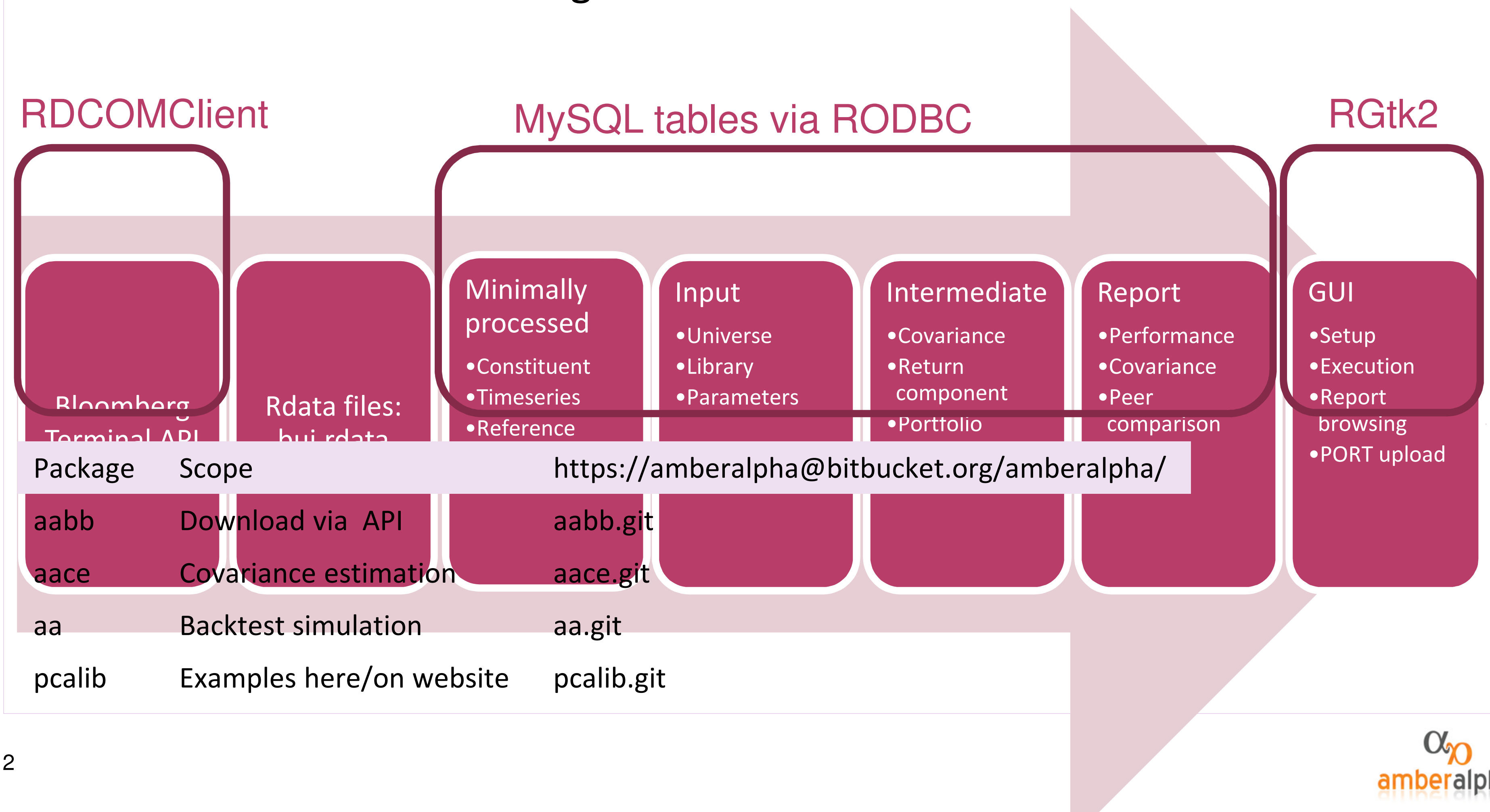
COVARIANCE FORECASTING FOR PORTFOLIO OPTIMISATION

R/Finance
Chicago
May 2013

Package aa: repeatable backtest simulation for equities

- Three inputs (universe, library, parameters) all logged in the database
- The user's 'alpha library' operates on zoo/xts objects from SQL database
- Estimation and forecasting of covariance is central



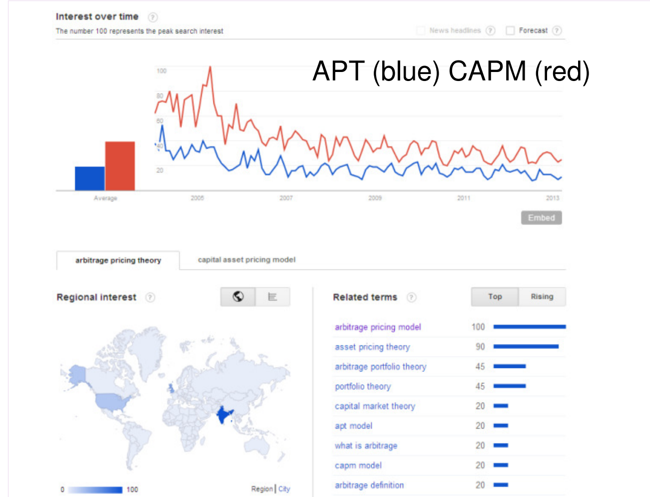
Package	Scope	https://amberalpha@bitbucket.org/amberalpha/
aabb	Download via API	aabb.git
aace	Covariance estimation	aace.git
aa	Backtest simulation	aa.git
pcalib	Examples here/on website	pcalib.git

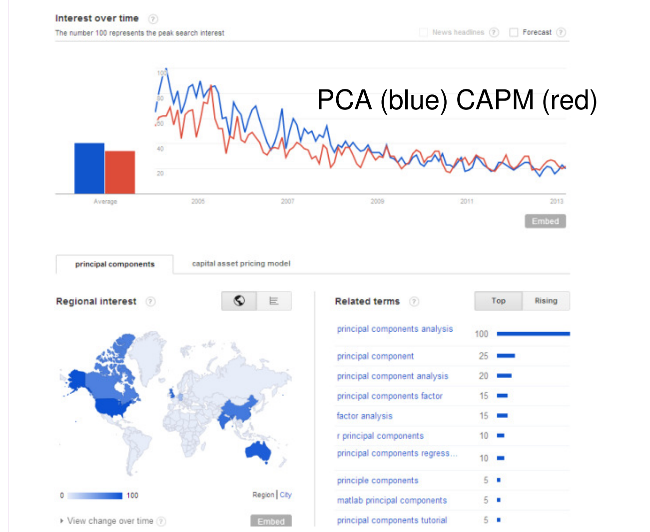
CAPM, the APT, PCA and all that

returns loadings scores specific returns

$$R = a + B \;.\; f + \epsilon$$

- Arbitrage Pricing Theory (APT)
  - If factor 1 is ‘the market’, CAPM nests into APT
  - factors 2:k remain to be specified
- Identifying the factors
  - Regression: cross-sectional loadings known a priori
    - Size
    - Value etc …
    - Industry (directional)
  - Regression: timeseries scores known a priori
    - Market (directional)
    - Bond yield changes
    - Oil price changes
    - Surprises in general
  - PCA-type: scores are ‘portfolio’ returns
    - Exists a choice: ‘the answer in finite samples is not clear’
      - Factor Analysis
      - Principal Components Analysis





These Google trends graphs are for entertainment only!

