

# This Blog is Systematic.

Also hydromatic, automatic: Rob Carvers blog. "It's Geeks Enlightening"

Home	Systematic investment	Systematic trading	LOBOs and politics	About me	Code	My Books	Talks and media	
------	-----------------------	--------------------	--------------------	----------	------	----------	-----------------	--

Wednesday, 18 November 2015

## Random data: Random wanderings in portfolio optimisation

Everyone knows that the usual naive method of portfolio optimisation is, well, a bit rubbish. This isn't because the method is flawed, but it relies on the inputs being 100% accurate, or to put it another way we need to know precisely what the mean, volatility and correlation of future returns are going to be. But most peoples crystal balls are somewhat foggy.

Fortunately there are a few methods we can use to deal with this problem. Some of these are straightforward; like the *handcrafted* method I describe in [my book](#). Others such as *bootstrapping* and *shrinkage* are more complicated. On any particular set of asset price history one method or another may perform better. It's even possible that just by fluke the simplest naive method will do best.

I believe that to get a good grasp of which portfolio optimisation is best you need to use *random data*; and I'll do exactly that in this post.

This is the third and final post in a series on using random data. The first, which is worth reading before this one, can be found [here](#). The [second](#) post is optional reading. You may also want to read [this post](#) which shows some optimisation with real data and explains some of the concepts here in more detail.

### Optimisation 101

I'm using straightforward Markowitz optimisation. We maximise the Sharpe Ratio of a portfolio, over some portfolio weights, given the expected mean, standard deviation, and correlation of returns.

Because I'm optimising the weights of trading systems I constrain weights to be positive, and to sum to 100%. I can also assume that my systems have the same expected volatility of returns, so all I need is the mean (or Sharpe Ratio) and correlation. Finally I don't expect to see negative correlations which makes the problem more stable.

The 'one period' or *naive* method is to take all past returns and estimate Sharpe Ratios and correlations; and then plug these estimates into the model. The flaw of this method is it ignores the huge uncertainty in the estimates.

Method two - *bootstrapping* - involves repeating our optimisation many times over different parts of our data, taking the resulting weights, and averaging them out. So the weights are the *average of many optimisations*, rather than *one optimisation on the average* of all data.

The logic for bootstrapping is simple. I believe that the past is a good guide to the future; but I don't know which part of the past will be repeated. To hedge my bets I assume there is an equal chance of seeing any particular historical period. So it's logical to use an average of the portfolios which did best over all previous periods.

*Shrinkage* is a Bayesian method which involves taking the estimated Sharpe Ratios and correlation, and then "shrinking" them towards a "prior". There are numerous ways to do this, but I'm going to use a relatively simple variation. "Shrinking" involves taking a weighted average of the estimated correlation matrix (or mean vector) and the prior; the weighting on the "prior" is the shrinkage factor. For "priors" I'm going to use the "zero information" priors - equal Sharpe Ratios, and identical correlations.

By the way the nemesis of all optimisations, the very simplest method of *equal weights*, is equivalent to using the shrinkage method of a shrinkage factor of 100%. It should also be obvious that using 0% shrinkage is the same as a naive optimisation.

*Handcrafting* is a simple robust method I describe in chapter four of my [book](#). We estimate

- [My website: systematicmoney.org](#)
- [twitter @investingidiocy](#)
- [linkedin/robert-stuart-carver](#)
- [facebook.com/Robert.Carver.Author](#)
- [github.com/robcarver17](#)
- [goodreads.com/Robert\\_Carver](#)

- [Favourite books](#)
- [Useful links](#)
- [Legal Disclaimer \(please don't sue me\)](#)

#### Blog Archive

- ▶ [2019](#) (9)
- ▶ [2018](#) (10)
- ▶ [2017](#) (14)
- ▶ [2016](#) (11)
- ▼ [2015](#) (32)
  - ▶ [December](#) (1)
  - ▼ [November](#) (5)
    - Random data: Random wanderings in portfolio optimi...
    - David Versus Goliath
    - Random data: Evaluating "Trading the equity curve"...
    - Using random data
    - [Typo in definition of carry rule for futures in "S...](#)
  - ▶ [October](#) (1)
  - ▶ [September](#) (5)
  - ▶ [August](#) (1)
  - ▶ [July](#) (7)
  - ▶ [June](#) (3)
  - ▶ [May](#) (1)
  - ▶ [April](#) (2)
  - ▶ [March](#) (3)
  - ▶ [February](#) (1)
  - ▶ [January](#) (2)
- ▶ [2014](#) (21)
- ▶ [2013](#) (1)

