Getting Started with Blueshift

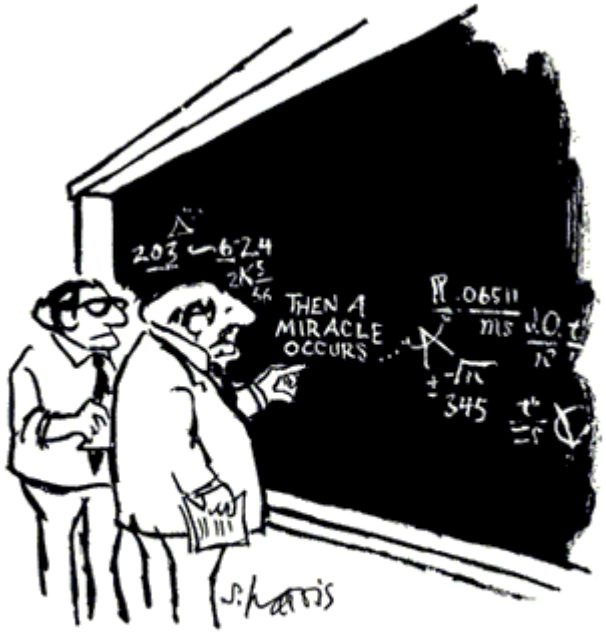
Introduction to Quantra Blueshift Platform

blueshift

blueshift.quantinsti.com

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Some Definitions



"I think you should be more explicit here in step two."

Systematic Trading: **Trading decisions** in a methodical way defining trading goals, risk controls and rules

Quantitative Trading: Trading **signal generation** based on quantitative analysis

Algorithmic Trading: **Execution of trades** using automated pre-programmed trading instructions

Hi-Frequency Trading: Algorithmic trading characterized by **high speed**, high turn-over rates and high order-to-trade

Three Degrees of Freedom

- How trading and or other decisions are **evaluated** (Quant or Qualitative signal generation)
- How trading and other decisions are **finalized** (Rule-based or discretionary decisions/ order generations)
- How those trades are **executed** (Automated or manual)

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

What is Blueshift

Systematic Trading Strategies – For Everyone!

IP protection

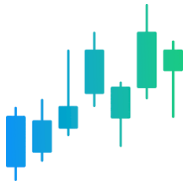
Strategy research

Financial datasets

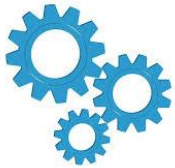
blueshift

Learning and community

On the cloud



Live trading



Back-testing

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Trading Strategy

Strategy: A set of rules or a well-defined algorithm to choose an action from the set of all possible actions in any state of a game.

Trading Strategy: Strategy (above). The game is stochastic (state transitions are probabilistic), massively multi-player and potentially long running

The Game Plan

1. Usually the number of participants is large, hence we create the concept of "market" (unlike smaller games like auctions or business negotiations)
2. Essentially the market manifests itself in terms of the tradable assets (what you can trade), their prices (at what levels) and the liquidity (and how much)
3. We make assumptions to understand how state changes happen - A set of theories from Economics that maps real-world to trading conditions and another set of theories on how the market process information to reflect changes in those conditions
4. The first set of theories are in general known as asset pricing theories
5. The second group of theories talks about market efficiency and behavioural economics

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

The Strategy Spectrum

Quant	Momentum (time-series or cross-sectional)	Pair-trading, most types of statistical arbitrage	Advanced models (e.g. HMM, regime switching)	HF Market-making, Cash-futures arbitrage	News-based automated trading
Technical	MA cross-over, Continuation patterns	Swing, Retracement, Pivot trading	Opening range, dual thrusts, patterns	Range-based short gamma (vol selling)	Nothing much here
Fundamental	Factor-based investing	value investing	value/ RV (relative value) strategies	Cross-asset, cross country RV/ short gamma	Usually discretionary
	Trending	Mean-reverting	Break-out	Carry	Event-based

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

The Power of Large Numbers

Tale of Two Styles

Asset manager A is a discretionary manager and follows 10 stocks diligently. Asset manager B is a systematic manager and have enough computational power to track 500 stocks. Investors seek outperformance and confidence.

Question #1: What is the hit ratio for A to achieve outperformance in at least 50% (total 5 stocks) of his portfolio with at least 95% probability

Question #2: What is the hit ratio for B to achieve outperformance in at least 50% (total 250 stocks) of his portfolio with at least 95% probability

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

The Power of Large Numbers

Tale of Two Styles

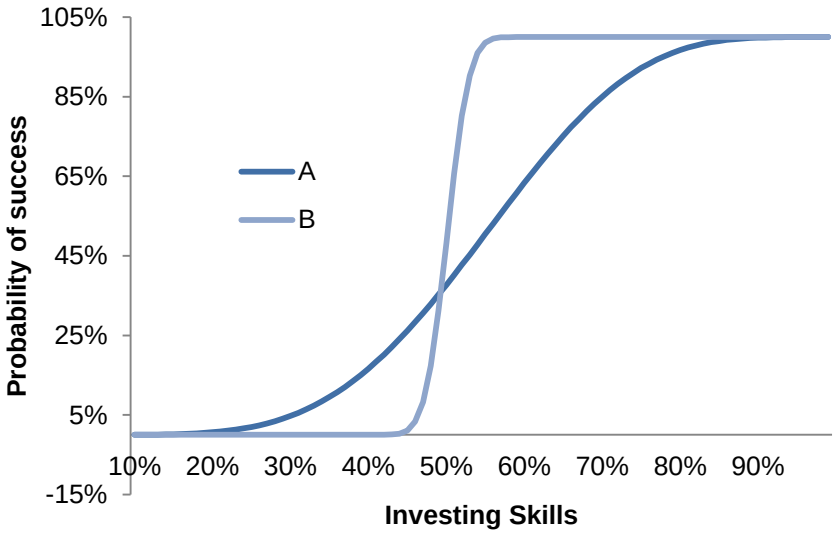
Solution: Let the success rate be p. Total required success n, total trial N. Probability of n success out of N trials is the binomial distribution

$$\Pr(n \mid N, p) = {}^{N}C_{n}.p^{n}.(1-p)^{N-n}$$

Probability of at least n success

$$P = 1 - \sum_{i=1}^{n} \Pr(i \mid N, p)$$

Compute the probability (1 - BINOM.DIST in Excel, with cumulative True). Solve for p, such that P is 0.95 in both cases.



CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Fundamental Law

The fundamental law of investment management: $IR \approx IC \cdot \sqrt{N} \cdot TC$

Information Ratio (IR)

Measure of success – risk adjusted returns over the chosen benchmark

Information Coefficient (IC)

Linear relationship between predicted asset returns and realized returns, usually in the range of 0 to 1 (unlike correlation range of -1 to 1, as you can always short the prediction of someone who is consistently wrong!!).

Number of Bets (N)

Number of **independent** bets that can be made in a given time-frame.

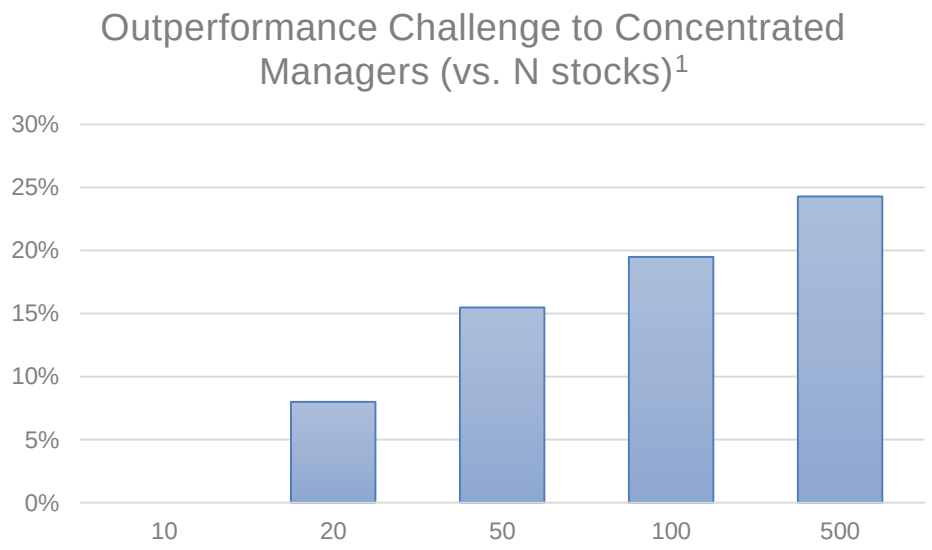
Transfer Coefficient (TC)

The linear relationship between predicted asset returns and actual trading positions. This ideally should be one, but can be lower (risk limits, capital limit, bandwidth problem due to work overload etc.)

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Systematic Fundamental?