Principal Components Analysis

Covariance Eigenvectors Eigenvalues

$$\Sigma \quad = \quad A^T \quad . \quad \Lambda \quad . \quad A$$

$$I \quad = \quad A^T . A$$ eigenvectors are orthonormal

- PCA as ‘dimension reduction’
  - eigenvalues are descending in magnitude
  - covariance can be summarised with the first k
  - precisely fits the bill for APT if stated as follows:

$$B \;=\; A_{.1:k} \;.\; \Lambda^{.5}_{1:k,1:k}$$ loadings

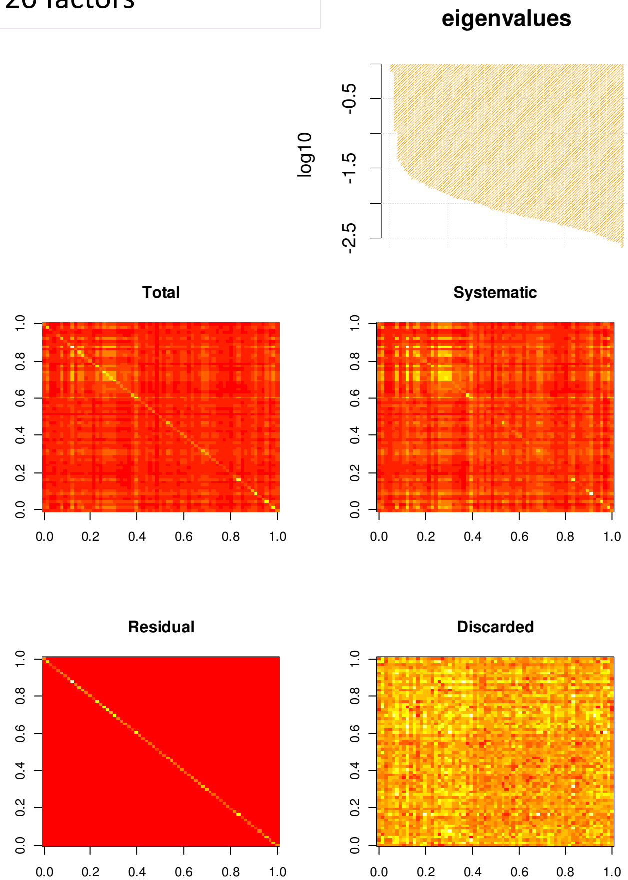
$$f \;=\; R \,.\, A_{.1:k} \;.\; \Lambda^{-.5}_{1:k,1:k}$$ scores (unit variance)

$$\epsilon \;= R - Bf$$ specific returns

- Remove off-diagonals for specific returns

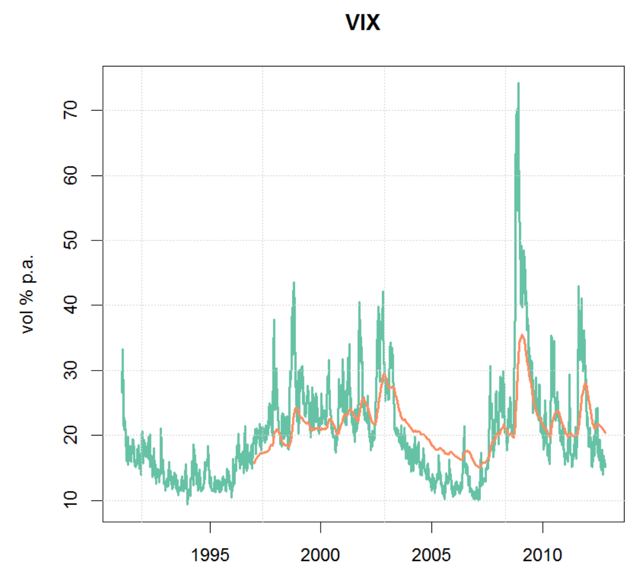
$$\Sigma^* \quad = \quad A^T_{.1:k} .\; E_{1:k} \,.\, A_{.1:k} + diag(A^T_{.k+1:n} .\; E_{k+1:n} \,.\, A_{.k+1:n})$$

Covariance illustration
Example:
67 utilities
230 weeks
20 factors



Tweaks

- Part 1 : see package BurStFin
  - Model correlation, not variance
  - Handling missing data
    - Model ‘individuals’ (stocks) with complete data
    - ‘Regress in’ incomplete stocks, factors 1:k’
    - Assign average loadings for lower-ranked factors
  - Iterate the process (repeat, starting from Σ*)
- Part 2 : more niceties
  - Optimise the ‘factor portfolios’
    - We had $\mathrm{f} = A_{.1:k} \ . \Lambda^{-.5}_{1:k,1:k}$
    - Satisfies orthonormality but does not minimise specific risk
    - Use package quadprog for constrained optimisation
  - Apply Bayesian shrinkage
    - prior: $B = \bar{B}$, stronger for low eigenvalues
  - Missing data in cross-section
    - Frequent occurrence in global data
    - Substitute NA with 0, compute systematic returns, iterate
  - Inverse VIX scaling to make vol stationary
    - VIX is forward-looking option vol and a better forecast
    - Some smoothing is required, though



amberalpha

Testing for breakdown: lag notation

- Rolling estimates
  - We update models regularly to reflect changes
  - Further information can be extracted from them

lag		return date		model estimation date
τ	$=$	t	$-$	T

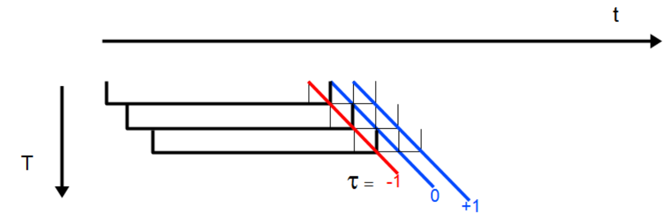
  - Applied to the factor model estimate

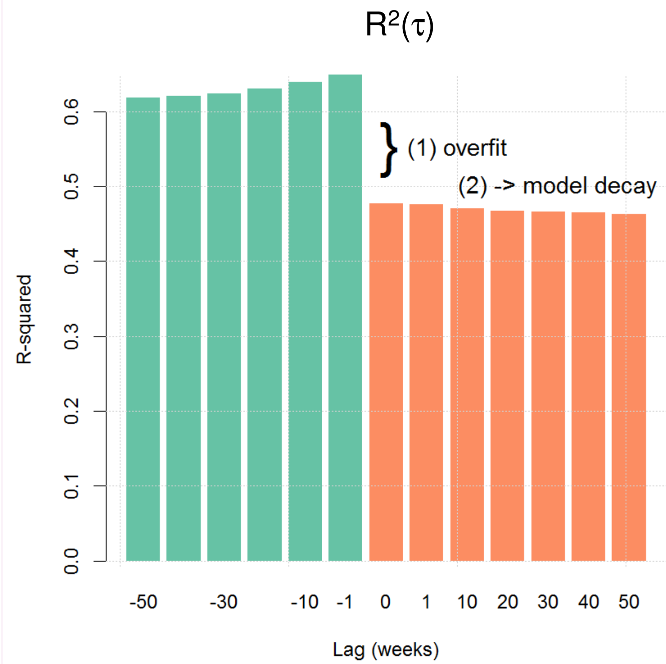
$$R(t) = B(T).f(T,t) + \epsilon(T,t)$$

- Example:
  - Analyse all return(t) through a single estimate T*
  - Analyse single return(t*) through all estimates (T)
  - Take constant-lag (τ*) panels
    - Negative lags are not feasible, they are in-sample
    - Zero lag is usual, means 'using the latest estimate'
    - Positive lags are feasible, mean 'using older estimates'

