INDEX TRACKING

What are these? Dow Jones

Hang Seng

FTSE All Share

S&P500

They are all stock market indices, which, in an easy to digest form, tell you how the stocks (companies) represented in the index have changed in value over time.

If you want you can invest your money in index tracking funds ("trackers") which aim to reproduce the performance of the index over time, perhaps by investing in all of the stocks that make up the index.

Question

How are index values arrived at (i.e. what formula is used)?

If a fund invests in all of the stocks in the index in such a way that its investment in each stock mirrors index composition (e.g. if a stock makes up 10% of the index then it makes up 10% of the investment) then the fund is said to be following a "**full/complete replication**" strategy.

Full replication is possible – but as the number of stocks in the index grows it can be an expensive strategy in terms of transaction cost. This is because:

- stocks typically enter/leave the index at regular intervals and so the entire fund must be **rebalanced** as this occurs to mirror the index as it changes
- any new money that is invested in (or money taken out of) the fund must be spread across all stocks to mirror the index

The largest index I am aware of that is tracked using full replication is the S&P500.

Suppose though we do not wish to adopt full replication. Then in essence we can view the index tracking problem as a ***decision problem***, namely to decide the ***subset*** of stocks to choose so as to (hopefully perfectly) mirror/reproduce the performance of the index over time. We call the subset of stocks we choose a **tracking portfolio (TP)**.

Question

It is estimated that approximately 10^F dollars are invested in index tracking funds in the USA alone - what is the value of F?

Questions

Where does the money to invest in index tracking funds come from?

How else might people choose to invest in stock (equity) markets if not via index tracking funds?

Why do people choose to invest in index tracking funds?

FORMULATION

The formulation (mathematics) you are about to see below is very general. Moreover it is relatively up to date, only since around 2003 have such general approaches been known.

Suppose that we observe over time 0,1,2,...,T the value of N stocks, as well as the value of the index we want to track. Further suppose that we are interested in deciding the best set of K stocks to hold (where K<N), as well as their appropriate quantities. In index tracking we want to answer the question:

*"what will be the best set of K stocks to hold, as well as their appropriate quantities, so as to best track the index in the **future** (from time T onward)?"*

Our basic approach in index tracking is a historical look-back approach. To ask the historical question:

*"what would have been the best set of K stocks to have held, as well as their appropriate quantities, so as to have best tracked the index in the **past** (i.e. over the time period [0,T])?"*

and then hold the stocks that answer this question into the future.

Although we did not stress it at the time the approach followed in Markowitz mean-variance optimisation is also of this type:

- look into the (immediate) past for relevant data
- use that data to form a portfolio
- hold that portfolio into the (near) future

NOTATION

Let:

N be the total number of distinct stocks (companies) in which we can invest

K be the desired number of distinct stocks in the TP

ε_i be the minimum proportion, and

δ_i be the maximum proportion, of the TP that must be held in stock i if any of stock i is held

X_i be the number of units of stock i in the current TP

V_{it} be the value of one unit of stock i at time t

T be such that we have observed historical values for stocks and the index over the time period 0,1,2,...,T. The time T represents a decision point, a time at which we **may** switch from our current TP $[X_i]$ to a new TP.

I_t be the value of the index at time t

C be the total value of the current TP $[X_i]$ at time T plus any **cash change** in the portfolio (either new cash available for investment or cash being withdrawn at time T), i.e.

$$C = \sum_{i=1}^{N} V_{iT}X_i + \text{cash change}$$

γ be the limit on the proportion of C that can be consumed by transaction cost

Then our decision variables are:

x_i the number of units of stock i that we choose to hold in the new TP

z_i = 1 if any of stock i is held in the new TP
= 0 otherwise

Without significant loss of generality we allow $[x_i]$ to take fractional values.

CONSTRAINTS

The constraints associated with the index tracking problem are:

$$\sum_{i=1}^{N} z_i = K$$

$$\varepsilon_i z_i \leq V_{iT} x_i / C \leq \delta_i z_i \quad i=1,...,N$$

$$C_{trans} = \sum_{i=1}^{N} \text{transaction cost}(X_i \to x_i \text{ at time } T)$$

$$C_{trans} \leq \gamma C$$

$$\sum_{i=1}^{N} V_{iT} x_i = C - C_{trans}$$

$$x_i \geq 0 \quad i=1,...,N$$

$$z_i \in [0,1] \quad i=1,...,N$$

OBJECTIVES - TRACKING

In time period t we get a return associated with the index, $R_t = \log_e(I_t/I_{t-1})$, where we define return using continuous time.

If, in each and every time period, the return associated with the TP:

$$r_t = \log_e[[\sum_{i=1}^{N} V_{it} x_i]/[\sum_{i=1}^{N} V_{it-1} x_i]]$$

was EXACTLY equal to R_t then this might seem ideal.

A possible objective in terms of index tracking is therefore:

$$\text{minimise} \sum_{t=1}^{T} (r_t - R_t)^2/T$$

i.e. minimise average squared error

Minimising average squared error is common – if you have ever done a linear regression that the equations for the slope and intercept of the regression line actually minimise average squared error.

