Automated pair construction

This document describes an approach to automatically identifying pair candidates.

Any related material has been stored in Insight's **Confluence** document hosting system and can be viewed by clicking on hyperlinks, which are presented in **blue, bold, underlined text**.

Contents

1 The approach

Pairs are identified in 3 steps:

- Identify longs
- Identify shorts
- Identify pairs

Both longs and shorts are identified based on 2 criteria: *efficacy* and *stacking*. Pairs are identified using the same criteria in addition to a *similarity* constraint: both legs need to belong to the same *category*.

1.1 Efficacy

This principle states that we want to be long well-performing stocks. This criterion is applies to a stock's *long term performance*. We typically look at decades-long of performance. All performances are relative to a market index. For UK stocks we use either `UKX` or `MCX` as our market indices.

1.2 Stacking

This criterion is an evaluation of a stocks recent price-action. We evaluate price action by looking at price, 10 week moving average and 40 week moving average charts. All price charts are, again, relative to a market index. For UK stocks we use either `UKS` or `MCX` as the relevant market prices. The term "stacking" is meant to imply that we would like to see moving averages that are ordered by window size. We would like to see (relative) price higher than the 10 week moving average while the 10 week moving average is higher than the 40 week moving average. This type of ordering is referred to as a "stack". Good longs will have good stacks.

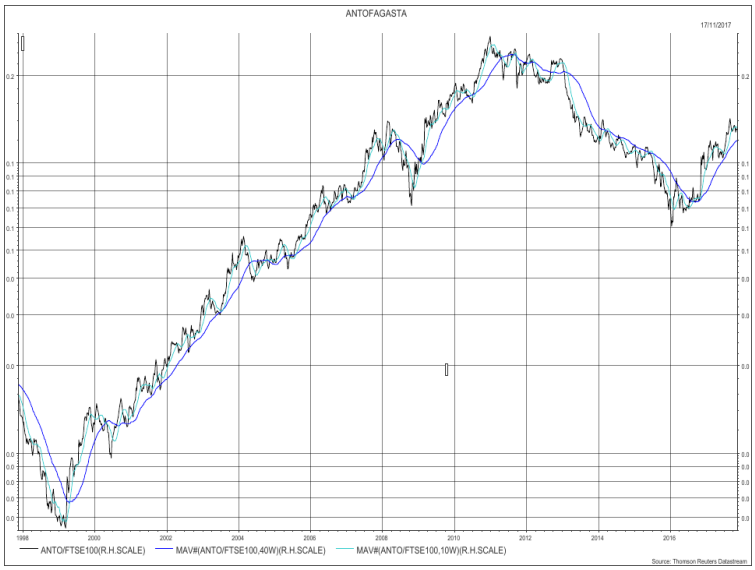
Additional items relating to the "stack" criterion:

- Charts need to be on logarithmic scale.
- The components of a stack need to be close as well as ordered. This indicates that a "base" is forming.
- A stack can change from "good" to "bad" over time. A stack that went "bad" and recovered is a better stack.
- The distance between the 10 week and the 40 week moving averages is a measure of a stock's risk.
- All charts are on relative prices (not div adjustemt)