- Application: consider the fit $R^2(\tau)$

Notation for rolling estimates





Testing for breakdown: regress returns on components

- An augmented factor regression

Define components of systematic return: $R_m = b_1 . f_1$ etc

$R(t) = a +$

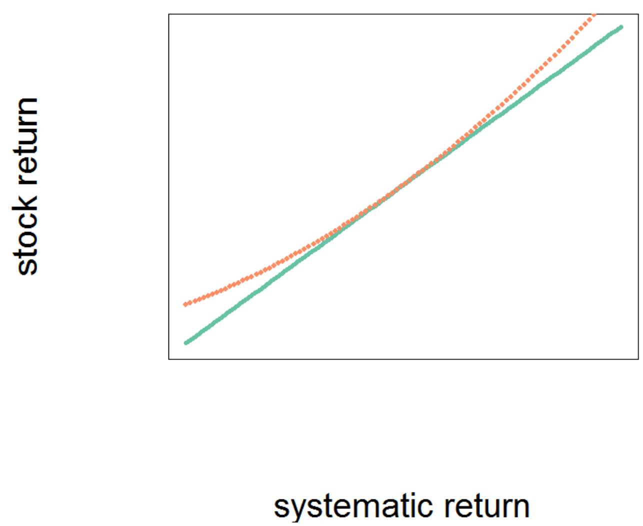
$c_m(\tau). R_m$	$+$	market (factor 1)
$c_s(\tau). R_s$	$+$	systematic (factors 2:k-1)
$c_l(\tau). R_l$	$+$	kth factor
$c_d(\tau). (R_m - \overline{R_m})$	$+$	deviation from mean
$c_q(\tau). (R_m - \overline{R_m})^2$		quadratic

Note: coefficients can be adjusted for 'errors in the variables'

- kth factor: the 'marginal' explanatory power
- deviation: c_d<0 implies mean-reversion
- quadratic: option-like payoff

```
#Testing the null
res <- summary(m1 <- lm(returns ~ Rm + Rs + Rl + Rd+ Rq -1))
linearHypothesis(m1,c(0,0,0,1,1))["Pr(>F)"]
linearHypothesis(m1,diag(5),c(1,1,1,0,0))["Pr(>F)"]
```

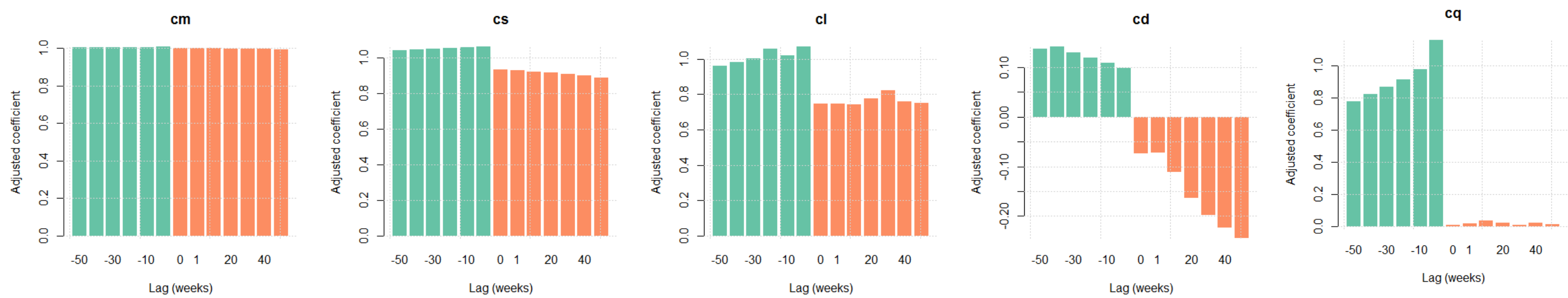
0<c_q for +ve gamma



amberalpha

Results from testing the model

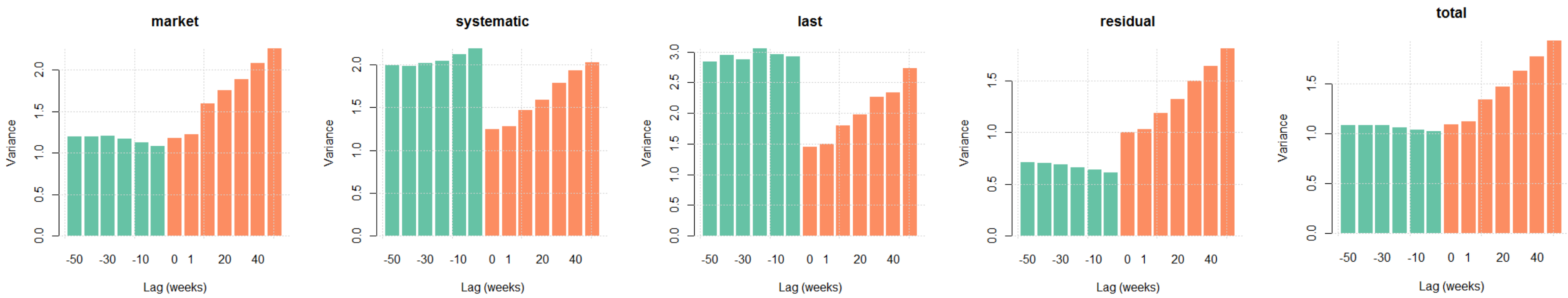
- Coefficients



- Market and systematic components have coefficients close to 1 at lag 0
- Deviation shows
    - In-sample, less shrinkage for better fit (as expected)
    - Out of sample, more shrinkage for better fit (about -0.07 more, on top of -0.3 applied)
    - Evidence of mean-reversion as the coefficient decreases for greater lags
- Quadratic term is pure noise

Results from testing the model

- Return variance/forecast variance

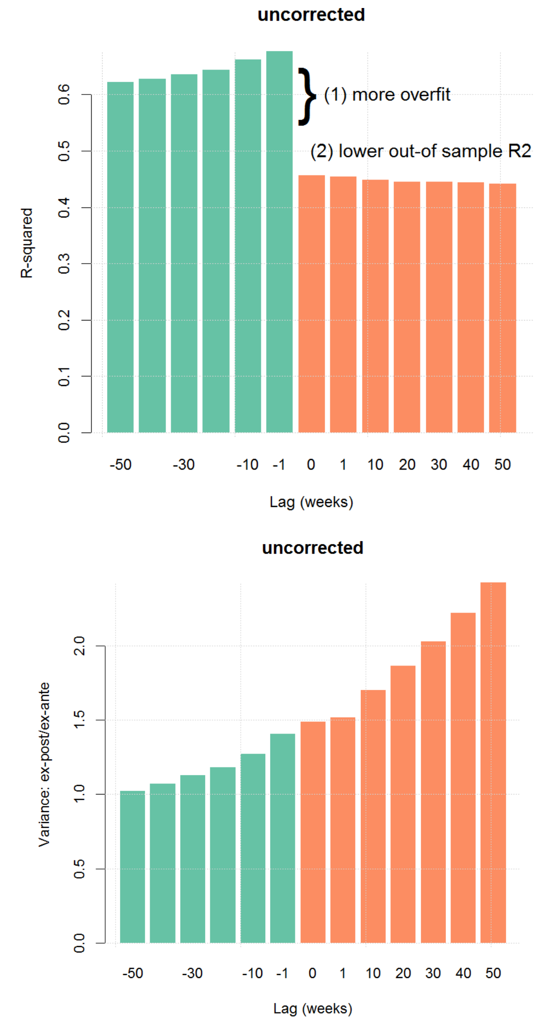


- In general the lag-zero normalised variance is very close to 1
- Some evidence of slight under-forecasting of systematic components
- (The trend upwards out-of-sample is period-specific, has no significance)

