Statistical Arbitrage

Ingredients of statarb model

- Alpha Model - forecasting
- Risk Model
- Trading Cost Model - estimating slippage of your trades
- Portfolio Construction
- Execution Algorithms - slicing VWAP/TWAP

- Attribution analysis

Alpha Model (1)

Alpha Model (aka Forecast Model) attempts to forecast returns for all the instruments in your trading universe.

The common approach to the problem is to identify reasons for possible inefficiency and come up with a set of features that allow to forecast residual/excess returns.

Philosophically there are few reasons for market inefficiencies:

- Cost of processing new information. Different market participants respond to the new information with different rate.
- Liquidity events. (Providing service)

Alpha Model (2)

Sources of underlying data:

- Price/Volume data (TAQ “Trades and quotes US” - dataset)
- Fundamental Data:

  EPS estimates by analyst I/B/E/S (Institutional Broker’s Estimate System) (Reuters)

  Bloomberg

  Actual earnings reported by companies

- Alternative data (for example foot traffic captured from mobile phones)

Alpha Model (3)

Data Sources:

- TAQ (Price/Volume data) https://www.nyse.com/market-data/historical/daily-taq
- I/B/E/S data https://wrds-web.wharton.upenn.edu/wrds/demo/demoform_ibes.cfm
- date,time,EPS,fiscal_year
- Analyst recommendations
- Actual earnings

Risk Models (1)

Different approaches to risk models:

- Statistical (eg. CAPM model)

  $E(Ri) = Rf + Bi * (E(Rm) - Rf) + E(Alpha), \quad Rf = 0$

  $Bi = Cov(Ri, Rm) / Var(Rm)$

- Fundamental (BARRA model - industry standard)
- Macro

Risk Models (2)

Risk Model is used in portfolio construction and possibly in Alpha Model.

r(i,t) = Sum(A(i,k)*m(k,t), k = 1...NR) + e(i,t)

A(i, k) - factor loadings

m(k,t) - factor return at time “t”

“i” - stock label, “ t” - time

i=1,..., N (N - number of stocks in trading universe)

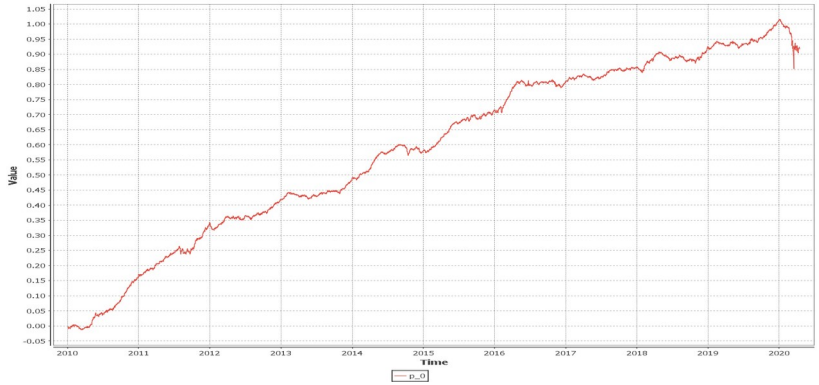
NR - number of risk factors

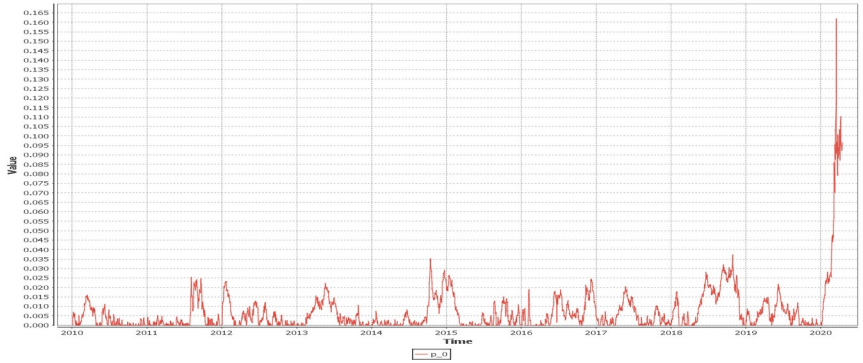
r(i,t) - return of stock “i” at time “t”, e(i,t) - residual

Alpha Factors versus Risk Factors

The distinction between Alpha factors and Risk Factors is conditional on our ability to forecast certain directions in the return space.