The fundamental law of investment management: $IR \approx IC \cdot \sqrt{N} \cdot TC$

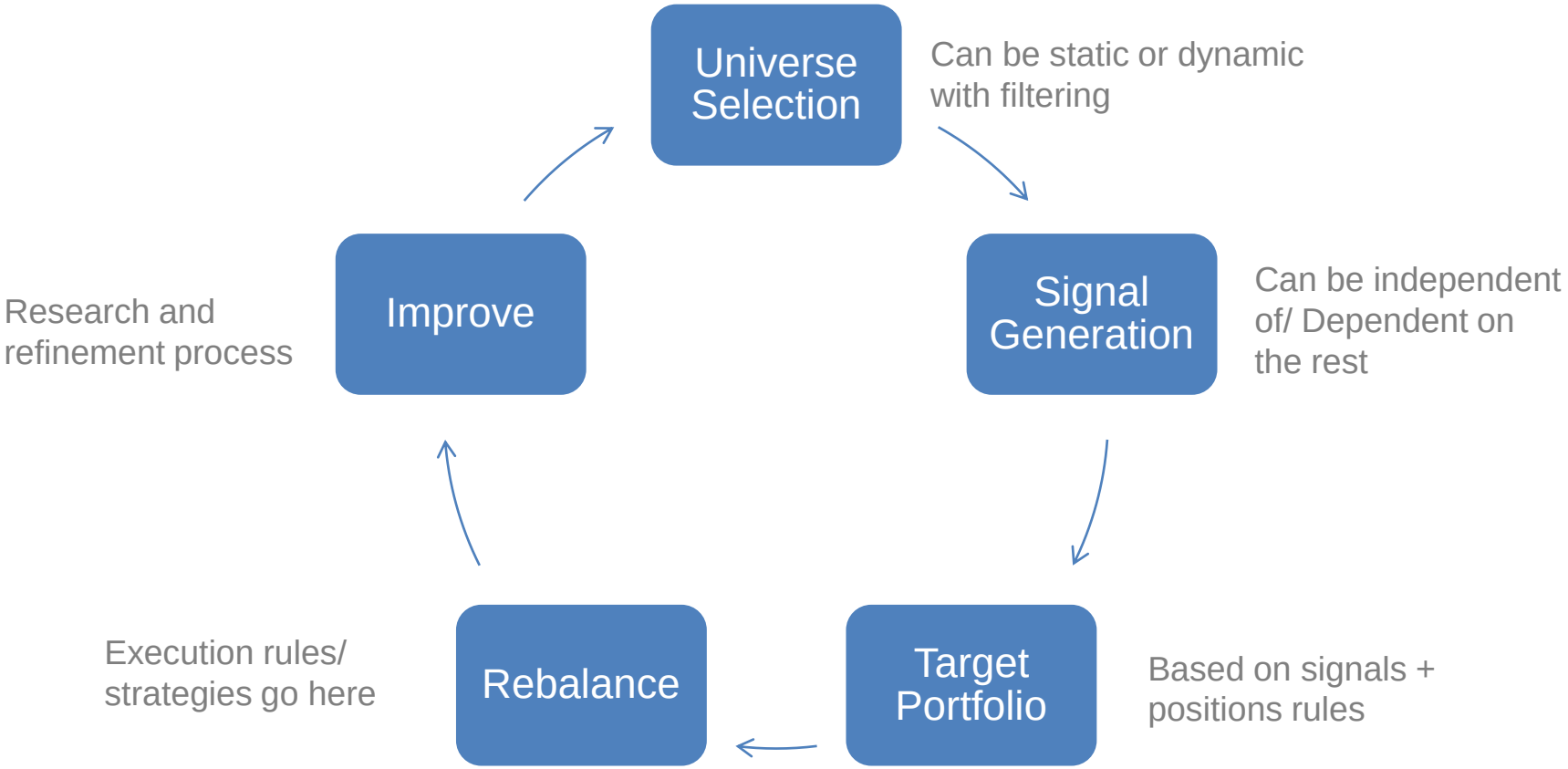


Discretionary	Systematic/ Quant
Small N (usually)	Large N (usually)
Edge is deep, proprietary insights	Edge is superior info processing
Ideally offers superior upside	Ideally offers superior consistency

1. Based on in-house computation, for outperformance in at least half of the portfolio with 95% probability

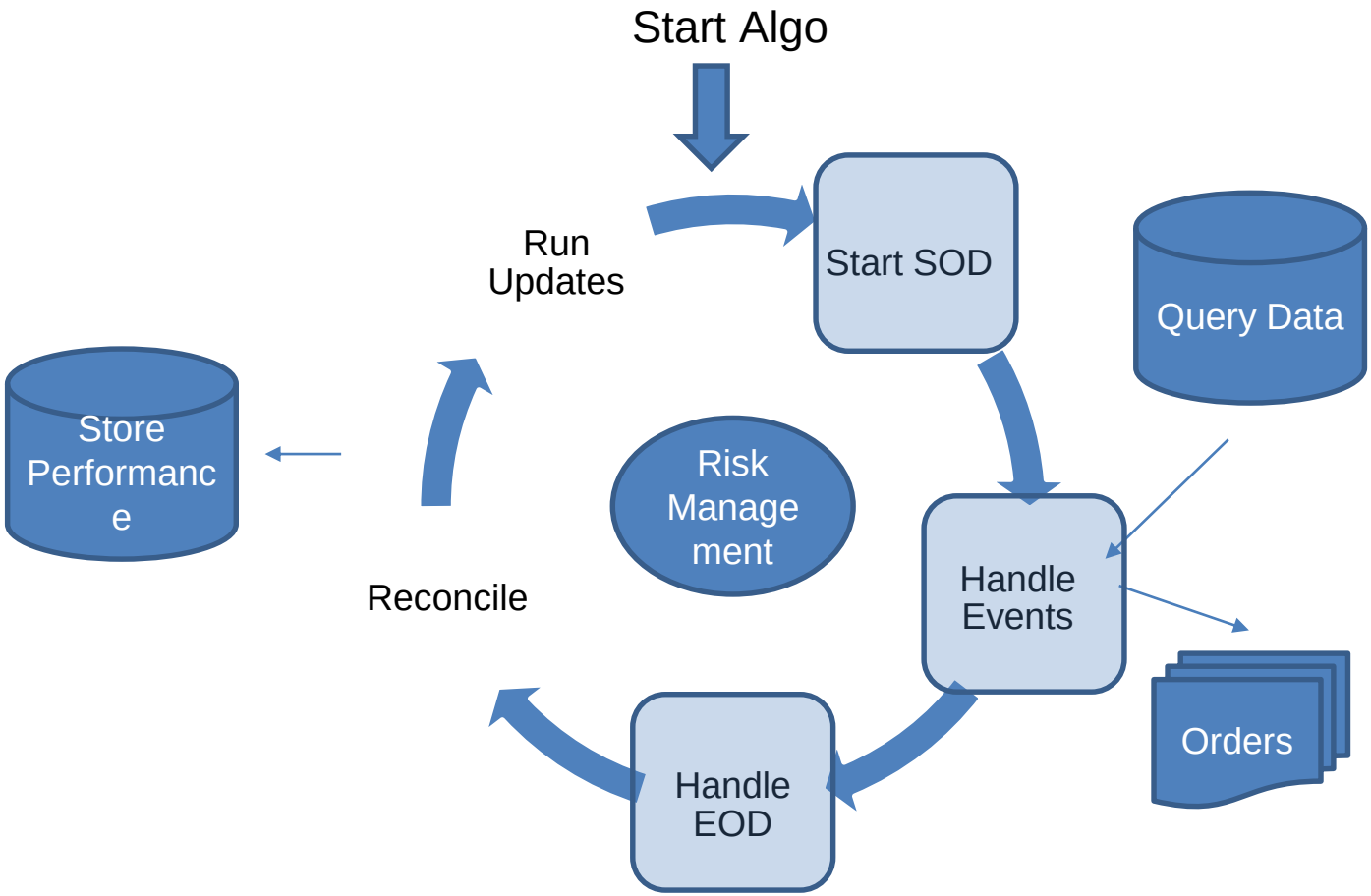
CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Systematic Strategy Design Cycle



CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Trading System from Users POV



A trading algorithm is essentially a process that consumes data and generates orders. Other things are additional but very important paraphernalia (risk and capital management, performance tracking etc.)