#### Search This Blog

Search

#### Popular Posts

- [Interactive brokers native python API](#)  
Until quite recently interactive brokers didn't offer a python API for their automated trading software. Instead you had to put up wi...
- [Historic data from native IB python API](#)  
This is the second in a series of posts on how to use the native python API for interactive brokers . This post is an update of the post I w...
- [A simple breakout trading rule \(pysystemtrade\)](#)

correlations, and then use a lookup table to determine what weights to use. I'm going to use the extended table which is here.

*\* There is an extension in my book to the hand crafted method which incorporates estimated Sharpe Ratios; but to keep life simple I won't be discussing it here.*

## The analysis

I'm going to use the technique I described in the [first post in this series](#) for producing random returns for 3 assets. I'm going to use three because it will make the results more tractable and intuitive. It also means I can use the handcrafted method without creating a complex grouping algorithm. The general results will still apply with more assets.

The portfolio I'm going to generate data from has identical Sharpes (in expectation), and correlations of 0.8, 0, 0. So it's very similar to the portfolio of two equity indices and one bond index I consider in this post.

*Note that the set of correlations is deliberately different (mostly) from the set used by the handcrafted method, just to ensure there is no favouritism here.*

The usual caveat about random data applies here. The expected sharpe ratio, volatility and correlation of returns in this random world is fixed; but in reality it varies a lot. Nevertheless I think we can still draw some useful conclusions by using random data. Be warned though that non robust methods (such as the classic naive method) will do even worse in the real world than they do here.

All my testing will be done on *out of sample* windows\*. So I'll use data in the past to estimate sharpe ratio and correlations, coming up with some weights, and then run those weights for a year to see how the portfolio behaves. The amount of history you have is critical, particularly with this stylised example where there is a fixed data generation process to be 'discovered'.

I ran tests with available asset price histories of 1 year, 5 years, 10 years and 20 years for the in sample period, with 1 year for the out of sample period.

*\* this term is explained more in chapter three of [my book](#), and in [this post](#).*

To keep things simple I'll just be using the average out of sample sharpe ratio as my measure of how successful a given method is. Just for fun I'll also measure the average *degradation* of Sharpe ratio for the optimised weights between in sample, and out of sample.

I'm going to going to basically compare the four methods against each other, and see which does best on at out of sample basis. Actually I'm going to compare 18 (!) methods- hand-crafted and bootstrapping; plus shrinkage with a shrinkage factor of 0% (which is equivalent to the naive method), 33%, 66% and 100% (which is equivalent to equal weighting); note because I allow different shrinkage factors on the mean and correlation that adds up to 16 possibilities.

## Results

All these graphs have the same format. The x axis shows which of the 18 methods are used. Note the abbreviations: H/C is handcrafted (which is also the benchmark), BS is bootstrapped and **Sa**b**** implies shrinkage of **a** on asset Sharpe Ratios and **b** on correlations; where 0,3,6 and 1 represent 0%, 33%, 66% and 100%. Naturally S00 is the naive method and S11 is equal weights. The y axis shows either the average out of sample SR versus a benchmark (S00), or the average degradation from in sample to out of sample (negative means returns got worse).

I plot the methods with shrinkage for the mean increasing as we go to the right (S11 - equal weights, BS and HC are on the extreme right; the naive portfolio S00 on the extreme left).

Naturally because this is random data I've generated a number of series of asset returns for each set of asset sharpe ratios and correlations; and the results are averaged over those.

There is some (messy) code [here](#).

### One year

B breakout . Not the classic home arcade game, seen here in Atari 2600 version, but what happens when a market price breaks out of a trading ...

[Can you eat geometric returns?](#)  
This post is about a slightly obscure, but very important, issue. Should we use geometric or arithmetic means of returns to evaluate investm...

[Placing orders in the native python IB API](#)  
This the fourth in a series of posts on using the native python API for interactive brokers . You should read the first , second , and thi...

[Streaming market data from native python IB API](#)  
This the third in a series of posts on using the native python API for interactive brokers . You should read the first , and the second , be...

[Using swigibpy so that Python will play nicely with Interactive Brokers API](#)  
The interactive brokers API is as far I know the only way that a non institutional client can access financial markets in a way which mak...