2 Examples: efficacy

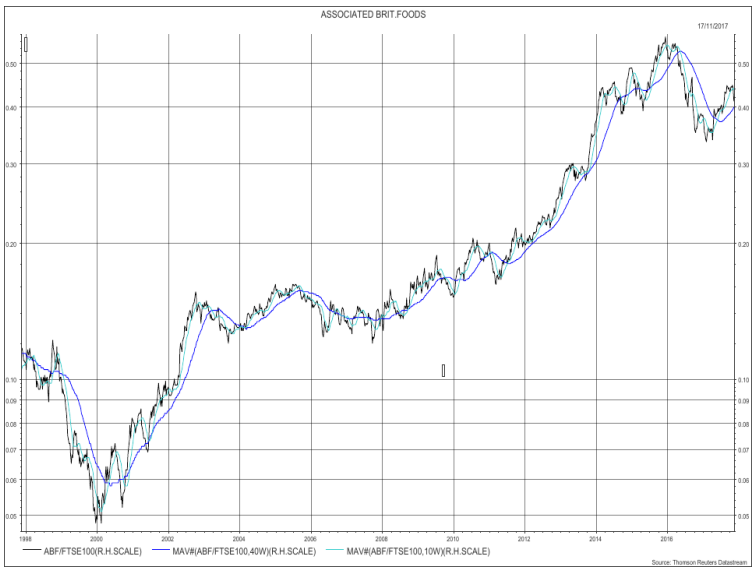
Good efficacy: ANTOFAGASTA

- ✓ Performance
- ✓ Persistent trends
- ✓ Low volatility



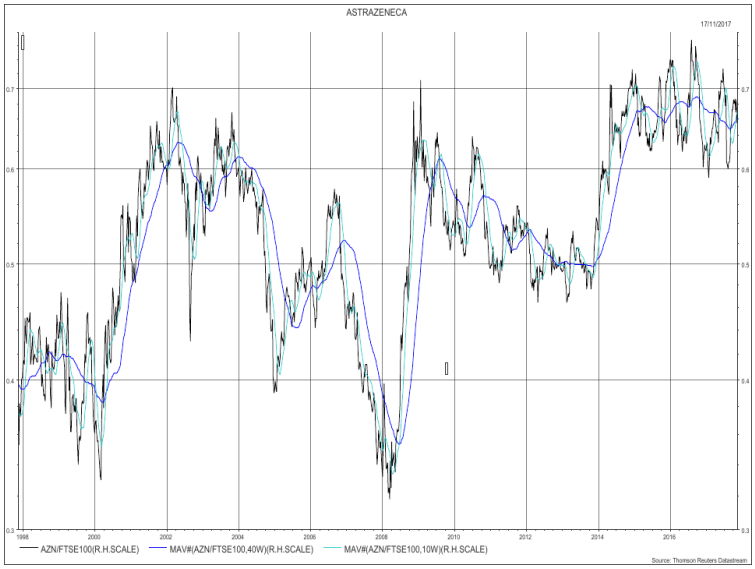
Good efficacy: ASSOCIATEDBRIT.FOODS

- ✓ Performance
- ✓ Persistent trends
- ✓ Low volatility



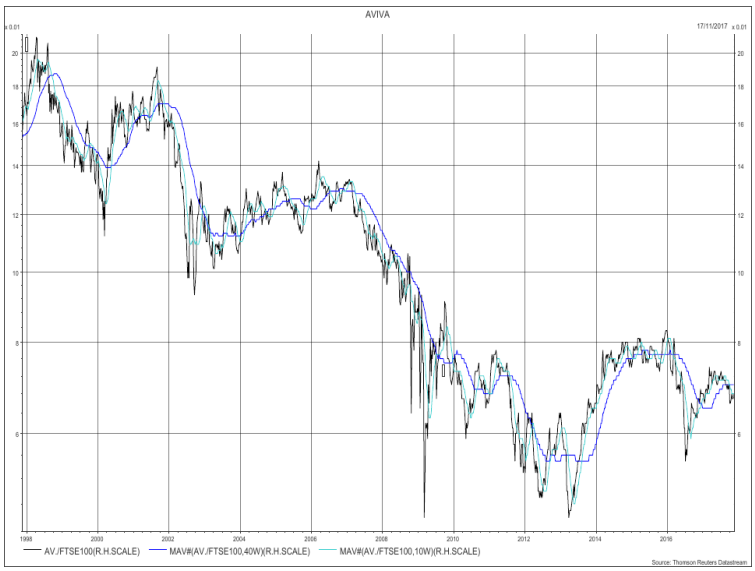
Bad efficacy: ASTRAZENECA