amberalpha

Testing for breakdown: without adjustments

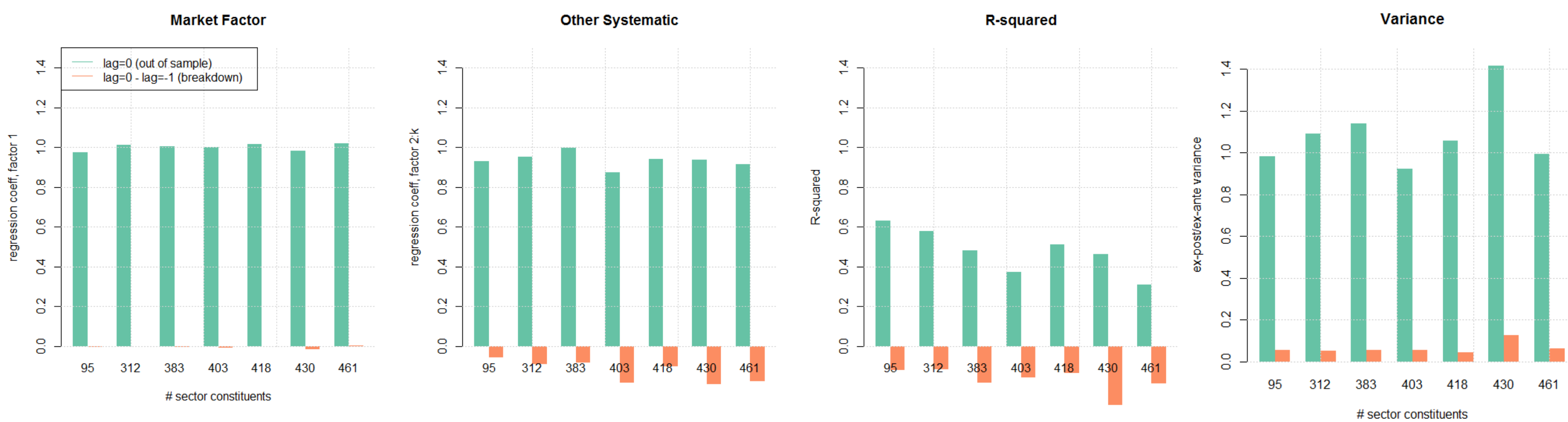
- Without shrinkage or VIX
  - $R^2(\tau)$ shows
    - Higher in-sample fit (0.68 vs 0.65)
    - Lower out-out-of-sample fit (0.46 vs 0.48)
  - $\sigma^2(\tau)$
    - underprediction is greater (1.49 vs 1.18)
  - These differences fit with expectations



amberalpha

Testing across estimation universes

- Results up to this point are averages across 7 global sectors
- Here examine impact of sector size on the results
- Report level of metric at $\tau = 0$ and change from $\tau = -1$

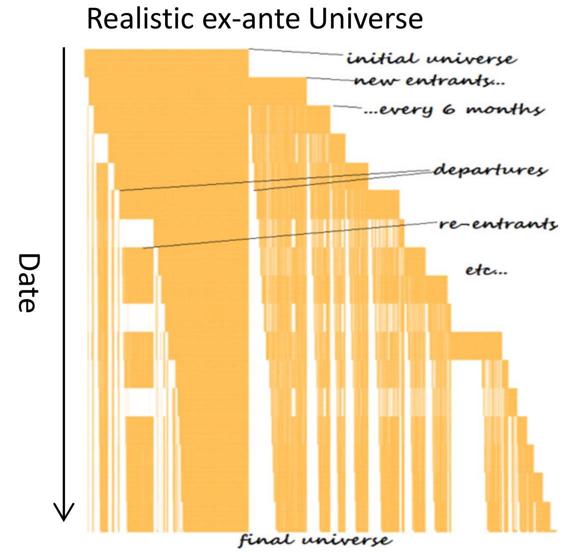


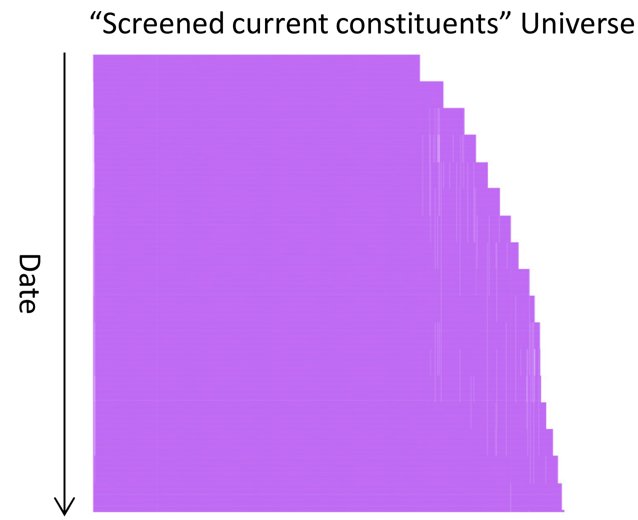
- R^2 is lower for larger universes, where k/n is lower
- No evidence of greater out-of-sample breakdown for larger universes

amberalpha

Universe selection

- Requirements
  - No survivorship bias
  - Stable identifiers
  - Stationary screening criteria
- A possible solution
  - Historical index constituents, screened
  - Issues with criteria, licencing
- A specific solution
  - Bloomberg bworld = bworldus+bworldeu+bworldpr
  - Identifier ‘unique identifier’ or ‘open symbology’
  - Screen on
    - Exchange, geographical, sector classification
    - Weight – maintains a fairly stationary universe composition
    - Liquidity
    - Alpha data coverage
- Impact of eliminating biases
  - ‘distress’ type performance is sensitive
  - Turnover is higher



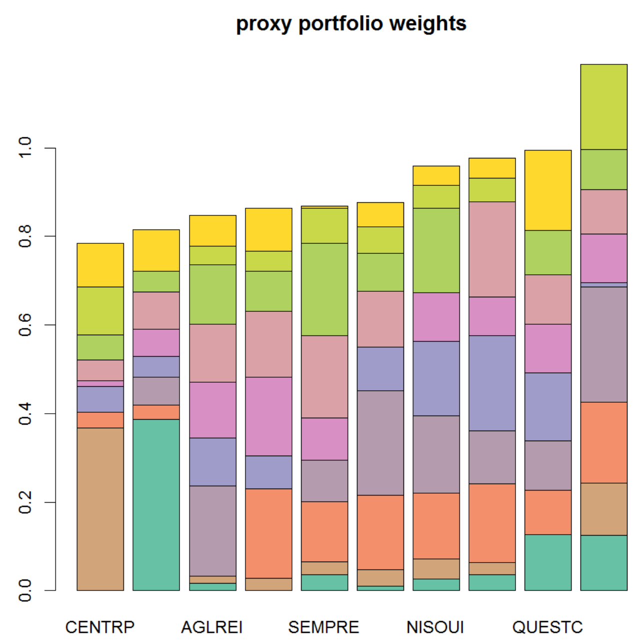


Application 1: forecast-free portfolio construction

- The 'proxy basket'
  - A small portfolio of closely matched stocks
  - Tracks the target stock with minimum variance
  - It is a constrained, optimal version of $\Lambda.\Lambda^T$
- Optional constraints
  - No shorting
  - Weights sum to unity
- Uses
  - Hedging
  - Alpha-generation: mean-reversion and statarb
  - Application to valuation using yield-like variables
    - the basket is a risk-matched benchmark in the same industry
    - It is the single best 'comp' for valuation
  - Regression tests: on return or yield

```
require(quadprog)
x <- diag(n)
solve.QP(Dmat=ce,dvec=0,Amat=cbind(x[,i],x[,-i]),
        bvec=c(1,rep(0,n)),meq=1)
```