Things I wish interactive brokers would do with their API software

As regular readers know I use interactive brokers (IB) to run an automated futures trading system . Now in many ways IB are great. They hav...

[Obligatory Bitcoin post](#)  
No, I'm not dead. You'd be forgiven for thinking I was given my lack of blog posts in the last few months, but I was busy; first p...

[Correlations, Weights, Multipliers.... \(pysystemtrade\)](#)  
This post serves three main purposes: Firstly, I'm going to explain the main features I've just added to my python back-testing pa...

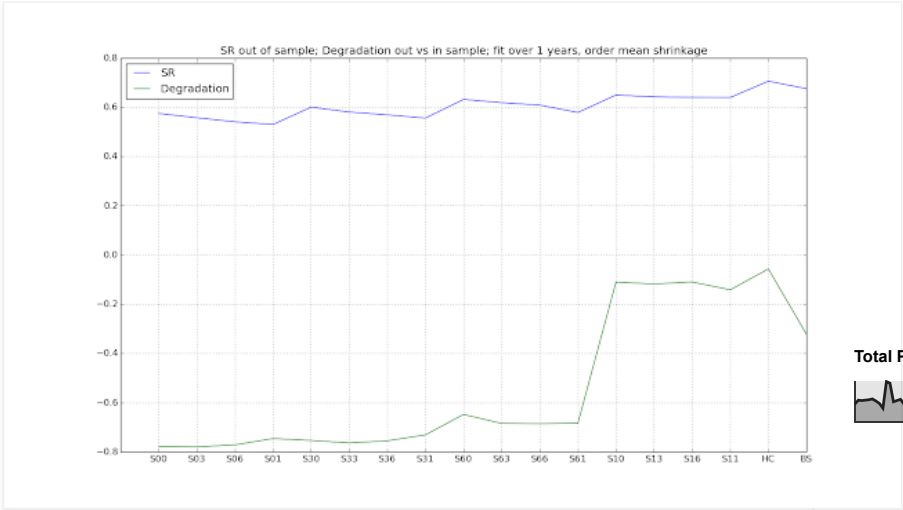
### Tags

[Systematic Trading](#) (38) [Python](#) (29) [Technology](#) (28) [Portfolio optimization](#) (26) [Interactive Brokers](#) (22) [pysystemtrade](#) (19) [Systematic\\_Trading\\_Book](#) (17) [Uncertainty](#) (17) [Statistics](#) (14) [Trend following](#) (14) [Investment idiocy](#) (11) [risk management](#) (11) [Systems building](#) (9) [Behavioural finance](#) (8) [Execution](#) (6) [Hedge funds](#) (6) [Random data](#) (6) [Volatility](#) (6) [Finance industry economics](#) (5) [High frequency trading](#) (5) [LOBO](#) (5) [Smart\\_Portfolios\\_Book](#) (5) [Conditional distributions](#) (4) [Equities](#) (4) [Kelly criterion](#) (4) [Novice investors](#) (4) [banks](#) (4) [carry](#) (4) [leverage](#) (4) [skew](#) (4) [sqlite](#) (4) [Git](#) (3) [Politics](#) (3) [housing associations](#) (3) [prediction](#) (3) [CAPM](#) (2) [Compound interest](#) (2) [ETF](#) (2) [Equity risk premium](#) (2) [Financial industry pay](#) (2) [asset allocation](#) (2) [bitcoin](#) (2) [handcrafting](#) (2) [ATR](#) (1) [Capital correction](#) (1) [Commodities](#) (1) [Docker](#) (1) [Factor models](#) (1) [Partial Correlations](#) (1) [QuantCon](#) (1) [Quantitative Easing](#) (1) [brexit](#) (1) [geometric returns](#) (1) [implicit fitting](#) (1) [kurtosis](#) (1) [leveraged trading book](#) (1)