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Systematic Strategy Development

- **Inputs**: Input to the strategies can come in many flavours[1] like
  - Price/ returns and it's transformation. Most of the common technical indicators are transformation of price returns[2]
  - Positioning information – volumes and open interest data and participant-wise positioning data if available
  - Fundamental information –macro-economic information, company fundamentals
  - Non-market information: Example twitter sentiments, analysts ratings etc.
- **Trading Rules/ Logic**: Can be either based on trader's hypothesis or inferred (learned) from data
  - Form hypothesis (e.g. moving average cross-over signals change in trends) and test – traditional approach
  - Feed data to infer rules (the subject matter of machine learning and artificial intelligence)
- **Back-testing and Forward-testing**: A crucial step and often over-looked step.
  - A good platform to guard against biases : data-mining bias, survivorship bias, market-impact modeling, look-ahead bias
  - Should be flexible, event-driven (to avoid look-ahead bias) and with built-in analytics
- **Portfolio Creation/ Optimization**: Never go all-in with a single strategy
  - Two strategies better than one – if they are uncorrelated (in terms of signals and/ or performance)
  - Various methods exists – ensemble, traditional portfolio theory, dynamic portfolio allocation etc.

1. Ideally sources should be disparate, uncorrelated and relevant.
2. For further technical details see here.

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

"Unfortunately for the case of sanity, that (momentum investing) seems to be true."

Cliff Asness (AQR Capital)

Blueshift Work Flow

Developing Strategies

blueshift

blueshift.quantinsti.com

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Data-set on Blueshift

Equities market data for NYSE (the US) and NSE (India) – minute levels with corporate actions. Updated once every day after market close.

FX data - 10 currency pairs - AUD/USD, EUR/CHF, EUR/JPY, EUR/USD, GBP/JPY, GBP/USD, NZD/USD, USD/CAD, USD/CHF, USD/JPY. Minute-level data since 2008. Updated once every day.

This is augmented by Macro data on 8 underlying economies (real GDP, Y-o-Y core inflation rate, short term (3-month) and long term (10-year) interest rates.

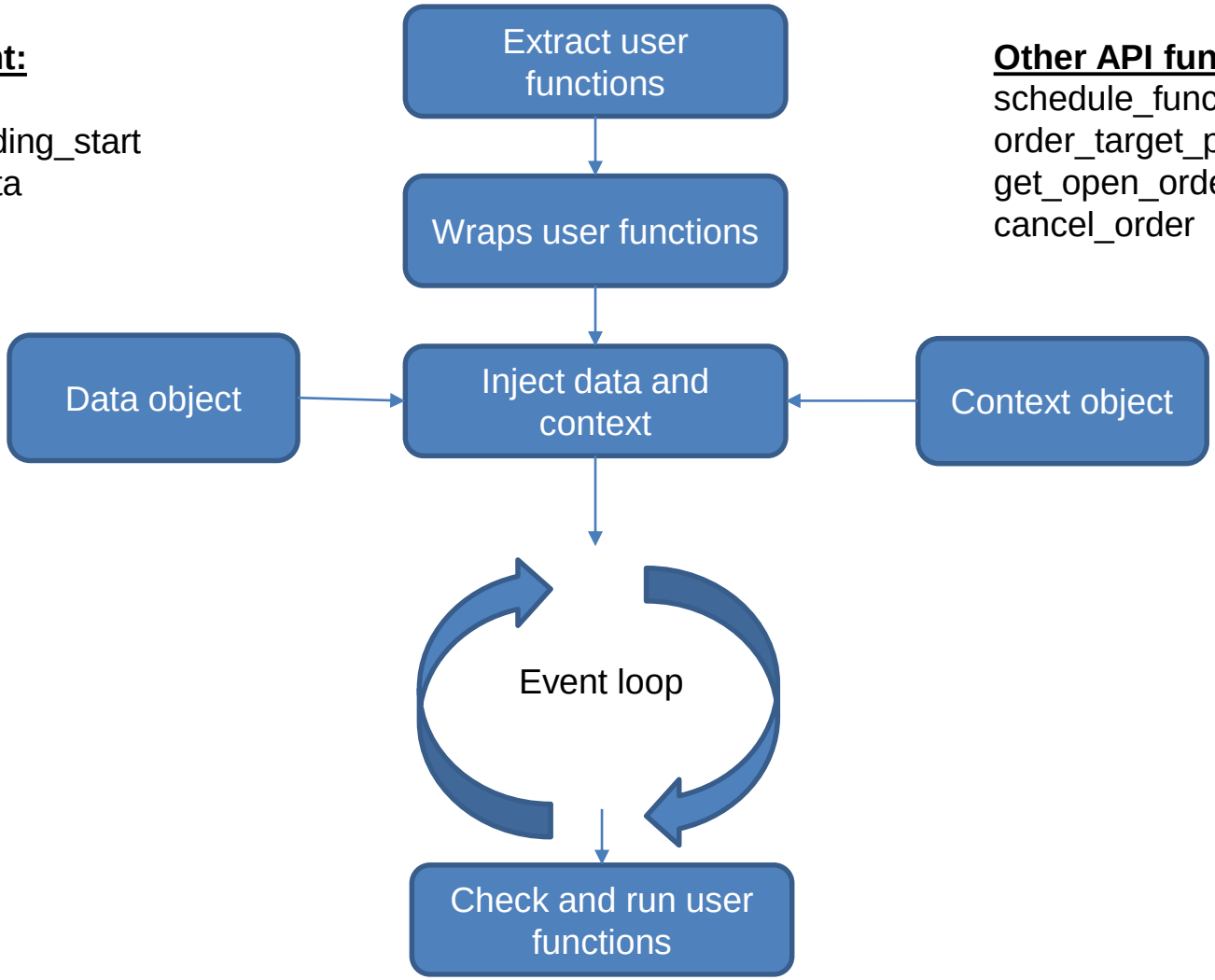
Price Data from the `data` argument passed. Two methods available – data.current() and data.history(). Macro data can be imported as python module, and have same functions (with slightly different arguments).

How to read data on Blueshift: Example as follows

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

How Blueshift API works

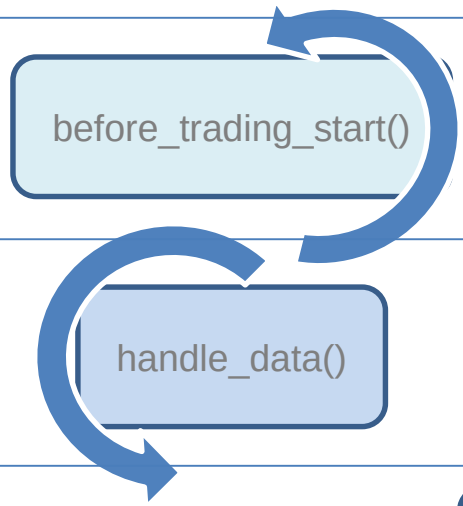
Entry Point:
Initialize
before_trading_start
handle_data
analyze



Other API functions
schedule_function
order_target_percent
get_open_orders
cancel_order

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Work-Flow on Blueshift



Called once at the start of the algorithm
Place to set up algorithm parameters and initialize them
Arguments: context – a dictionary object to store all user parameters and variables and available to all other functions

Called once at the start of the trading session (i.e. once a day)
Place to set up daily pre-open stuff
Arguments: context – as above, and data – a object with current date-time, can be queried for current or historical data on symbols

Called once at every bar (every minute)
Place to set up logic on how to respond to market data
Arguments: context – as above, and data – as above

Other Useful API Functions

schedule_function() date_rules() time_rules()	order_percent_target() get_open_orders() cancel_order()	set_commission() set_slippage() set_account_currency()
Schedule any user-defined functions at a given time and frequency	Functions to place, and cancel orders	Functions to set trading and other parameters

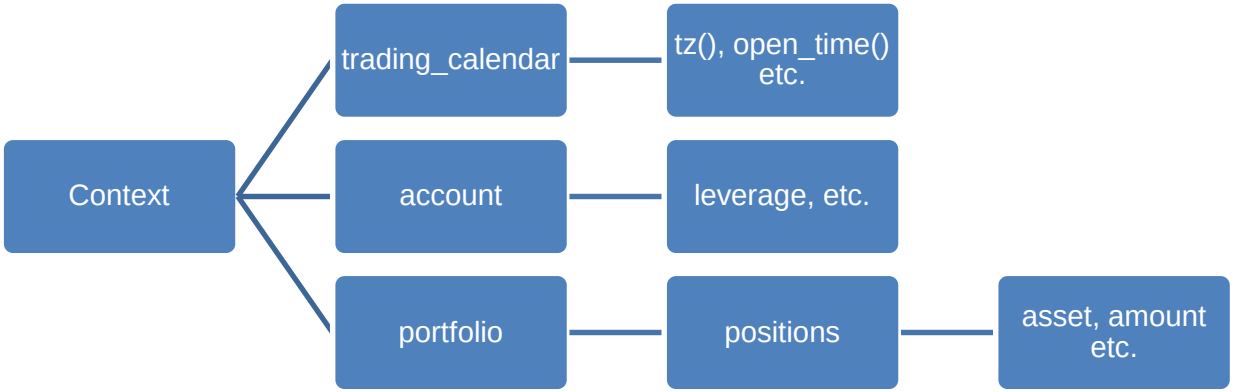
CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Blueshift Environment

Data contains everything needed to feed data: current(), history(), as well as current_dt

Context contains everything else: Portfolio (context.portfolio), account (context.account), calendar (trading_calendar)

*See **here** for more details*



CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Blueshift Workflow Demo

1. *Signing up and Platform Tour*
2. *Creating Strategy*
3. *Running Back-test with a template*
4. *Adding some library functions*
5. *Using our library*

blueshift.quantinsti.com

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

A Simple Strategy

Buy (sell): if close near lower (upper) bands
Buy (sell): if close away from extremes and positive (negative) momentum

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

"A rising market and a long position is a sure sign of investment genius"

John Kenneth Galbraith (Economist)

Alpha Strategies

Understanding Source of Risks and Returns

blueshift

blueshift.quantinsti.com

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

How The Market Works!