- ✗ Performance
- ✗ No trends
- ✗ High volatility
- ✗ Large price reversals



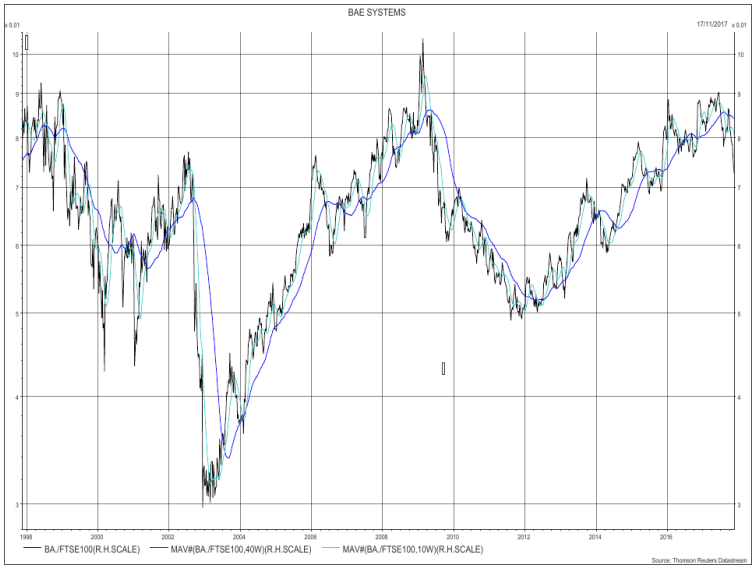
Bad efficacy: AVIVA

- ✗ Performance



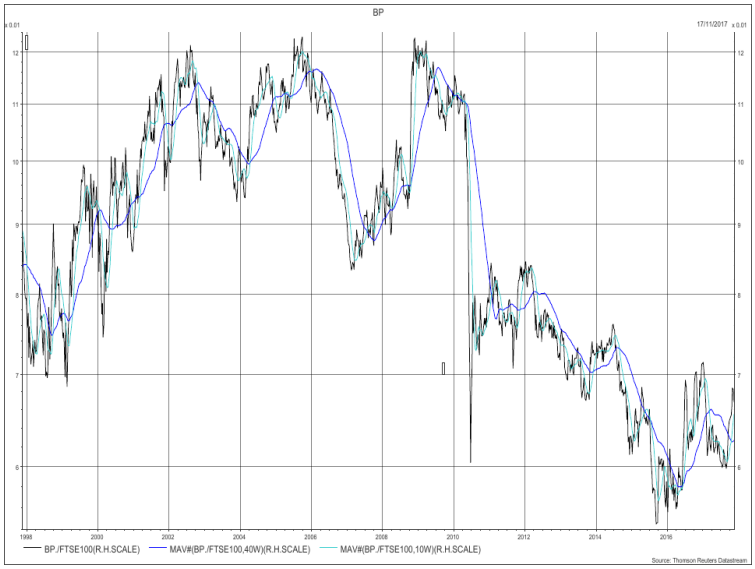
Bad efficacy: BAESYSTEMS

- ✗ No long-term performance
- ✗ Large price reversals



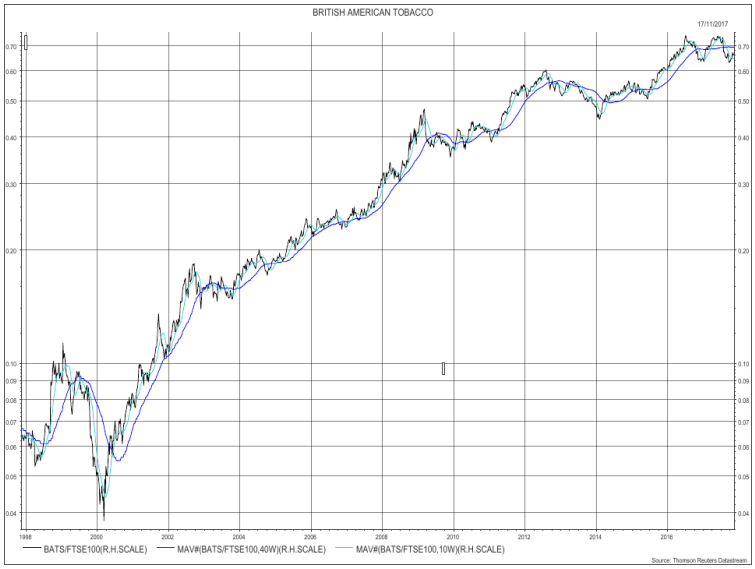
Bad efficacy: BP