Minimise:	$w'\Sigma w$
Subject to:	$w_i = 1, \forall\, w_i > 0$
Example:	Gas Utilities, 2012-10



amberalpha

Application 2: market-neutral tilt portfolio

$U = E[R] - \lambda\sigma^2 = w.\hat{r} - \lambda\, w^T.\Sigma.w$

$\lambda^* = \Sigma^{-1}.\,\hat{r}/x_g^*$ unconstrained solution defines $\lambda^*(x_g^*)$

- Forecast options
  - centred ranks: uniform distribution
  - ‘normalise’ using inverse cdf
- Constraint options
  - Factor-1 neutral within sector/region
  - Factor 2:k neutral
  - Position size
- Leverage options
  - Gross exposure (x_g^*)
  - Volatility

```
library(quadprog)
sol <- uniroot(f = tgtqp, interval = c(5, 0.05) *
             objfun(w = solve(Dmat, dvec), Dmat = Dmat)/tgt,
             Dmat = Dmat, dvec = dvec, constr = constr,
             tgt = tgt, tol = 0.1, objfun = objfun)#(schematic)
```

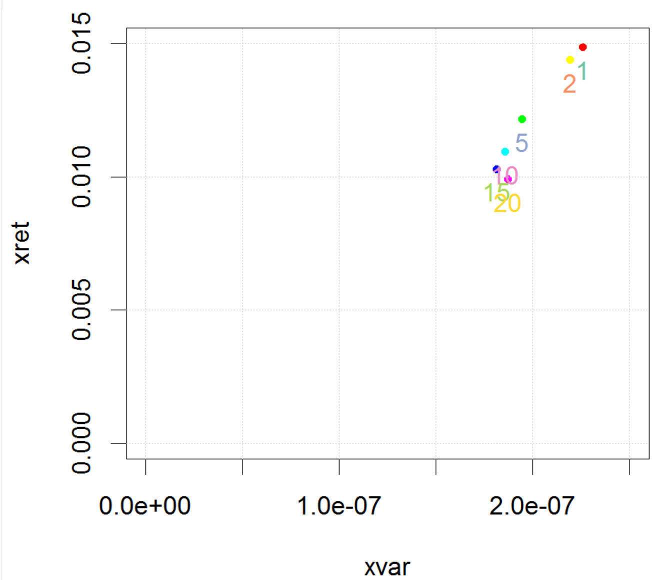
Solve for λ, w

Subject to: $w.B_{.1:n} = 0$

$\Sigma|w| = 0.01$

Equal forecasts

Example: Utilities, 2011-10

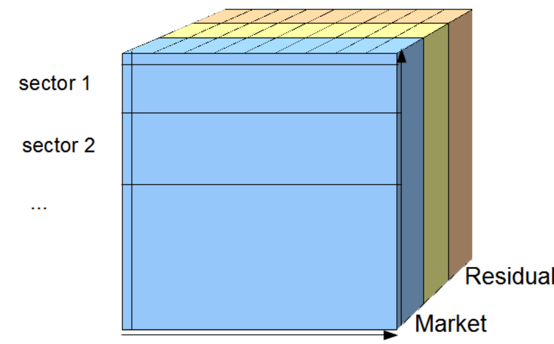


amberalpha

Attribution

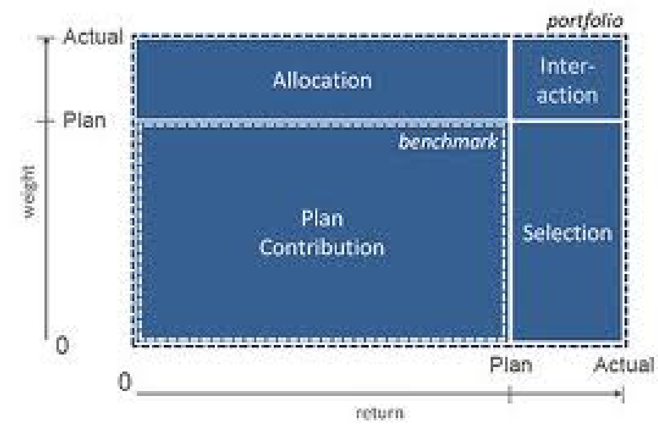
- For a single period
  - Variance and return have the same components
    - Market (factor 1)
    - Systematic (factors 2:k)
    - Residual
  - Drilldown into category trees
    - Geographical tree (region, country, state)
    - Industrial classification tree (GICS, ICB, BICS)
    - Long, short subportfolios
  - Option to apply different trees on the two axes
- Multi-period treatment of returns
  - 'Simple' with no compounding, or…
  - 'Smoothing' scheme, redistributing interactions
- Contrast with Brinson/Fachler and extensions
  - No benchmark (cash benchmark)
  - Consequently no selection/allocation/interaction
  - Currency easy:
    - local returns vs local cash (hedged)
    - $ returns vs $ cash (unhedged)
  - Leverage easy: premia are self-financing

Cash benchmark



$$\sigma^2_{port} = \sum_i \sum_j w_i \,.\, w_j . \sigma_{i,j,M|S|R}$$
$$R_{port} = \sum_i w_i . R_{i,M|S|R}$$

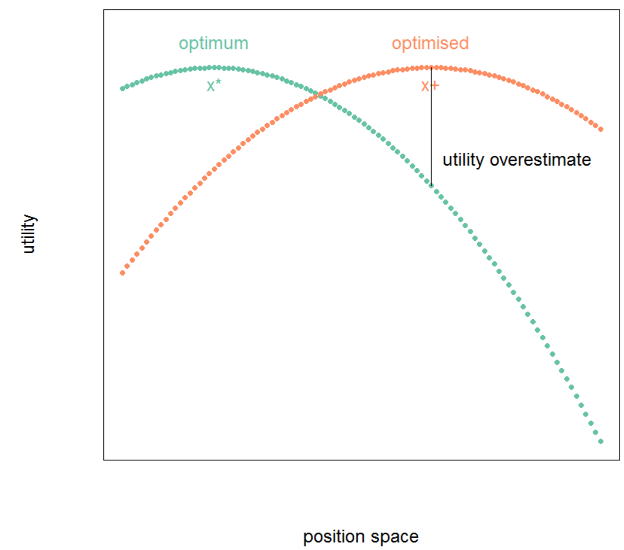
Brinson/Fachler



amberalpha

Mean-variance optimisation criticisms

- Markowitz optimisation
  - Ignores sampling errors in the covariance matrix
  - Solves a mis-specified problem
- Portfolio weights
  - Sensitive to sampling errors in covariance
  - Do not follow intuition
  - Are distant from true optimality
- Utility and Sharpe Ratio
  - Variance is underestimated
  - Expected return is overestimated
- ‘Solutions’ have been proposed
  - But how serious is the problem?
  - Might the answer depend on the risk model type?