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Asset Pricing: Basics

Fundamental Value of An Asset : The value of the asset such that it gives owner a **reasonable profit**[1]

Market Price: The price at which (a large number of) players **can agree** to buy or sell an asset.

Question: What is **the fundamental Value** of any thing? Expected Pay-off?

A Simple Game of Luck[2]:

You pay $X dollars to play a game against the “house”. The game starts with $1 in the pot. At each turn the house toss a fair coin. If it is tails, you get whatever in the pot, else we move to the next turn with the house doubling the pot amount (i.e. $4, then $8 then $16 and so on).

What is X. How much you will pay to play this game?

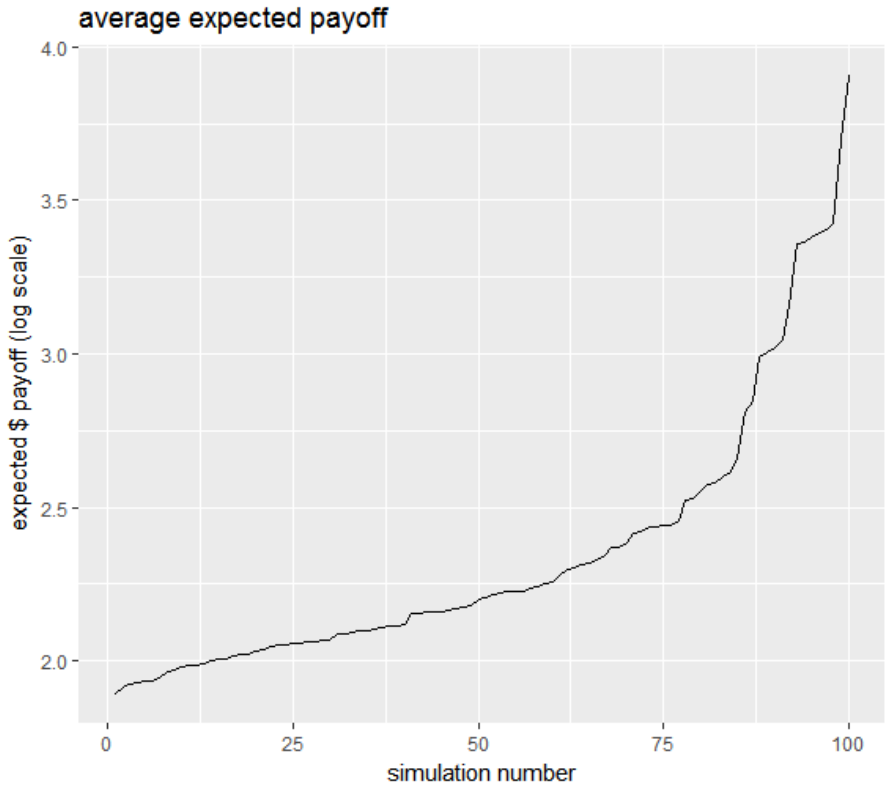
1. In line with the definition from Adam Smith, Wealth of Nations (1776)
2. Also known as St. Petersburg paradox

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Asset Pricing: Paradox

Question: What is **the fundamental Value** of any thing? Expected Pay-off?

Unbounded Expectation!!



1. 100 runs of simulations with 50K tosses each

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Asset Pricing: Concepts

Utility: the pleasure or satisfaction derived. A mathematical concept (a function) to arrange a set of choices in a given time and/ or over time in order of preference.

Value of anything is the present expected utility (Expected Utility Hypothesis)[1]
Utility brings in to the picture the risks preference as well as time preference

Present Value Models: Fundamental value is **Net Present Value** (NPV) or **Discounted Cash Flow** (DCF) appropriately adjusted for costs (including opportunity costs) and risks.

1. Time Value of Money: People prefer now than later
2. Risk Aversion: People prefer less risky option when the pay off is same
3. Assumes functional capital market (lending and borrowing)

Costs: Usually zero for financial assets
Risk Free Rate: Captures the opportunity costs (in a sense captures temporal preference)
Risk premium: Captures risk aversion – how much investors demand to be compensated for risks in equilibrium
Discount Factors: Include both risk premia and time premia. Depends on individual utility, but in presence of capital market converges to give **"Law of One Price"**

1. The alternative (arguably more realistic) hypothesis is the Prospect Theory
2. Also known as St. Petersburg paradox

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Asset Pricing: Maths

Asset Prices: Payoff under different scenarios (s) multiplied with some "**discount factor**", that takes in to account utility

$$P = \sum_{s=1}^{S} x^s \varphi^s$$

If x = 1 for any scenario s, then the asset is **risk-free asset**. Obviously

$$P^0 = \varphi^0 = \sum_{s=1}^{S} \varphi^s \, , and \; R_f = \frac{1}{\varphi^0}$$

Risk free rate

Alternatively we can write

$$P = \sum_{s=1}^{S} x^s \varphi^s = \varphi^0 \sum_{s=1}^{S} x^s \frac{\varphi^s}{\varphi^0} = \frac{1}{R_f} \sum_{s=1}^{S} x^s q^s$$

Risk neutral pricing

Again, alternatively

$$P = \sum_{s=1}^{S} x^s \frac{q^s}{R_f} = \sum_{s=1}^{S} p^s x^s \left(\frac{q^s}{R_f p^s} \right) = \sum_{s=1}^{S} p^s x^s m^s = \mathbb{E}(mx)$$

Expected "utility" angle

Risk Premium: From basic stats we get

$$P = \mathbb{E}(mx) = cov(m, x) + \mathbb{E}(m)\mathbb{E}(x)$$

Re-arrange this as

$$\mathbb{E}R = R_f \_ R_f cov(m, R), R = \frac{x}{P}$$

Risk premium

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Asset Pricing: Theories

Arbitrage Pricing Theory (APT):

$$\mathbb{E}R = R_f - \sum_{i=1}^{K} \beta_i \left(R_f \mathbb{E}(m.F_i) \right), where\ m = \frac{1}{R_f} + \sum_{i=i}^{K} b_i F_i$$

Here **F** are the factors and beta are the factor exposure of that particular asset

Capital Asset Pricing Model (CAPM)

This assumes there is only one factor, the market exposure: High returns because higher risks!

$$\mathbb{E}R = R_f - \beta.\mathbb{E}R_M$$

Fama French 3-FacorModel

Some other factors than market, like market capitalization (size factor) and book-to-market ratio (value factor) are important too!

Note:

- These models are useful for portfolios (large N, like investor B). We are ignoring idiosyncratic risks based on the assumptions that they can be diversified away!!
- These are relative valuation (relative to factors), NOT absolute valuation (NOT how much is Infosys worth to your, but what the value of average investors put on it)
- Not useful for pricing derivatives, hard to find stable factors (or appropriate discount rates)! Rather price in terms of the underlying instead (risk neutral pricing)

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

What is Alpha

Modern Asset Pricing Theories are essentially as follows:

$$R_{i,t} = a_i + \sum_j \beta_{i,j} \Delta F_{j,t} + \epsilon_{i,t}$$

Alphas will be arbitraged away

idiosyncratic risks not paid for

$$\mathbb{E}(R_{i,}) = R_f + \sum_j \beta_{i,j} RP_j$$

1. *ΔF is k-variate I(0), ε is n-variate white noise (total instruments n, total factors k). RPs are risk premia (k-variate I(0))*
2. *Objective: Create portfolio to beat the market!*
3. *First try: Capture any factor other than market – Jensen's alpha!*
4. *Holy Grail: RP (and ΔF too!) not time-dependent – pure alpha!*
5. *Tough luck: try RP not time-dependent – Beta is the new alpha?*
6. *Next: RP is stable and consistent – Focus on factor allocation!*
7. *Time-varying but predictable? – Focus on market timing!*
8. *Unstable and unpredictable on all scale? – go back to picking stocks!*

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

"Successful investing is anticipating the anticipations of others."