### Quantocracy - featured blog



### Top 60 UK investment blogs

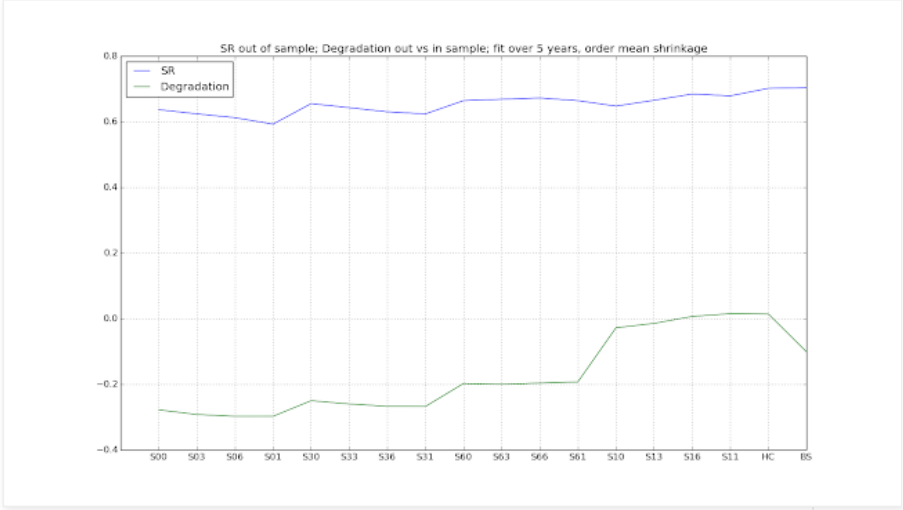


Total Pageviews  
 1,224,394

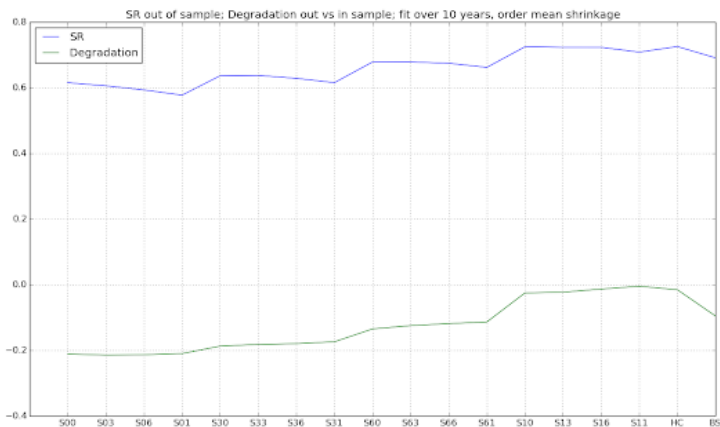
The results above are from using just one year of data. Adding shrinkage to correlations seems to make things slightly worse, but shrinking the means improves things. The handcrafted method does best, with bootstrapping coming in second. The scale is a little misleading; going from the worst to the best method is an improvement in SR from around 0.53 to 0.68.

Unsurprisingly without sufficient shrinkage on the means there is massive degradation in the Sharpe Ratio going from in sample to out of sample. Look at S00 the naive method. Out of sample it has an average SR of 0.58; in sample it is 0.78 higher than this, at 1.36! Using a non robust optimisation method over such a short period of data is going to give you seriously .