- ✗ No long-term performance
- ✗ Violent sell-offs
- ✗ Volatile



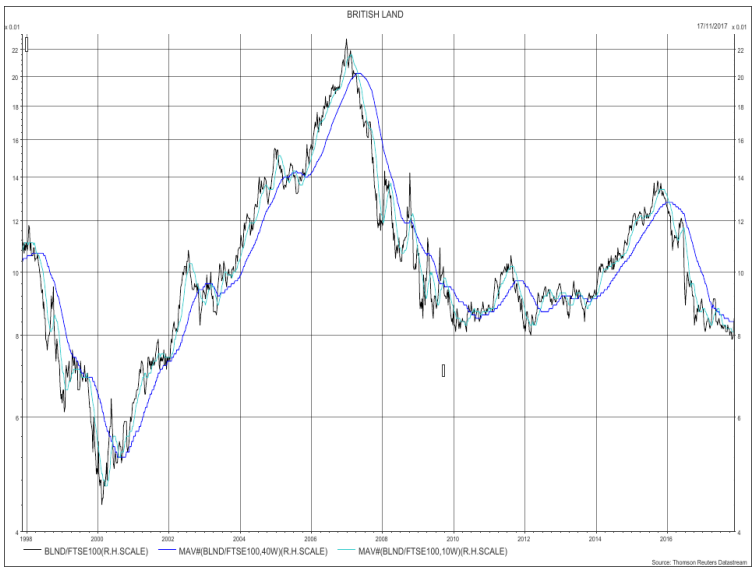
Good efficacy: BRITISHAMERICANTOBACCO

- ✓ Good long-term performance
- ✓ Low vol vs trend



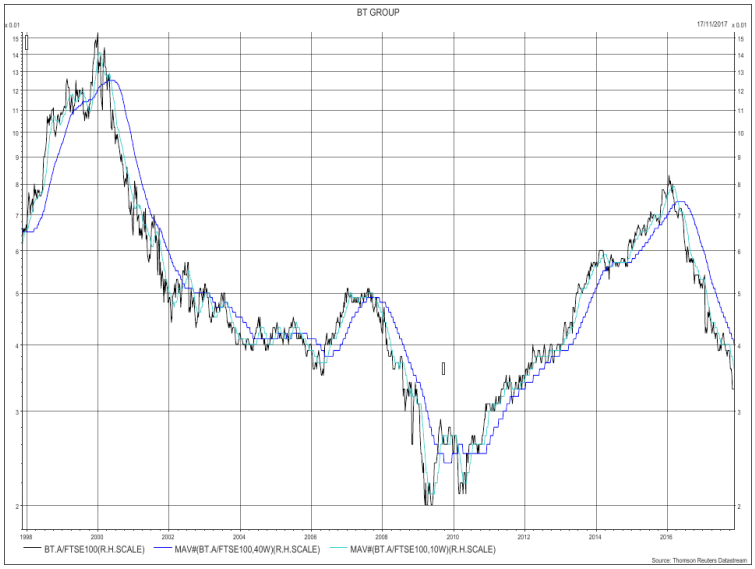
Bad efficacy: BRITISHLAND

- ✗ No long-term performance
- ✓ Trends over extended period
- ✗ Violent selloffs.