John Maynard Keynes (Economist)

Systematic Equities

A very brief overview

blueshift

blueshift.quantinsti.com

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Momentum: Concept

What is Momentum: Exploiting time series persistence

Strategy-I: Buy (sell) assets that have been going up (down) in prices in the past, hold them in the portfolio for a defined period of time (holding period), and then rebalance. This is known as time-series momentum

Strategy-II: Rank all assets based on their past returns. Buy the top x-percentile and sell the bottom x-percentile (even if they are going up in prices). Rebalance after the holding period

Systematic Asset Management Before Momentum:

1. Passive - Portfolio Optimization: $\min(w^T.\Sigma.w - q.R^T.w), s.t. q \geq 0, \sum_i w_i = 1$
2. Active - Market Timing and signal generation (digital signal processing, technical analysis etc.)

Systematic Asset Management after: Factor investing:

1. Passive - Portfolio Optimization (mean-variance)
2. Passive – Portfolio Optimization (risk factors optimization)
3. Active - Market Timing and signal generation (digital signal processing, technical analysis etc.)

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Momentum: Everywhere

Equities: classic cross-sectional and time-series momentum (we have just seen it)

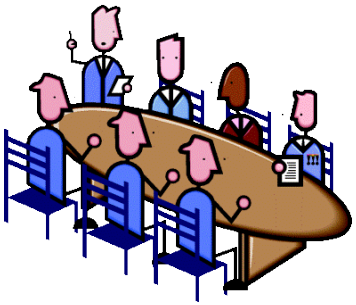
Fixed income: Central banks tend to hike (cut) rates in steps, leading to momentum effect on short term rates. Also economies undergoes business cycles that get reflected in rates and inflations – leading to trends in long term rates

Commodities: One of the first markets to trade momentum (pioneer of technical analysis). A strong drivers are business cycles and persistence of under or over capacity (setting up factories require time). Tend to mean revert over long period.

Foreign Exchange: A large part of momentum is driven by the same factors as rates.

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Momentum: Drivers



Behavioural

Under-Reaction: Slow reactions to company specific or macro-economic news (the opposing theory is known as Efficient Market Hypothesis or EMH). Conservative behaviour, bounded rationality, liquidity constraints etc.

Delayed Over-Reaction: Behavoiral feedback mechanism. This also gets reinforced by under-reaction

Disposition Effect: Investors tend to stick to losers too long and book profit too quick. Creates momentum.

Risk-based

Risk Compensation: Market value of firm growing faster than fundamental increases the risk, hence momentum profit should compensate. Momentum is usually stronger for companies with higher growth opportunities and riskier cash flows

Correlation Structure: Momentum assets strongly correlate with each other, which implies a common factor which can experience changes in future (**momentum crash**)

Liquidity Risks: Recent winners have greater exposure to liquidity risks (everyone tries to exit at the same time)

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Blueshift Pipelines

Pipelines are efficient ways to compute a factor on a large universe

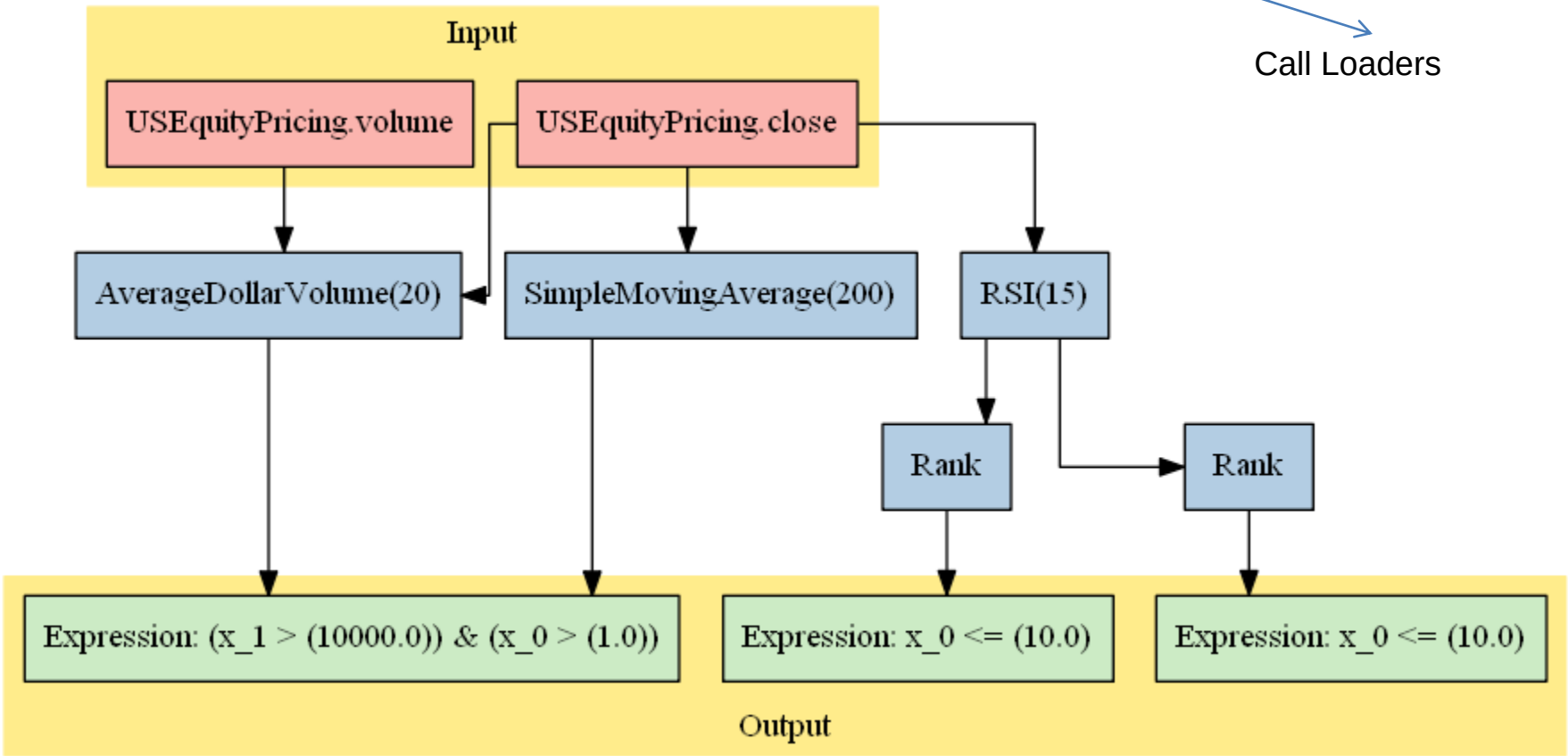
We attach a pipeline at the onset (initialize). This is attach_pipeline()
Then call the results periodically: pipeline_output(). This returns usually a pandas dataframe with multi-index (indexed by stocks and then timestamps).

To define pipeline output, we define the factor. This is basically defining a class with a `compute` function that must return the factor we are interested in. It takes a `input` (usually the price), a `window length` to compute, and all assets in the database. It returns the computed value

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Blueshift Pipelines

Define factors → Compute Dependencies → Create Graph → Topological Sort → Call Loaders



Screen

Columns

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

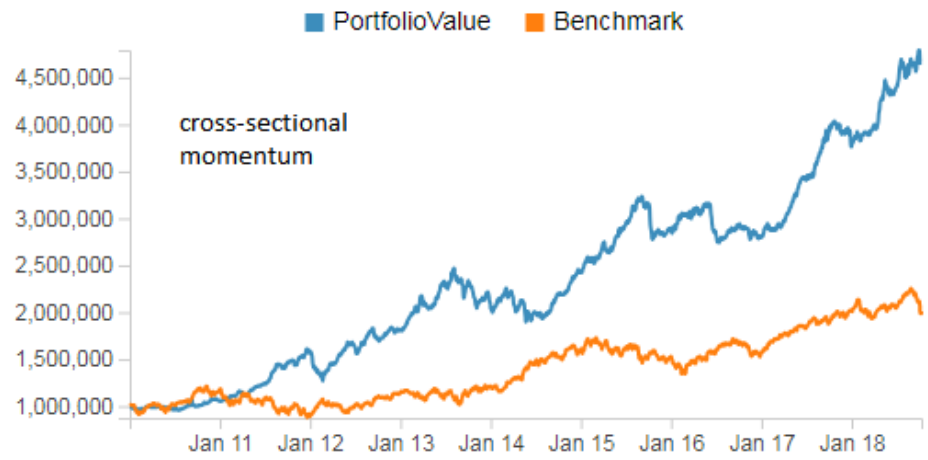
Blueshift Workflow Demo

1. *Create a library for momentum factors*
2. *Import that to develop a cross-sectional strategy*
3. *Use same library to develop a time-series strategy*
4. *Compare results*