$$U(w) = E[R] - b.\sigma^2$$

$$SR(w) = \frac{E[R]}{\sigma^2}$$

The trivial read-across from Utility(w) -> Sharpe(w)

amberalpha

Optimisation: a monte-carlo test (1)

- Using 'vanilla' PCA without shrinkage
- From an estimated covariance matrix
  - Generate synthetic data
  - Re-estimate covariance from this
  - Draw expected return from a uniform distribution
  - Optimise
    - subject to industry group neutrality constraint
    - Adjusting risk-aversion to target volatility
  - Repeat for seven global equity sectors
- Results
  - optimum and optimised weights correlate highly
  - True vol is 1.16x expected vol from the estimate
  - Expected return is 1.09x the true optimum

sector	number of stocks in model	euclidean length of optimal portfolio	euclidean length of error	correlation of x* and x+	expected return of optimal	expected return of optimised	true variance of optimised
1	213	0.37	0.13	0.95	1.48	1.63	1.41
2	246	0.31	0.09	0.97	1.32	1.44	1.32
3	306	0.39	0.10	0.97	1.77	1.91	1.29
4	179	0.36	0.10	0.97	1.25	1.35	1.30
5	290	0.37	0.10	0.97	1.71	1.86	1.31
6	196	0.36	0.10	0.97	1.34	1.45	1.30
7	60	0.67	0.26	0.95	1.40	1.57	1.46
mean	**213**	**0.41**	**0.13**	**0.96**	**1.47**	**1.60**	**1.34**

Optimisation: a monte-carlo test (2)

- Apply Bayesian shrinkage
  - Prior: $B = \bar{B}$, stronger for low eigenvalues
  - Prior weight is 0.3 for factor 1 loading
- Underestimate of volatility is reduced
  - $\overline{\left(\frac{\sigma_{true}^2}{\sigma_{estimated}^2}\right)} = 1.23$; $\overline{\left(\frac{\sigma_{true}}{\sigma_{estimated}}\right)} = 1.11$
- Conclude
  - Underestimation of volatility has been exaggerated
  - Higher return partially compensates
  - Optimisation is not ‘eating’ alpha or IR in this case

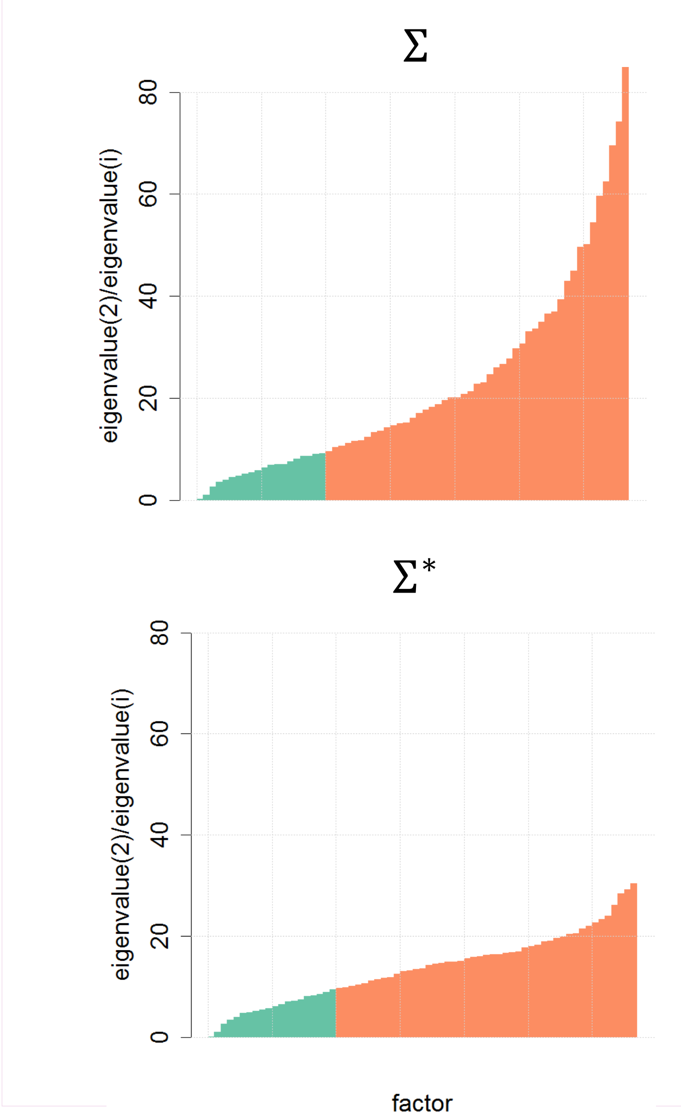
sector	number of stocks in model	euclidean length of optimal portfolio	euclidean length of error	correlation of x* and x+	expected return of optimal	expected return of optimised	true variance of optimised
1	213	0.38	0.11	0.96	1.56	1.59	1.20
2	246	0.31	0.09	0.96	1.25	1.30	1.19
3	306	0.41	0.10	0.97	1.86	1.89	1.16
4	179	0.36	0.11	0.95	1.25	1.29	1.21
5	290	0.36	0.11	0.96	1.77	1.82	1.22
6	196	0.37	0.11	0.96	1.41	1.45	1.21
7	60	0.69	0.28	0.93	1.23	1.32	1.43
mean	**213**	**0.41**	**0.13**	**0.95**	**1.48**	**1.52**	**1.23**

Optimisation as a projection

- An observation
  - Recall that for unconstrained optimisation

$$\Sigma \quad = A^T.\Lambda.A$$
$$\Sigma^{-1} = A^T.\Lambda^{-1}.A$$
$$W^* \ = \lambda^{-1}.\Sigma^{-1}.E[R] = const\,.\ A^T\ .\ \Lambda^{-1}.(A.E[R])$$

