eta^2 \leq 4$. At higher values of σ_η^2, returns are higher conditional on observing two smaller informed purchases.

With random assignment of informed traders and an inability to observe shorts, expected returns may be higher after observing more purchases or after observing fewer purchases. The noise in informed traders' signals determines the relative ranking.

[64] Mathematically, this solution applies to any two purchases that add up to the magnitude of the single larger purchase.

Figure F.1: Competition - parametrized example

This figure displays expected returns, $E[\epsilon - p]$, conditional on different observed patterns of trading in a one-period Kyle model with two randomly assigned informed traders, as in Appendix F.2. Expected returns are displayed as a function of σ_η^2. $\sigma_\epsilon^2 = 1$ (the variance of information) and $\delta = 0.5$ (the probability a given informed trader is "active" in a stock).



F.5 Competition – Results

Table F.1: Competition and future returns

This table displays information involving competition and future monthly characteristic-adjusted returns. $\#funds_{s,t}$ is the number of hedge funds that purchased a stock s in quarter t. The sample is limited to stocks with $\#funds_{s,t} > 0$. Average volume consumed is the average volume consumed in that stock: $\frac{volconsumed_{s,t}}{\#funds_{s,t}}$. VCQ is the quintile of volume consumed (aggregation method 1; 1-5 for stocks with hedge fund purchases, and 0 for stocks with no hedge fund purchases) for stock s. $ME_{s,t}$, $V_{s,t}^{-1}$, $IOR_{s,t}$, and $BEME_{s,t}$ are the log of market cap, the log of the inverse of dollar volume, the level of institutional ownership, and the log of the book-to-market ratio of stock s at the end of quarter t (t-1 for volume), respectively. All variables are winsorized at the 1%/99% levels. Calculations are based on 13F filings from 12/31/1989-9/30/2012. T-statistics are displayed in brackets. ** and * denote significance at the 5% and 10% levels, respectively.

Column:	(1)	(2)	(3)	(4)
Dependent variable:	Char.-adj ret (t+1)	Char.-adj ret (t+1)	Char.-adj ret (t+1)	Char.-adj ret (t+1)
VCQ (t)	0.0010	0.0015		
	[3.77]**	[5.21]**		
$\#funds_{s,t}$	0.0004		0.0002	
	[2.53]**		[0.68]	
Average volume consumed (t)		-0.0341		-0.0187
		[-2.35]**		[-1.07]
$ME_{s,t}$	-0.0009	-0.0007	-0.0031	-0.0031
	[-0.75]	[-0.61]	[-2.58]**	[-2.60]**
$V_{s,t-1}^{-1}$	0.0004	0.0003	0.0024	0.0022
	[0.38]	[0.33]	[2.60]**	[2.37]**
$IOR_{s,t}$	0.0022	0.0015	0.0053	0.0046
	[0.91]	[0.59]	[1.78]*	[1.63]
$BEME_{s,t}$	0.0012	0.0012	0.0006	0.0007
	[2.57]**	[2.60]**	[0.71]	[0.86]
Fama-MacBeth	Y	Y	Y	Y
Only top quintile of volume consumed?	-	-	Y	Y
Observations	148,996	148,996	30,278	30,278
R-squared	0.024	0.024	0.033	0.033