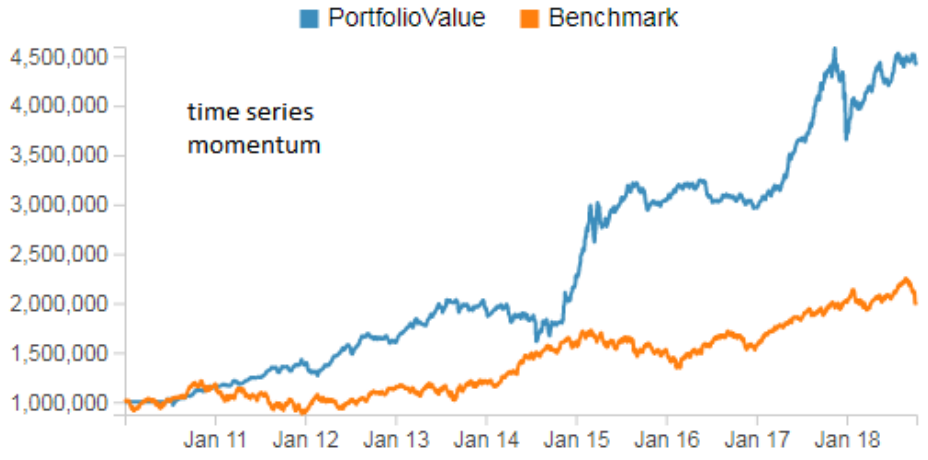
CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Momentum: Performance

RETURNS	ALPHA	BETA	SHARPE	DRAWDOWN
366.54%	0.19	-0.14	1.55	-24.06%



RETURNS	ALPHA	BETA	SHARPE	DRAWDOWN
344.00%	0.17	-0.04	1.53	-21.89%



1. The side-ways markets during 2011-2013 saw better performance from XMOM
2. Strong post-election rally was better captured in TSMOM
3. XMOM had a shorter “flat” period post election rally
4. In general, “beta” is higher for TSMOM (especially for absolute beta)
5. Overall performance, including alpha, Sharpe ratio, drawdowns are similar

Let’s see the demos

1. Based on NSE 500 stock universe, with look back period of 12 months and rebalance period of 1 month

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

“Trial and error is freedom.”

Nassim Nicholas Taleb

Statistical Arbitrage

Statistics, and Practice

blueshift

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Stat Arb: Concepts

- Choose a pair of stocks that move together very closely, based on a certain criteria (i.e. TCS & Infosys)
- Wait until the prices diverge beyond a certain threshold, then short the "winner" and buy the "loser" (contrarian approach)
- Reverse your positions when the two prices converge --> Profit from the reversal in trend
- Pairs trading strategy requires few key decision points:
  - **Pair (or basket) Selection Algorithm**
  - **Market Timing (Entry and Exist)**
  - **Optimal Position Sizing Mechanism**
  - **Spotting the Scarce Opportunity**

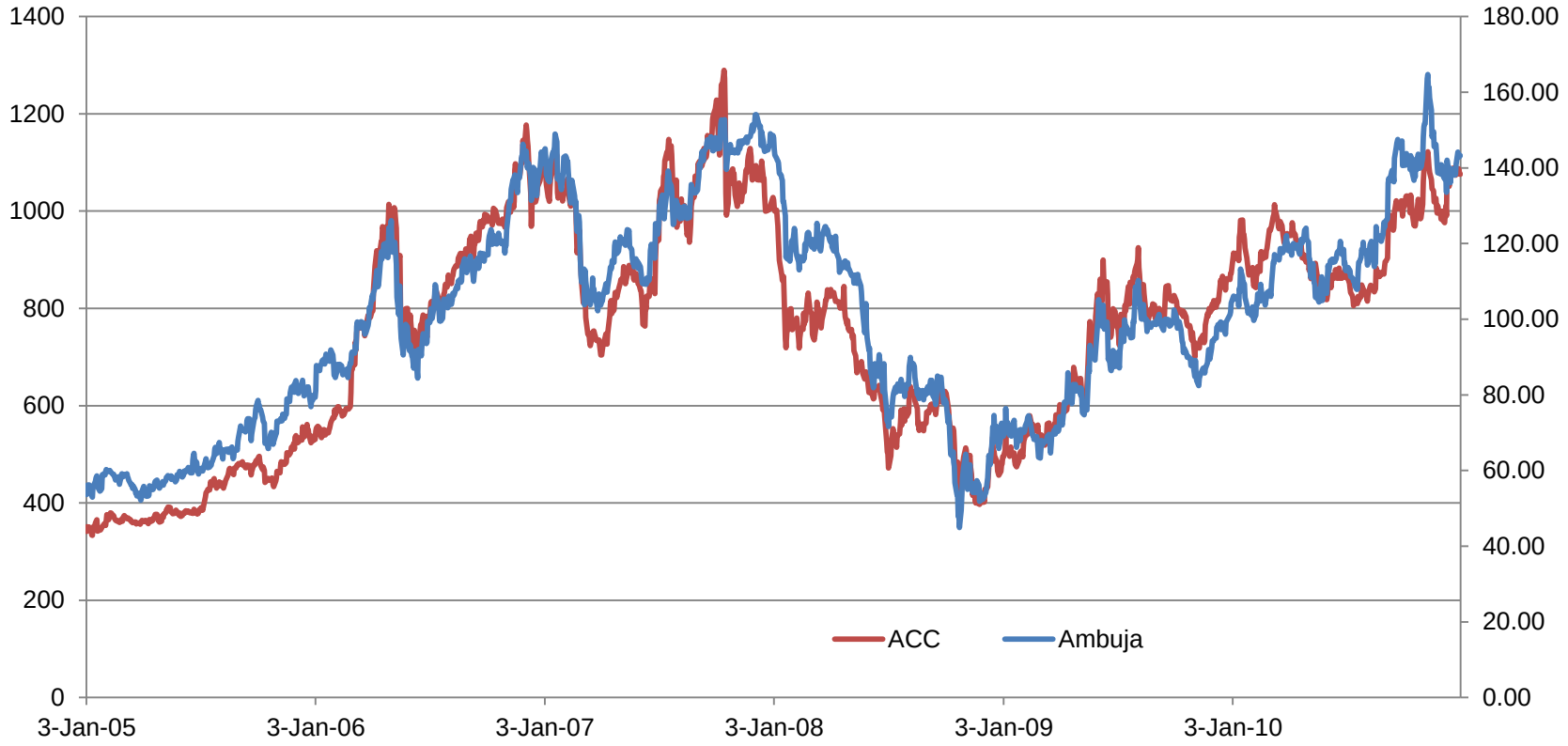
CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Stat Arb: Risks

- Statistical Arbitrage is not a risk-free strategy. Key risk involved in trading are as follows:
  - Rather than converging, prices of the two securities can begin to diverge (drift apart), i.e. the spread picks up trend rather than mean-reverting to the original mean and the cointegration can be broken.
  - Sometime, an event in a security can trigger extreme movement in pair ratio. Entering into such a security trade should be avoided.
  - Strict risk management of Stop Loss is required to handle adverse situations once the mean-reverting behavior is invalidated.
  - Inherently stat-arbs are like short gamma trades. They tend to win most of the time and sometimes blow up!