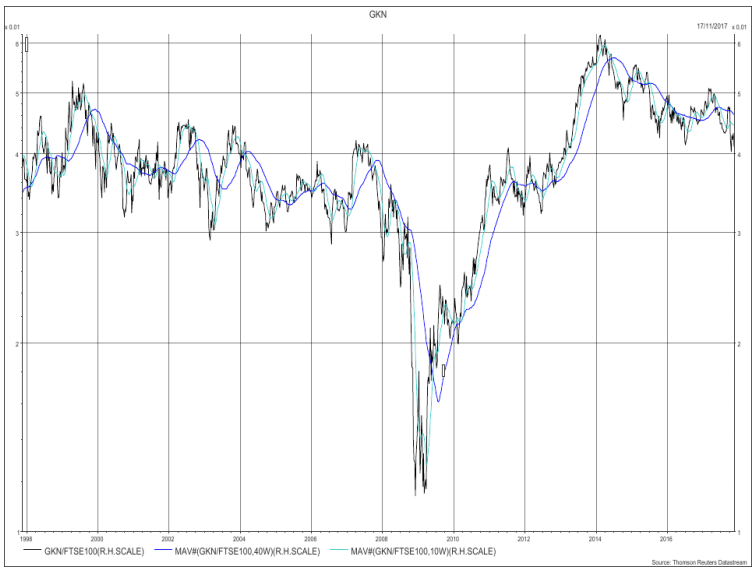
Bad efficacy: BTGROUP

- ✗ No long-term performance
- ✓ Trends over extended period
- ✗ Violent selloffs.



Bad efficacy: GKN

- ✗ No long-term performance
- ✗ Violent selloffs.



3 Examples: stacking

Bad stack: ADMIRALGROUP

✗ 40w MA ¿ 10w MA

Good stack: ANTOFAGASTA

✓ 40w MA ¡ 10w MA

✗ 10w MA roughly equal to Price

✓ stack was bad but recovered

- ✓ 40w MA ¡ 10w MA
- ✓ 10w MA ¡ Price
- ✓ stack was bad but recovered
- ✗ 10w MA not close to 40w MA, potential mean reversion

4 Efficacy indicators

To construct an automated filter for "efficacious" stock charts we can use one of 2 appoaches:

- construct a quantitative metric that closely tracks a human's subjective judgement of the same chart

  Advantage We know exactly what we are measuring

  Drawback Performance usually lags more modern techniques

  Drawback Features are hard to construct

- train a classifier on a large number of syntetic examples of "efficacious" and "non-efficacious" stock charts and then use it to classify real charts

  Advantage Easier to construct

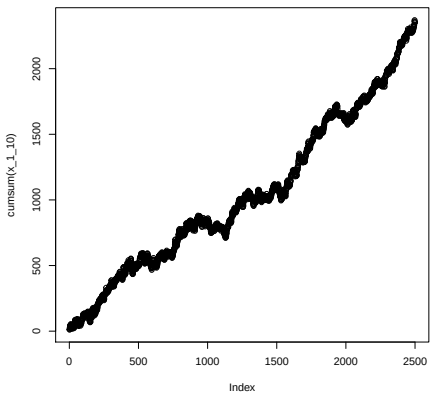
  Advantage Known to perform better than hand-crafter features

  Drawback "Black box" we wont know the exact reason for a classification decision

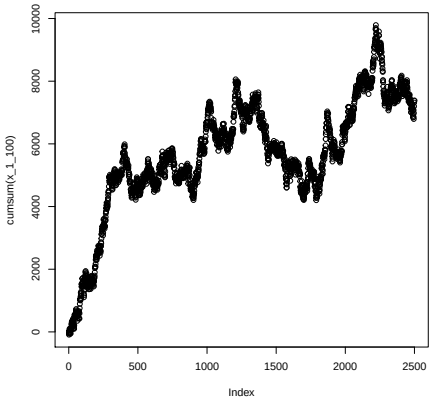
Designing metrics that are effective chart classifiers is challenging. Consider, for example, the sharpe ratio. It is very popular because of its simplicity. However, it has serious problems discrimination between very different performance profiles, as can be seen in the below (highly contrived) counter-example:

Chart classification example: Sharpe vs Calmar ratio

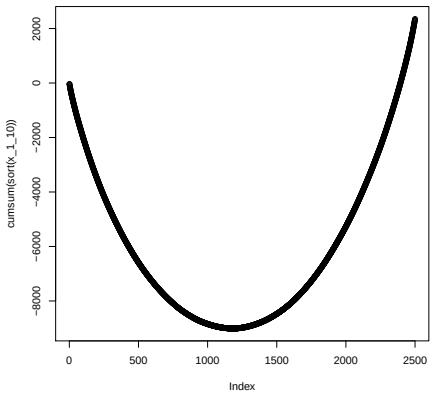
Sharpe 1.48
Calmar 0.138



Sharpe 0.48
Calmar 0.019



Sharpe 1.48
Calmar 0.003



Sharpe 0.48
Calmar 0.001