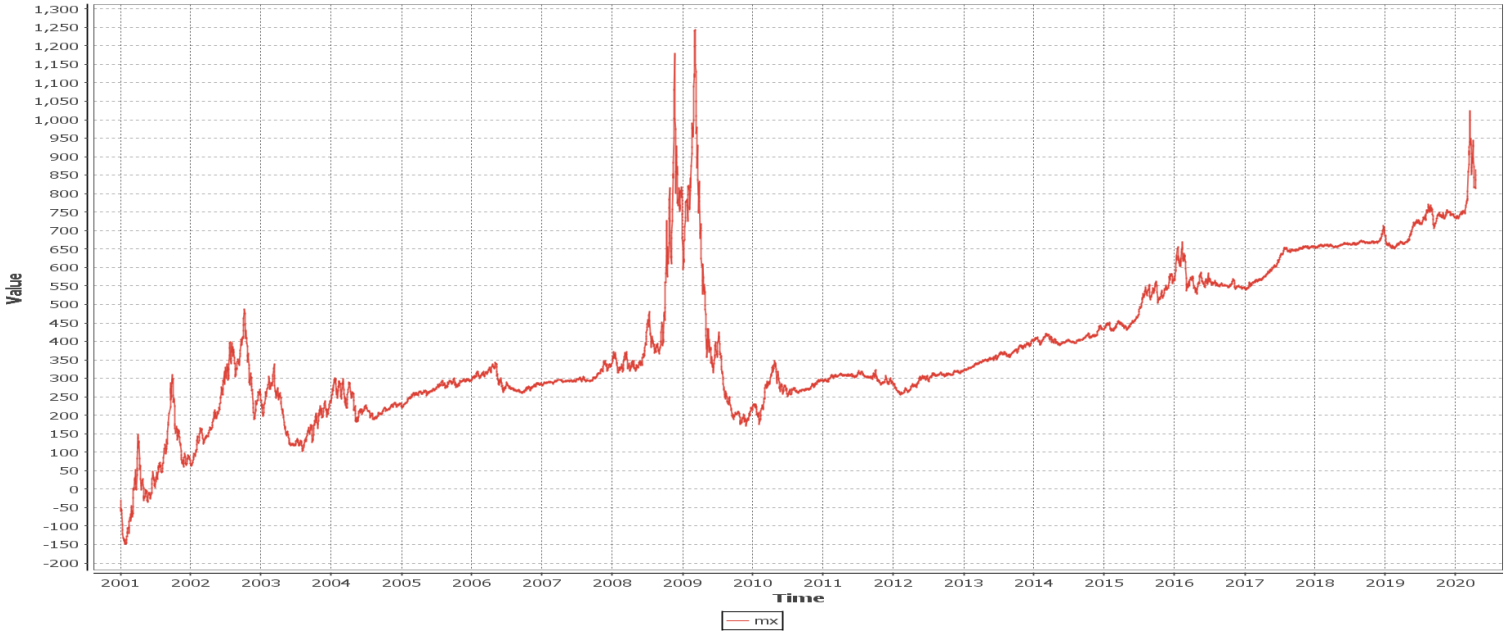
Example Alpha Factor (short term reversion):





Example of Risk Factor Performance

1 Year Price Momentum



Trading Cost Model (1)

Definition of Trading Cost

Notation:

n - number of shares to trade (sign > 0 for buy order, sign <0 for sell order)

E - entry price when the decision to trade was made

F - average fill price

Assume we are trading during time interval t=1...T

VWAP = Sum (trade_size(t)*P(t) , t=1….T) / Sum(trade_size(t), t=1...T)

Cost2Entry = n*(F-P)

Spread Crossing = n*(F-VWAP)

Trading Cost Model (2)

Spread Crossing = SC * RelativeSpread * dollarsTraded

MarketImpact = MI * f(dollarsTraded / dollarsTradedInTheMarket) * dollarsTraded

SC - spread crossing coefficient

MI - market impact coefficient

ParticipationRatio = dollarsTrade/dollarsTradedInTheMarket

$f(x) = x, \quad x < 0.01$

f(x) = sqrt(x) for large “x”