Cross Sectional Momentum

Oh Kang Kwon† • Stephen Satchell†,* (ses999gb@yahoo.com)
Paris – March 2019

† Discipline of Finance, The University of Sydney
* Trinity College, University of Cambridge
Outline

- General thoughts
- Distribution of momentum returns
General Thoughts
What is Cross Sectional Momentum (CSM)?

- We have a universe of assets
- We rank their returns/earnings/prices over a ranking/formation period.
- We go long assets in top $m$-tile, short bottom $m$-tile, weighted equally or otherwise
- We then hold our position over a subsequent holding period
- Resulting return on this position is cross sectional momentum (CSM) return
- There are many variations on this structure
Profitability of Momentum

- Some observations:
  1. Momentum is profitable if returns exhibit strong deterministic trend
  2. Momentum is profitable if returns have some autocorrelation
  3. High risk positions sometimes have higher returns
  4. Observation 3 compatible with market efficiency
We argue that cross sectional momentum (CSM) is profitable when there are large differences in expected returns (high factor cross sectional volatility (CSV))

1. Europe/Asia should be good for CSM (different countries and industries)
2. UK/US should be bad for CSM (homogeneous) but UK good for momentum?
3. Japan?
Practitioners of behavioural finance would say, implicitly, Brits/Europeans and would have persistent psychological problems that do not correct.

Americans do not have these problems.

Quantitative explanation seems more plausible.
• When quantitative finance became ugly (2007–2008), it re-emerged as behavioural finance
• Academics were hired to tell tales about investors’ incurable psychological issues
• For example, Hong and Stein (1999) with different trader types under-reaction to overconfidence and overreaction to biased self-attribution
• For a prospect theoretical interpretation of momentum returns, see Menkhoff and Schmeling (2006)
Distribution of CSM Returns
• Based on 2017 JEDC paper by Oh Kang Kwon and Stephen Satchell
• Considers the CSM returns as two-period problem – ranking period and holding period
• Assumes the stock returns over the two periods are multivariate normal
• If two periods are independent and returns are stationary, then markets are efficient and high momentum returns are a consequence, presumably, of higher risk
• Construct portfolios consisting of $m$ long and $m$ short assets from a universe of $n$ assets – more generally $m_+$ long and $m_-$ short
Consider the special case \( n = 2 \) and \( m = 1 \), and let \( r_{1,t} \) and \( r_{2,t} \) be the returns on two assets over the ranking period, and \( r_{1,t+1} \) and \( r_{2,t+1} \) the corresponding returns over the holding period.

Then for CSM strategy:
- if \( r_{1,t} > r_{2,t} \), viz. in ranking period, then long asset 1 and short asset 2
- if \( r_{1,t} < r_{2,t} \), then do the opposite

This implies for resulting CSM return, \( r_{\text{csm},t+1} \), over holding period

\[
\text{pdf}(r_{\text{csm},t+1}) = \text{pdf}(r_{1,t+1} - r_{2,t+1} \mid r_{1,t} > r_{2,t}) \\
+ \text{pdf}(r_{2,t+1} - r_{1,t+1} \mid r_{1,t} < r_{2,t})
\]

If markets are efficient, this is a mixture of univariate normals

\[
\text{pdf}(r_{\text{csm},t+1}) = \text{pdf}(r_{1,t+1} - r_{2,t+1}) \text{prob}(r_{1,t} > r_{2,t}) \\
+ \text{pdf}(r_{2,t+1} - r_{1,t+1}) \text{prob}(r_{1,t} < r_{2,t}),
\]

in this case, kurtotic and skewed for plausible parameter values.
• If markets are not efficient (predictable), then structure is more complicated but given in terms of truncated normals

\[ \text{pdf}(r_{\text{csm}, t+1}) \]

\[ = \phi_1 \left( r; -\mu_{t+1}, \sigma_{t+1}^2 \right) \Phi_1 \left[ 0; \mu_t - \frac{\rho_{t,t+1} \sigma_t}{\sigma_{t+1}} (r + \mu_{t+1}) , \sigma_t^2 \left( 1 - \rho_{t,t+1}^2 \right) \right] \]

\[ + \phi_1 \left( r; \mu_{t+1}, \sigma_{t+1}^2 \right) \Phi_1 \left[ 0; -\mu_t - \frac{\rho_{t,t+1} \sigma_t}{\sigma_{t+1}} (r - \mu_{t+1}) , \sigma_t^2 \left( 1 - \rho_{t,t+1}^2 \right) \right] , \]

where for \( i, j \in \{1, 2\} \) and \( u \in \{ t, t+1 \} \)

\[ \mu_{t+1} = \mathbb{E}[r_{2,t+1} - r_{1,t+1}], \quad \sigma_{i,u}^2 = \text{var}(r_{i,u}), \quad \rho_u = \frac{\text{cov}(r_{1,u}, r_{2,u})}{\sigma_{1,u} \sigma_{2,u}} , \]