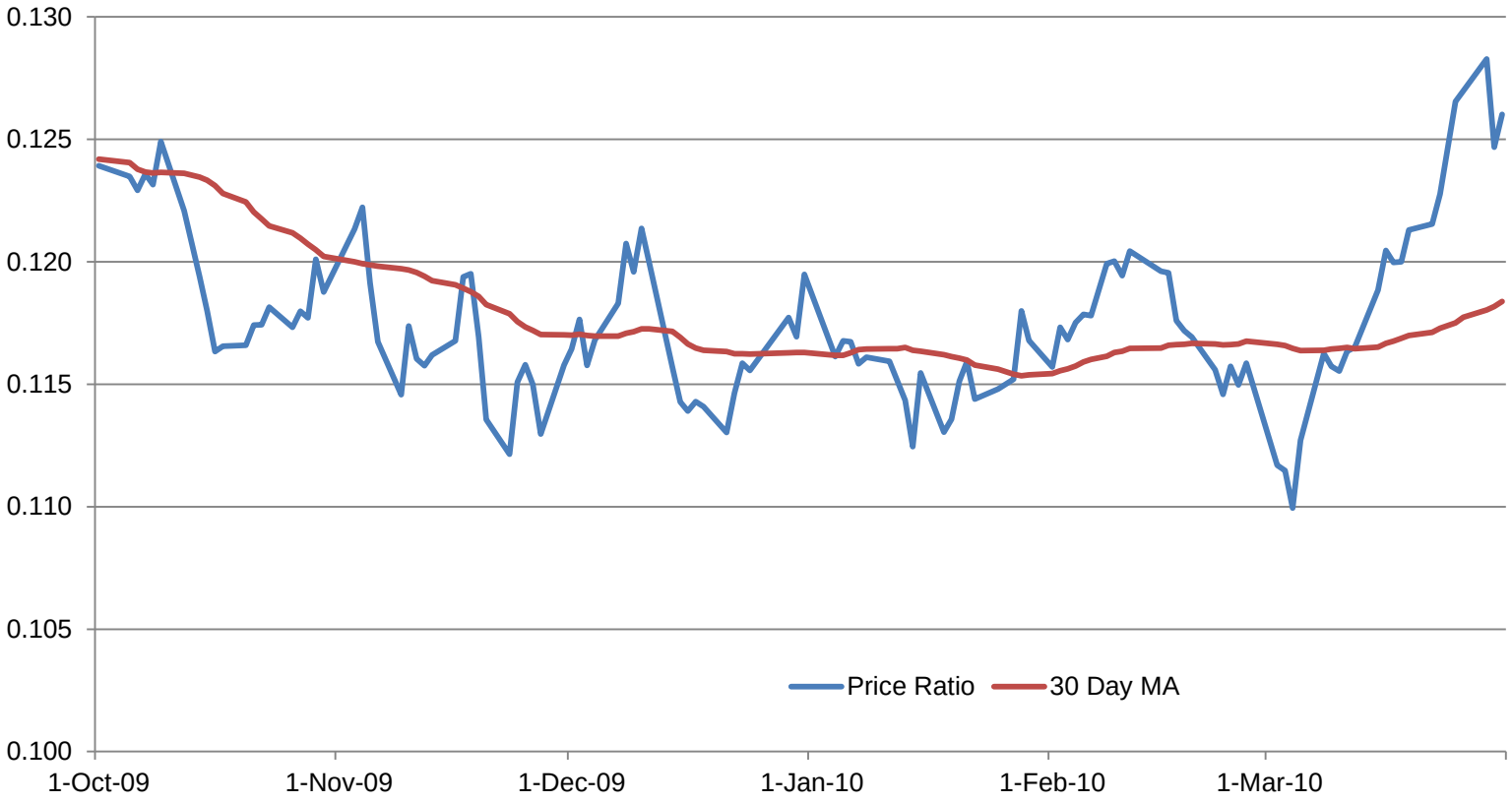
CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Stat Arb: Example



CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Stat Arb: Example



CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Stat Arb: Risks

- Statistical Arbitrage is not a risk-free strategy. Key risk involved in trading are as follows:
  - Rather than converging, prices of the two securities can begin to diverge (drift apart), i.e. the spread picks up trend rather than mean-reverting to the original mean and the cointegration can be broken.
  - Sometime, an event in a security can trigger extreme movement in pair ratio. Entering into such a security trade should be avoided.
  - Strict risk management of Stop Loss is required to handle adverse situations once the mean-reverting behavior is invalidated.
  - Inherently stat-arbs are like short gamma trades. They tend to win most of the time and sometimes blow up!

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Stat Arb: Cointegration

- What is co-integration?
  - If there exists a stationary linear combination of non-stationary asset prices, the combined prices are said to be co-integrated.
- What are the co-integrated process?
  - Cash and Future
  - Stocks in same industry and similar business model
  - Term structure of yields, etc.
- How are they useful for Pair Identification?
  - Return-based pair-trading only works at high frequency!!
- Process of Determining Co-integration
  - Techniques used to identify Co-integration are:
    - Augmented Dickey-Fuller Test
    - Jonhansen Test

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

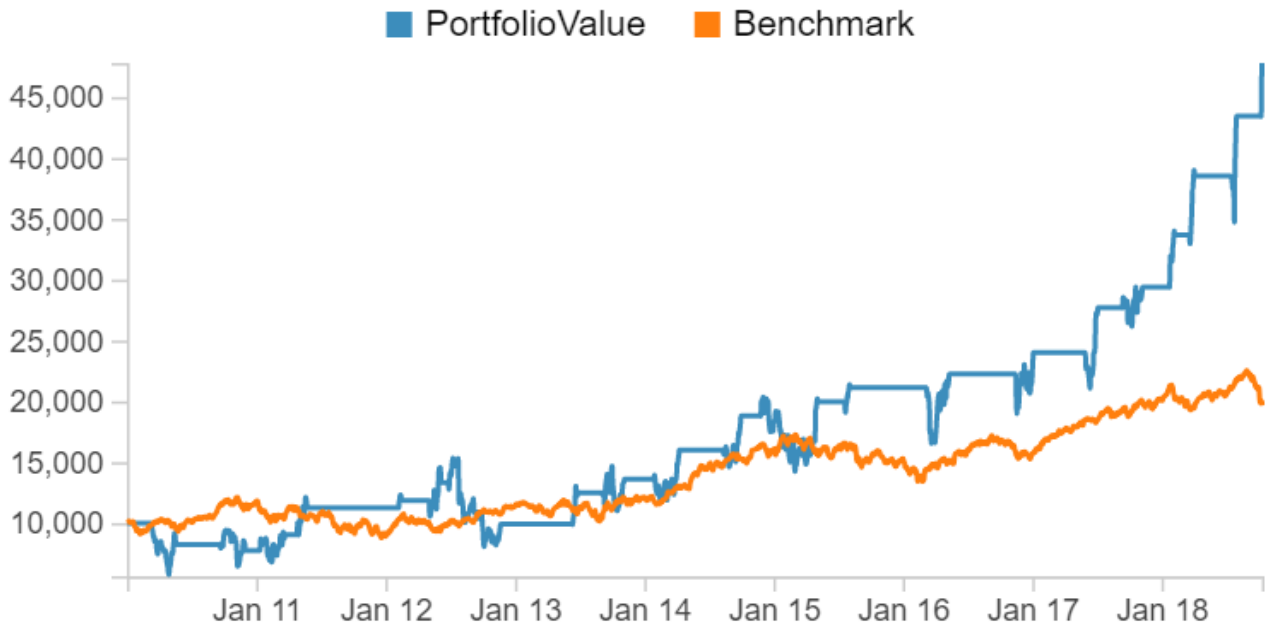
Stat Arb: Steps

- Pair Identification
  - Based on prior knowledge
  - Based on quantitative analysis (example: cluster analysis)
  - Finally confirm hypothesis by cointegration test
- Rich cheap identification?
  - Track the cointegrated pair and compute z-spread
  - Define threshold of entry and exit
  - Check for idiosyncratic nature (news flows, company specific announcements etc)
- Risk Management
  - Position sizing: based on expected time to convergence, spread volatility over the period and strength of the signal
  - Step in vs all-in sizing

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Stat Arb: Performance

RETURNS	ALPHA	BETA	SHARPE	DRAWDOWN
377.95%	0.26	-0.08	0.63	-49.45%



CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

"A rising market and a long position is a sure sign of investment genius"

John Kenneth Galbraith (Economist)

Systematic Trading in FX

FX Styles and Alpha Strategies

blueshift

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

The Fundamental Beta in FX

Value: “cheap” currencies will go up in value, and reverse for “expensive” ones. Measure of richness/ cheapness requires economic or empirical models (like purchase power parity, REER or statistical regression-based etc.)

Momentum: Past returns are persistent. Rank currencies on past returns, long the top ones, short the bottom ones(cross-sectional momentum). Alternatively, long if past returns are positive (above threshold) and short if the reverse (time-series momentum)

Carry: Positioning to profit if nothing changes in the universe except passage of time. Long currencies with higher yields (interest rates) and funded by ones with low yields.

Defensive: Low risk currencies (better sovereign ratings) produces better risk-adjusted returns. Long currencies with higher sovereign ratings and short the bottom ones.