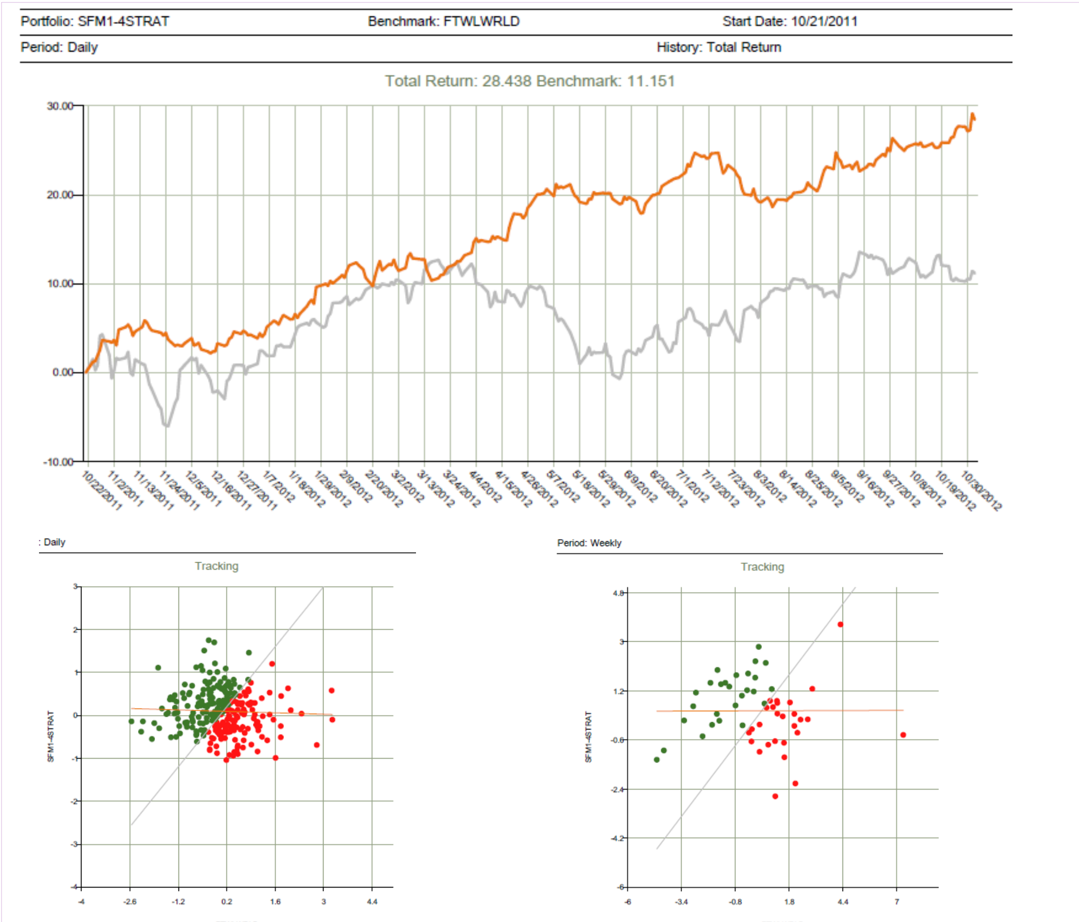
  - Or in recipe form:
    1. Project E[R] onto eigenvectors by dotproduct
    2. Divide by eigenvalues **(!)**
    3. Scale eigenvectors accordingly and sum
  - Step 2 is the error-maximisation property
    - It relates to the condition number of the matrix
    - The correction $\Sigma \rightarrow \Sigma^*$ has reduced this
    - Shrinkage reduces it further
- Modified PCA reduces error amplification
- Industry-level neutrality constraints
  - Reduce systematic risk
  - Condition the solution, largely driven by $E[R]/A_1$.



amberalpha

A live application in real time: global equity market neutral

- A contrarian low-frequency strategy run in real time as a ‘paper portfolio’ performed in-line with backtest
- But too low vol: 8% for 8x leverage and net long, not $-neutral so not ‘market neutral’ despite beta=0
- Moral: be careful what your client asks for (diversification, hedging) - you might get it and still not like it



Performance: Daily	Portfolio
Total Return 1 Month(s)	2.59
Total Return MTD	0.93
Total Return QTD	1.67
Total Return YTD	23.20
Total Return 3 Month(s)	7.71
Total Return 6 Month(s)	6.95
Total Return 1 Year(s)	23.90
Total Return 2 Year(s)	--
Total Return 3 Year(s)	--
Risk: Weekly	
Standard Deviation 1 Year(s)	8.07
Semivariance 1 Year(s)	8.35
Beta 1 Year(s)	--
Correlation 1 Year(s)	--
R-Squared 1 Year(s)	--
Information Ratio 1 Year(s)	--
Sharpe Ratio vs Risk Free 1 Year(s)	2.95
Tracking Error 1 Year(s)	--

source: Bloomberg PORT

Sharpe Ratio	2.91
Return/trade	0.72%
Return/gross	3.30%
Volatility/gross	1.14%
Holding period	23 weeks
Net/gross	6.3%

source: internal

amberalpha

Application: the open source talent contest

- There exists a severe barrier to entry in the capital management industry
  - Backtests have no credibility due to deliberate or unwitting exercise of hindsight options
  - Paper portfolios are a waste of time – trade at the touch, and one can run many

No track record

No capital

- Move to a clear and transparent process with detailed reporting
  - Anonymous enrolment into accredited paper portfolios, executed at VWAP
  - Detailed drill-down reporting on positions, risk, liquidity , performance
  - Capital introduction / seeding from early investors

