Zooming In on Equity Factor Crowding

Bence Tóth
Capital Fund Management

OECD NAEC-EconophysiX workshop on econophysics and policy
Jan 21, 2020
Collaborators

- Valerio Volpati
- Michael Benzaquen
- Zoltán Eisler
- Iacopo Mastromatteo
- Jean-Philippe Bouchaud

What is crowding and why do we care?

Crowding
- Different investors following the same (or very similar) signals
- Expect it for well known strategies ("alternative $\beta$")

Possible effects
- Erosion of performance due to competition for the same excess returns
- Increased transaction costs due to similar trade flows
  - Co-impact (market reacts to total net flow!)
- Systemic risk due to overlapping portfolios: liquidation can trigger further ones; cascades
  - Quant Crunch 2007
If there is a crowd, it should be visible in the order flow: can we identify it?
If yes, how does it behave in time?
We will look at correlations between expected order flow (signal) and actual order flow
▶ **Standard Fama-French factors + Momentum**
  - SMB: Small [market capitalization] Minus Big
  - Momentum: Long term trends in asset prices

▶ **Scope:** Russell3000 1995-2018

▶ **Signal:** ranked, normalised

▶ **Trading is costly:** need to slow down the signal

▶ **Quadratic trading costs lead to exponential slowing down**

▶ **Expected position:**

\[
\pi_{i,t} = A \sum_{t' \leq t} s_{i,t'} \exp \left( -\frac{t - t'}{D} \right)
\]

▶ **Expected order flow**

\[
\Delta \pi_{i,t} = \pi_{i,t} - \pi_{i,t-1}
\]
▶ Standard Fama-French factors + Momentum
  – SMB: Small [market capitalization] Minus Big
  – Momentum: Long term trends in asset prices

▶ Scope: Russell3000 1995-2018

▶ Signal: ranked, normalised

▶ Trading is costly: need to slow down the signal
▶ Quadratic trading costs lead to exponential slowing down

▶ Expected position:

\[ \pi_{i,t} = A \sum_{t' \leq t} s_{i,t'} \exp \left( -\frac{t - t'}{D} \right) \]

▶ Expected order flow

\[ \Delta \pi_{i,t} = \pi_{i,t} - \pi_{i,t-1} \]
Equity-Factor Data

- Standard Fama-French factors + Momentum
  - SMB: Small [market capitalization] Minus Big
  - Momentum: Long term trends in asset prices

- Scope: Russell3000 1995-2018

- Signal: ranked, normalised

- Trading is costly: need to slow down the signal

- Quadratic trading costs lead to exponential slowing down

- Expected position:

  \[ \pi_{i,t} = A \sum_{t' \leq t} s_{i,t'} \exp \left( -\frac{t - t'}{D} \right) \]

- Expected order flow

  \[ \Delta \pi_{i,t} = \pi_{i,t} - \pi_{i,t-1} \]
Standard Fama-French factors + Momentum
- SMB: Small [market capitalization] Minus Big
- Momentum: Long term trends in asset prices

Scope: Russell3000 1995-2018
Signal: ranked, normalised
Trading is costly: need to slow down the signal
Quadratic trading costs lead to exponential slowing down
Expected position:

$$\pi_{i,t} = A \sum_{t' \leq t} s_{i,t'} \exp \left( \frac{t - t'}{D} \right)$$

Expected order flow

$$\Delta \pi_{i,t} = \pi_{i,t} - \pi_{i,t-1}$$
Microstructure data
- All trades and order book updates
- 1600 US stocks, 2011-2018
- This is **anonymous** flow
- Daily trade imbalance

\[ I_{\text{volume}}^{i,t} = \frac{\sum_{n \in t} \text{sign} \cdot \text{volume}}{\sum_{n \in t} \text{volume}} \]

- Daily order book imbalance

\[ I_{\text{book}}^{i,t} = \frac{V_{\text{bid}}^{i,t} - V_{\text{ask}}^{i,t}}{V_{\text{bid}}^{i,t} + V_{\text{ask}}^{i,t}} \]

Metaorder data
- Ancerno dateset: institutional investors
- 1999 - 2014
- Covers \( \approx 10\% \) of volume
- Data contains client identification
- Possible to **group trades to metaorder**
  - same client
  - same start/end date
  - same direction
  - same product

\[ I_{\text{meta}}^{i} = \frac{\sum m \in t \epsilon_{i,m}^{\text{meta}}}{\sum m \in t |\epsilon_{i,m}^{\text{meta}}|} \]

- Executed via limit or market orders
Quick stats

Imbalance autocorrelation

- Positive autocorrelation
- Decays very slowly
- Power-lawish: $C(\ell) \sim \ell^{-\gamma}; \gamma \approx 0.5 - 0.8$

Imbalance$_t$-return$_{t+\ell}$ correlation

- Positive at zero lag
- Zero for $\ell > 0$: past public info does not predict future returns
- Positive for $\ell < 0$: compatible with autocorr
Results

- We find a significant correlation
- Peak at 3-4 months (in line with autocorr time of signal)
- Positive correlation to OB imbalance
- Negative correlation to trade imbalance

Possible interpretations

1. Aggressive flow following mean-reversion
   - Unlikely: MR profitable on much shorter times

2. Momentum followed by passive orders
   - Quite possible (AQR paper: 80% of volume executed via limit orders)
Results

▶ We find a significant correlation
▶ Peak at 4-6 months
▶ Correlation is positive
▶ Note: very similar shape as anonymous data

Interpretation

▶ Institutionals follow Momentum strategies
▶ Execution via limit orders
Momentum III. - Price returns

Results
- Significant correlation
- But much weaker than with imbalances

Interpretation
- It is very hard to see things on prices!
- This correlation can help estimate costs
  \( C(Q) = \int_0^Q p(q) dq \)
- Need to compare to correlation of position and returns (\( \approx 0.1\% \))
- In a quadratic cost model: price paid is half of instantaneous impact
- Suggests: costs \( \approx \) gains!
- Note: other implementations still work
**Anonymous flow**

- Consistently significant correlation
- Clear upward drift since 2012
- Increased crowding

**Metaorder flow**

- Consistently significant correlation
- No clear trend can be found
- Note: periods only partially overlap
Anonymous flow

- Significant correlations
- Positive for OB, negative for trade imbalance
- Weaker than for Momentum
- Expected, since longer holding period (less rebalancing)

Metaorder flow

- Correlations are barely significant

Time evolution

- Too noisy to say anything
Conclusions and next steps

Conclusions

▶ Crowding can be quantified through correlations between imbalances and expected flow from equity factors
▶ Signals are particularly significant for Momentum: it is crowded
  – In line with recent low profitability
▶ Crowding appears to have increased in the past years
▶ Results are robust
  – Across stocks (liquidity, tick size, …)
  – Changing the mechanism for slowing down
  – Looking at only long/short leg
▶ HML and SMB: clear signs of crowding in anonymous flow data

Open questions

▶ How to relate weakening performance to crowding
  – Front-running until signal disappears?
  – And if crowd leaves, signal returns?
▶ Quantify risk of de-leveraging spirals → policy implications