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Creating A Carry Strategy

1. *A bit more on the logic*
2. *Which one we should buy and which ones to sell*
3. *Coding the strategy – Starting from the template*
4. *Adding the momentum logic*
5. *Adjusting the re-balance function*
6. *Running the back-test*

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Creating Other Factor Strategies

1. *Pick up from where we left*
2. *Which one we should buy and which ones to sell*
3. *Coding the strategy – Starting from the template*
4. *Running the back-test*

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Creating A Factor Basket

1. *Pick up from where we left*
2. *Simple allocation of 1/3 in to each of carry, value and momentum*
3. *Combining the strategies*
4. *Running the back-test*

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

"Past performance is the best predictor of success"

Jim Simons (Renaissance Technologies)

Back-test Evaluation

Maths , Pseudo-maths and Overfitting

blueshift

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Risk Reward Metrics

Equity Curve – picture tells a thousand metrics

Algorithm Metrics *:*

Risks and Performance: Total returns, annualized returns, volatility, Sharpe ratio, Sortino ratio, Omega ratio, time-series stability, skew, kurtosis, max drawdowns, alpha, beta, net and gross leverage,

Trade Metrics*: Average number of trades, per trade profit, Win rate, Total trades, Profit-loss ratio, average wins, average loss, average holding period, average costs, maximum adverse excursion, maximum favorable excursion*

Scenario Analysis*: Strategy performance in defined periods – 2007 quant crisis, 2008 financial crisis, 2011 US Budget crisis, 2010-2012 European sovereign crisis, EM crisis, Other identifiable periods of idiosyncratic risks*

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Risk Reward Metrics

Sharpe Ratio:

$$Sharpe = \frac{R_P - R_f}{\sigma_P}$$

Sortino Ratio:

$$Sortino = \frac{R_P - R_B}{\sigma_{down}}$$

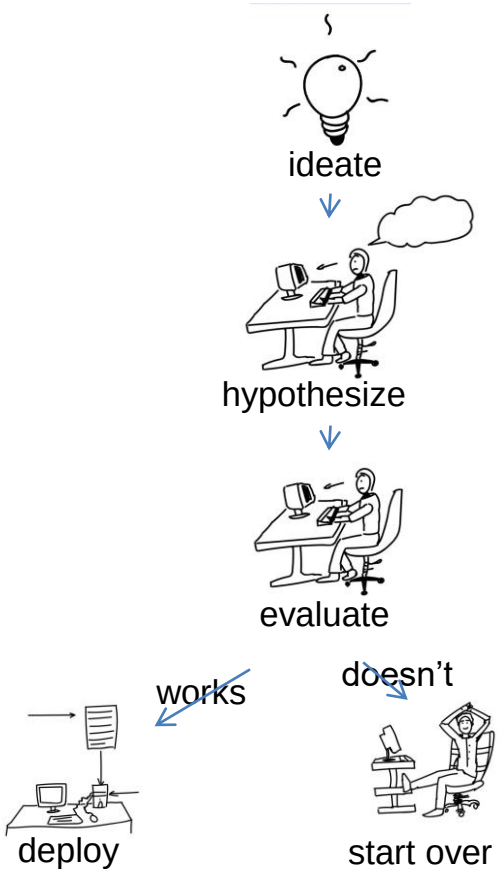
Omega Ratio:

$$\Omega = \frac{\int_r^{\infty} \big(1 - F(x)\big).dx}{\int_{-\infty}^{r} F(x).dx}$$

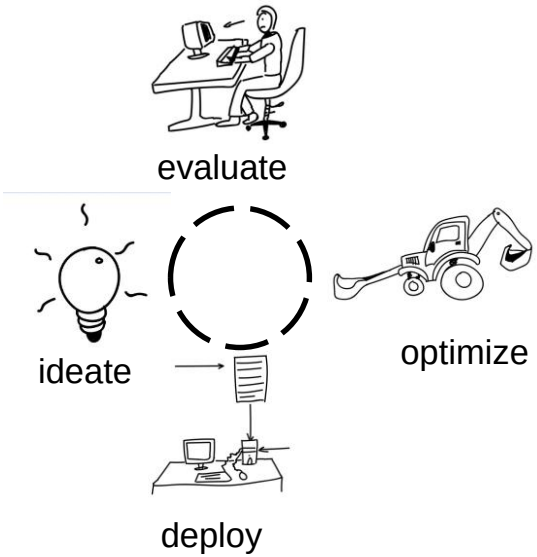
CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Back-testing: Optimize?

Scientific way of developing a strategy:



The way most of us end up with:

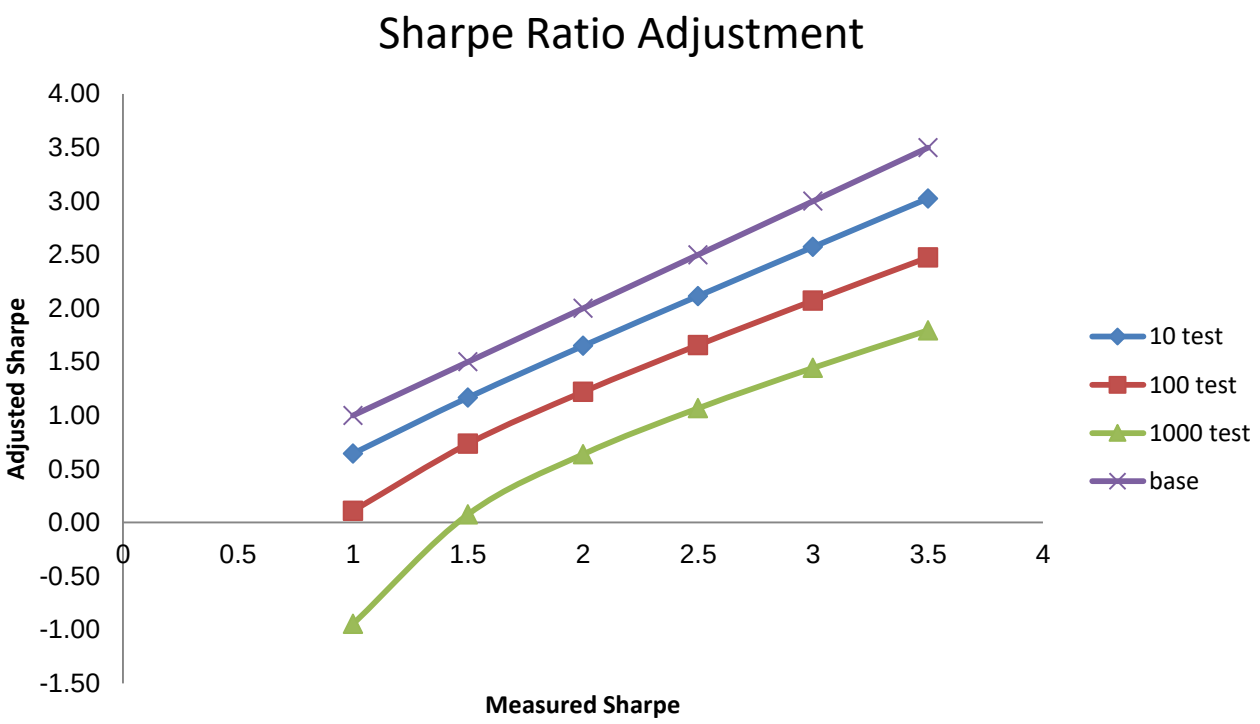


Back-tests are useful for hypothesis testing
Not a data-fitting tool

We do not want to optimize our backtest performance. We want to optimize expected live performance.

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Problems with Optimization



Alternative to Optimization

1. Adaptive Strategies: Example Change Point Analysis
2. Stable Factor research: With economic/ behavioural justification
3. Ensemble Strategies: Works best with diverse strategies [not in demo]
4. Validation – cross validation or paper-trading

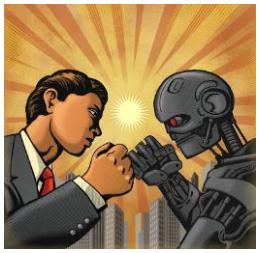
CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

What It Is Not

Black-box: A systematic strategy should **NOT** be a black-box. We should **ALWAYS** be able to explain why we buy what we buy (and sell); and why does it make money.

Risk-free Profit: **No** systematic strategy (or any strategy!) is risk free, and will **never** make money all the time, every time. Risk management is very important.

Fire-and-Forget: A systematic strategy is an **evolving** one. It must pass through the constant cycle of performance measure, tuning and re-risking.



Man vs. Machine: Systematic strategies are not about man vs. machine, but man **and** machine. Human brain is far better evolved to develop hypotheses. Machines are far better at testing and executing them.

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

Strategy Development - Recommendations

Always save all your strategy variables in the `context` environment
- helps tracking variables and re-starts

Use schedule_function if your algo does not need to respond to every bars
- will run faster

Good practice to check account leverage to make sure the algo is running as intended

Do not over-fit – build a hypothesis and test. Data-mining (p-hacking) generates good performance in the past and bad going forward

Check the stats in the backtest results. A strategy with low Sharpe, concentrated wins (low stability), large downside risks (negative skew), and large drawdowns and high beta are not one you can actually invest in

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

List of Demo Strategies

1. *Strategy Template*
2. *Value Strategy*
3. *Carry Strategy*
4. *Time-series momentum strategy*
5. *Cross-sectional momentum strategy*
6. *Factors basket strategy*
7. *Bollinger band based break-out strategy*
8. *RSI based mean-reversion strategy*
9. *Correlation based statistical strategy*
10. *FX daily computation template*
11. *Change point based regime switching strategy*

Find them at: https://github.com/QuantInsti/blueshift-demo-strategies

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.

THANK YOU!

For further queries reach out to blueshift-support@quantinsti.com

CONFIDENTIAL. NOT TO BE SHARED OUTSIDE WITHOUT WRITTEN CONSENT.