Here we have a formulation of the index tracking problem as a mixed-integer nonlinear problem.

To get some insight into how we might try and solve this formulation let us assume that:

- we have zero transaction cost, i.e. buying and selling stocks costs us nothing
- we have no limits on the proportion of the tracking portfolio in each stock (i.e. ε_i=0 and δ_i=1

The index tracking problem then becomes:

$$\text{minimise} \sum_{t=1}^{T} (r_t - R_t)^2/T$$

subject to

$$\sum_{i=1}^{N} z_i = K$$

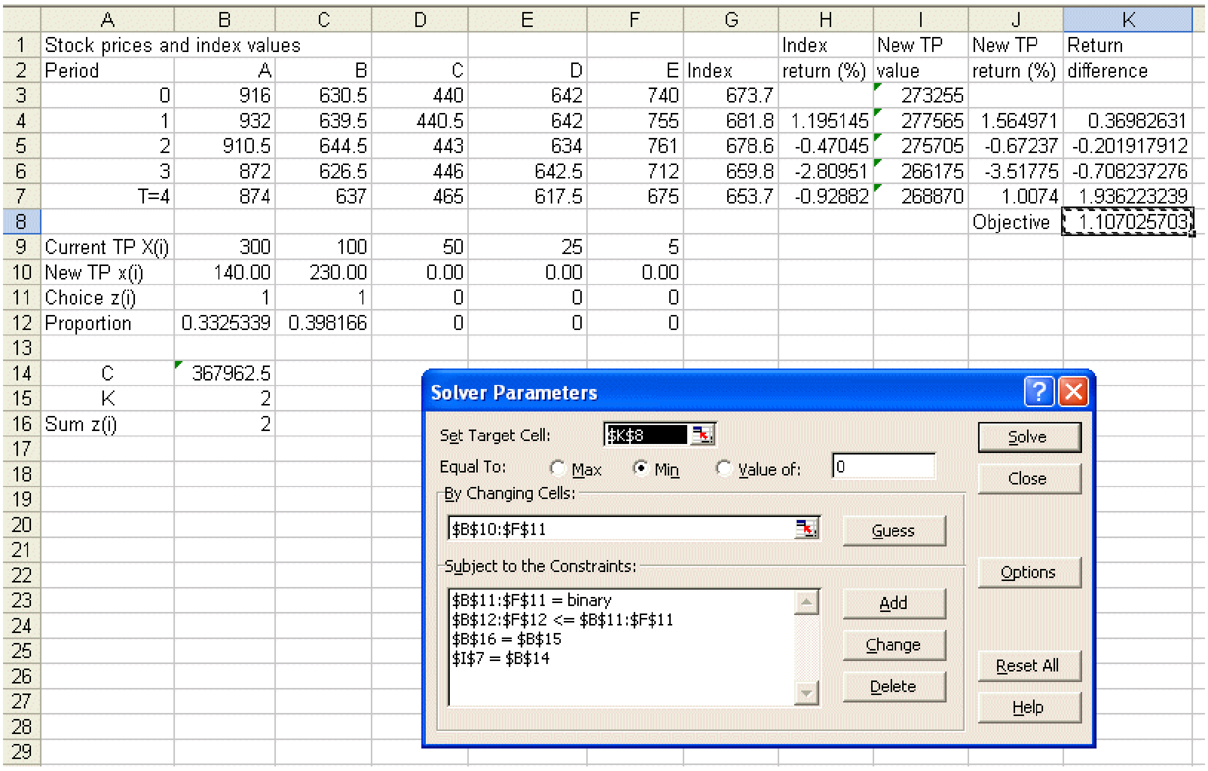
$V_{iT}x_i/C \leq z_i$ $\quad$ i=1,...,N

$$\sum_{i=1}^{N} V_{iT}x_i = C$$

$x_i \geq 0$ $\quad$ i=1,...,N

$z_i \in [0,1]$ $\quad$ i=1,...,N

The example below is the one we will deal with in class, with the Solver model for this example also being shown below



	A	B	C	D	E	F	G	H	I	J	K
1	Stock prices and index values							Index	New TP	New TP	Return
2	Period	A	B	C	D	E	Index	return (%)	value	return (%)	difference
3	0	916	630.5	440	642	740	673.7		273255		
4	1	932	639.5	440.5	642	755	681.8	1.195145	277565	1.564971	0.36982631
5	2	910.5	644.5	443	634	761	678.6	-0.47045	275705	-0.67237	-0.201917912
6	3	872	626.5	446	642.5	712	659.8	-2.80951	266175	-3.51775	-0.708237276
7	T=4	874	637	465	617.5	675	653.7	-0.92882	268870	1.0074	1.936223239
8										Objective	1.107025703
9	Current TP X(i)	300	100	50	25	5					
10	New TP x(i)	140.00	230.00	0.00	0.00	0.00					
11	Choice z(i)	1	1	0	0	0					
12	Proportion	0.3325339	0.398166	0	0	0					
13											
14	C	367962.5									
15	K	2									
16	Sum z(i)	2									

Again note that Solver provides a heuristic, not an optimal, solution.