Process Development

Anonymous Enrolment

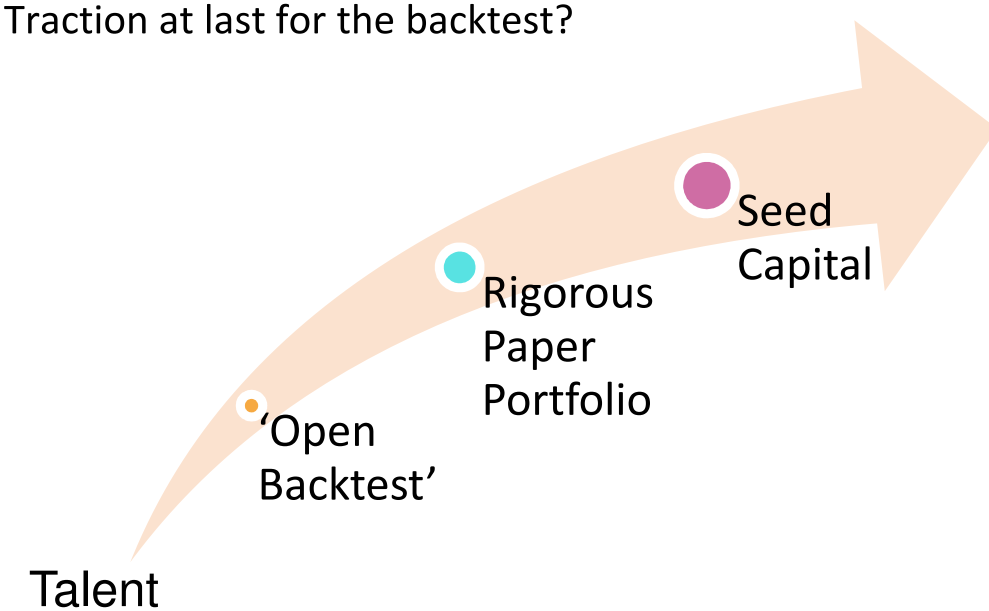
Trading At VWAP

Selection Introduction

Capital

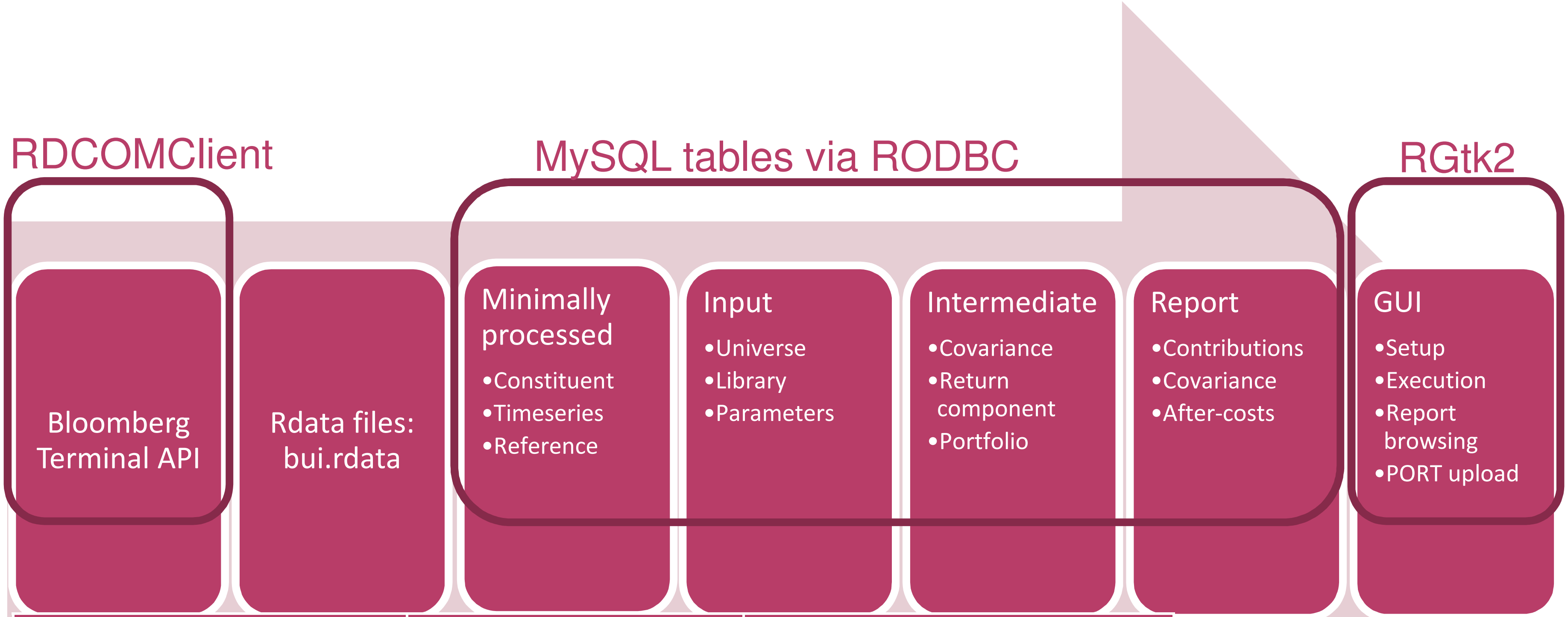
Not just another backtest

- We have seen
  - A framework for rigorous testing of an equity covariance model
  - Applications in backtest simulation of equity market neutral strategies
- The latent demand for an open source equity risk and backtest system
  - The market for trading talent is highly imperfect due to secrecy
  - This is the missing mechanism for matching talent with capital whilst respecting IP
  - A bias-free backtest + paper portfolio executed at VWAP can form the reference point
  - Traction at last for the backtest?



amberalpha

Process: review and timings

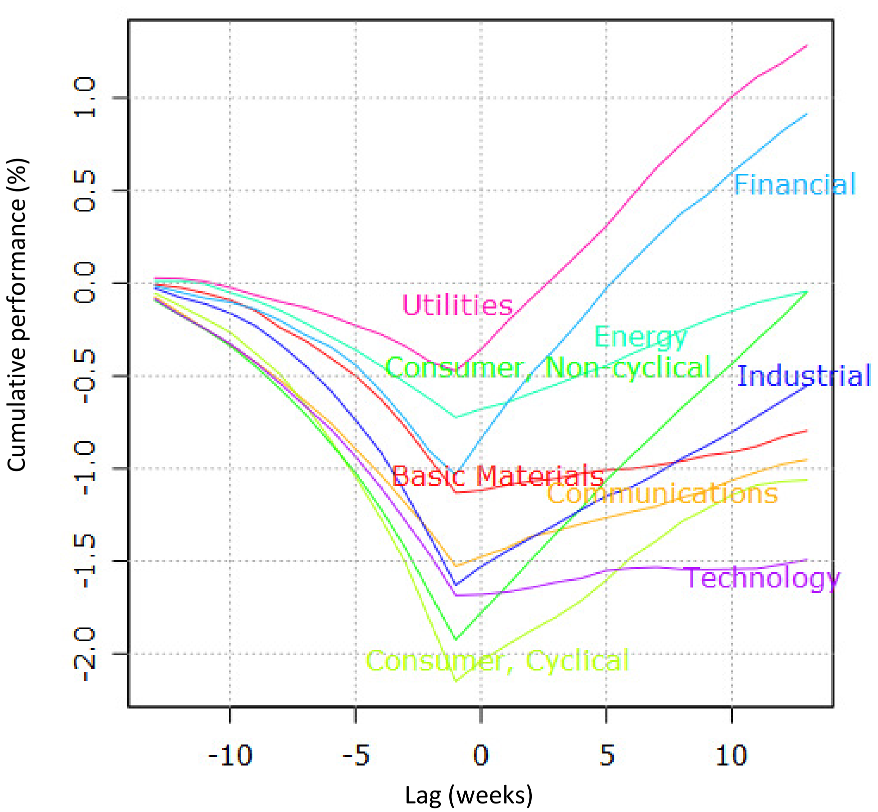


Step	Time (100-stocks 500 periods)	Possible improvement
Download timeseries updates	1 min	alternative delivery (Datalicence)
Update & extend historical data	1 min	- ditto -
Covariance estimation	2 min	In-memory
Return and yield decomposition	2 min	data.table / other in-memory
Portfolio construction and trading	2.5 min/strategy	Remove stopifnot() checks
Report generation	1 min/strategy	data.table

Attribution of performance, risk, position, directional vol

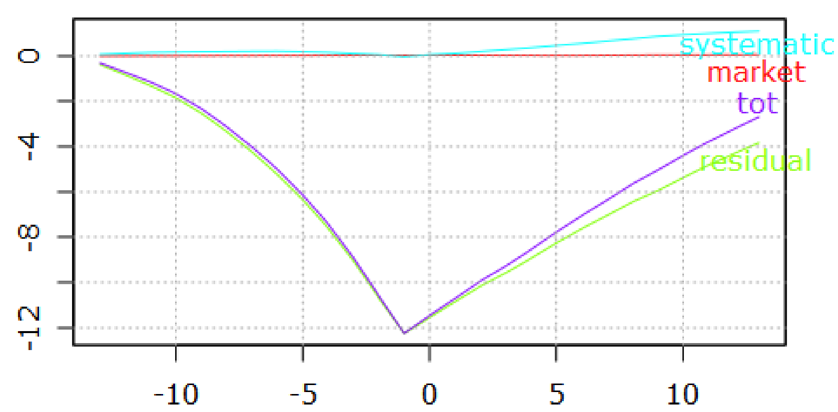
- Attributions shown are derived by lagging position with respect to returns, similar to earlier lag of covariance
- Shows that the strategy is contrarian, market neutral, residual risk, and symmetrical across long and short
- The three shown are a subset of approximately 300 tabulations warehoused for rapid browsing

Sector contribution

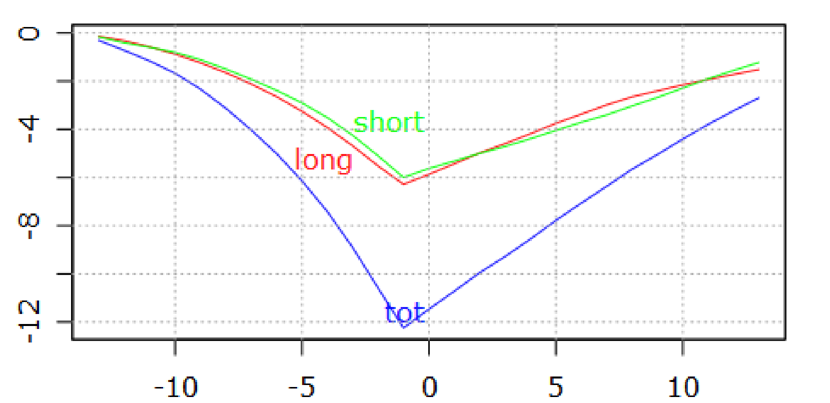


Live running 1-year period, all sectors average, bworld US + EU screened

Risk component contribution



Long / Short contribution



ATTRIBUTING PERFORMANCE

amberalpha