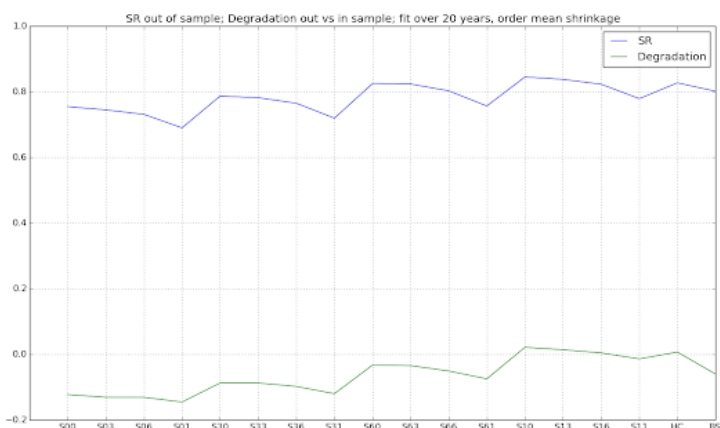
Five years



Ten years



## Twenty years



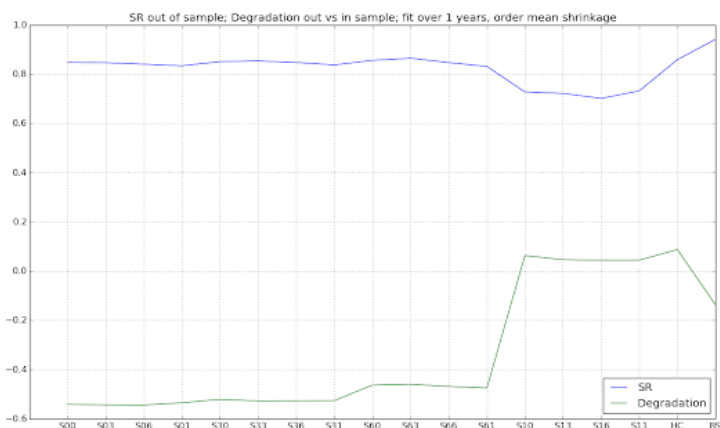
With twenty years of data the naive method is doing a little better. Shrinking the correlations definitely penalises performance - we now have enough data to be confident that the correlations are significant. However there is still a benefit from shrinking the means; even completely away. The advantage of handcrafting and bootstrapping has been whittled away, but they are still amongst the best methods.

There is still a degradation in performance going out of sample, but it is much smaller than before.

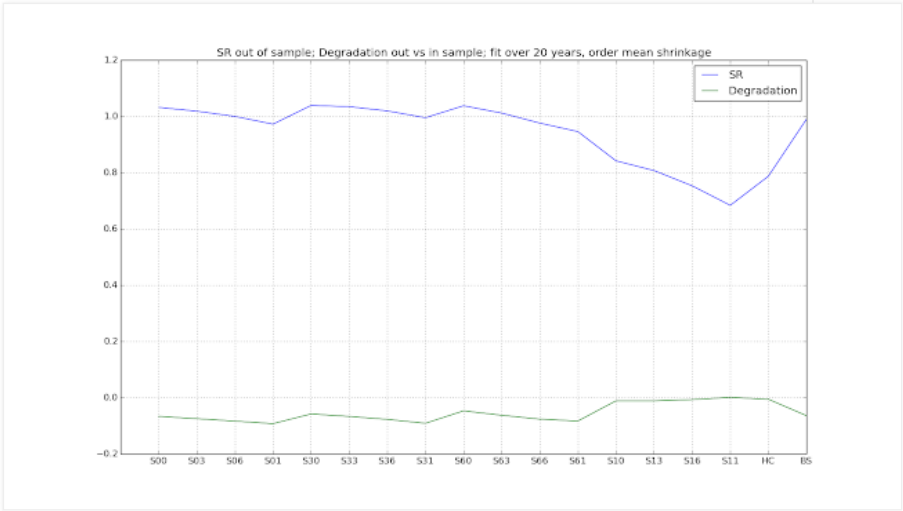
## Postscript

To an extent these results are a function of the underlying portfolio. If we ran these tests with a portfolio that had significant mean differences then shrinking the mean wouldn't be such a good idea. Here for example are the results with the same correlations as before, but Sharpe Ratios of [0.0, 0.5, 1.0].

First one year:



Now after twenty years:



Here the optimal shrinkage for the mean seems to be between 33% and 66%. Handcrafting, which in the simple form here does not account for differences in sharpe ratios, doesn't do as well as bootstrapping which does. It also loses out once the shrinkage methods have enough data to use the difference in Sharpe ratios properly.

However we don't know in advance what kind of portfolio we have... and significant differences in correlations are more common than statistically different Sharpe Ratios.

Conclusion

I naturally have a soft spot for my preferred method of bootstrapping and handcrafting. Shrinking can be a good alternative, but it's hard to get the shrinkage factor correct. In general you need to shrink mean estimates more than correlations, and shrink more when you have less data history. Using insufficient shrinkage; or none at all with the naive method, will also lead to massive degradation from in sample to out of sample returns.

This was the final post in a series on using random data.


First post: [Introducing random data](#)  
Second post: [Does equity curve trading work?](#)

Posted by [Rob Carver](#) at 10:35  
Labels: [Portfolio optimization](#), [Python](#), [Random data](#), [Statistics](#), [Systematic\\_Trading\\_Book](#), [Uncertainty](#)

No comments:

Post a Comment

Enter your comment...

 Comment as: handsomehu1f ▾

Sign out

Publish

Preview

☐ Notify me

[Newer Post](#) [Home](#) [Older Post](#)  
Subscribe to: [Post Comments \(Atom\)](#)

Contact Me (Spam will be politely ignored)

Name

Follow by Email

Email address...




Submit

Subscribe To

Email \*

Message \*

Send

-  Posts 
-  Comments 