\[ \rho_{i,j} = \frac{\text{cov}(r_{i,t}, r_{j,t+1})}{\sigma_{i,t} \sigma_{j,t+1}} , \quad \sigma_t^2 = \sigma_{1,u}^2 + \sigma_{2,u}^2 - 2 \rho_u \sigma_{1,u} \sigma_{2,u} , \]

\[ \sigma_{t,t+1} = \rho_{1,1} \sigma_{1,t} \sigma_{1,t+1} + \rho_{2,2} \sigma_{2,t} \sigma_{2,t+1} - \rho_{1,2} \sigma_{1,t} \sigma_{2,t+1} - \rho_{2,1} \sigma_{2,t} \sigma_{1,t+1} , \]

\[ \rho_{t,t+1} = \frac{\sigma_{t,t+1}}{\sigma_t \sigma_{t+1}} \]

• Analytic expressions for first four central moments of \( r_{\text{csm}, t+1} \) available
• When are momentum returns positive?

• Consider again the simple case $n = 2$, $m = 1$, and market efficient

• Letting $p = \text{prob}(r_{1,t} > r_{2,t})$,

$$
E[r_{\text{csm}, t+1}] = p (E[r_{1,t+1}] - E[r_{2,t+1}])
+ (1 - p) (E[r_{2,t+1}] - E[r_{1,t+1}])
= (2p - 1) (\mu_{1,t+1} - \mu_{2,t+1}),
$$

where $\mu_{i,t+1} = E[r_{i,t+1}]$ and

$$
p = \Phi \left( \frac{\mu_{1,t} - \mu_{2,t}}{\sqrt{\sigma_{1,t}^2 + \sigma_{2,t}^2 - 2\rho_t \sigma_{1,t} \sigma_{2,t}}} \right)
$$

• So you have to be able to pick the stock with the higher expected return more than 50% of the time – not surprising!
Would high cross sectional volatilities (CSV) be good/bad for CSM?
This depends on whether it is factor CSV (good) or idiosyncratic CSV (bad).
We can see this from previous formula, factor CSV increases the numerator while idiosyncratic CSV the denominator.
Special Case $n = 3$ and $m = 1$

- Ordering of asset returns in the ranking period corresponds to $6 = 3!$ permutations of $\{1, 2, 3\}$, viz. $(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1)$
- If $(1, 2, 3)$, we long asset 1 and short asset 3, if $(1, 3, 2)$ we long asset 1 and short asset 2, etc.
- For a permutation $\pi$ of $\{1, 2, 3\}$, write $\pi_i$ for the image of $i$ so that, for example, if $\pi = (2, 1, 3)$ then $\pi_1 = 2$, $\pi_2 = 1$ and $\pi_3 = 3$. Then

$$
\text{pdf}(r_{\text{csm}, t+1}) = \sum_{\pi \in S_3} \text{pdf}(r_{\pi_1, t+1} - r_{\pi_3, t+1} \mid r_{\pi_1, t} > r_{\pi_2, t} > r_{\pi_3, t}),
$$

where $\pi$ ranges over all permutations $S_3$ of $\{1, 2, 3\}$
- Resulting distribution is from the unified skew-normal (SUN) family considered in Arellano-Valle and Azzalini (2006)
General Case of $n$ Assets

- Above results generalize naturally to universe of $n$ assets, $m_+$ long and $m_-$ short:
  - pdf for CSM return is a sum over the permutations of $\{1, 2, \ldots, n\}$
  - each term in the CSM return pdf consists of univariate normal and truncated multivariate normal
  - total number of distinct orderings of asset returns over ranking period is
    \[
    \frac{n!}{(n - m_+ - m_-)!m_+!m_-!}
    \]
  - if we want to investigate S&P500 long top 100 and short bottom 100, this number is vast
• So is this profoundly useless?
  - Perhaps for direct practical applications
  - We at least understand why momentum returns should be kurtotic
  - Even with normal returns we get non-normal momentum returns
  - Too early to link volatility spikes with momentum crashes, but framework may be able to address this \( r_{csm,t+1} \) skewness as function of correlation
  - fails to explain why long CSM makes most of the money
Link between Skewness and Correlation

- Special case where return correlations over holding period are all $\rho_{t+1}$

- S&P500 and Fama-French data suggests skewness of CSM returns tend to be negative

- $\rho_{t+1} \rightarrow 1$ related to market crashes as Sancetta and Satchell (2007) show that $\rho_{t+1} \rightarrow 1$ in a CAPM framework when market vol goes up
• Notion that expected utility maximisers take expected values over such a distribution becomes fanciful without access to modern MC
• For a long 50 and short 50 momentum portfolio from S&P500, distinct orderings over the ranking period is

\[
\frac{500!}{400!50!50!} \approx 10^{160}
\]

which is huge!
• To put things into perspective, number of seconds in the history of the universe is approximately $10^{20}$
Conclusion

- We have derived the pdf of CSM returns
- This pdf is recognisable as a density from a known family of distributions
- Results are practically usable only for small $n$
- For $n = 2$, we can capture many of the stylised facts of CSM returns