A Handbook of Management Cybernetics. 1972. G. T Foulis, Henley-on-Thames

Investment Analysis-A New Approach

ROGER T. HARTLEY*

Introduction

When someone invests money on the stock-exchange, and this applies especially to institutional investors such as the giant international companies (and the government) lie expects to make more money than the value of his original investment To *offset* this golden vision of receiving large amounts of cash for doing absolutely nothing, he has given up (temporarily at least) the possibility of using those resources which form the investment. In other words, investment is simply 'pay now, live later'.

The obvious need for resources to start an investment gave rise to the old phrase 'you need money to make money', but this contains a suspicion of envy which is no longer valid. The rise of the Unit Trusts means that for the cost of a holiday anyone can 'play the market', albeit one step removed from more direct forms of investment. The Daily Mirror has a column devoted to news in the City of London, the stronghold of Conservatism (and conservatism). Investment is an industry in itself and wherever a large market is involved new and improved methods are bound to arise.

This is, of course, where investment analysis comes in. When large sums of money are involved, the best available techniques are employed; and so, as in many other fields, 'art' is replaced by 'science'. This simply means that more thorough, more costly, more time-consuming, more everything methods are used because the risks are also greater, and people hate to lose money.

However, the art of investment still holds forth in many quarters and some small piece of evidence for this is embodied in the following from Cohen and Zinbarg (1967) about American investors:

> The reader will probably acknowledge readily that most individual investors have little notion of what rate of return they have earned on their investment ... But most readers will probably find it difficult to believe that professional investors have almost as little knowledge as the proverbial man-in-the-street. Yet our experience suggests that this is truly the case.

This then sets the scene for this paper, which apart from presenting a short account of part of my own work, is also a brief survey of' the field of investment analysis seen from a scientific viewpoint. Perhaps before going any further the ambiguity present in the literature of the meaning of 'investment analysis' should be cleared up. Investment is usually described as either portfolio investment, which is the buying and selling of stocks and shares on the stock-exchange; or capital investment, which is the process

*Brunel University.

carried out within a company to enable it to expand. As Ginsberg (1968) says, and this makes the distinction very clear

> ... Portfolio investment being very often the means by which capital investment is financed ...

This means that, as mentioned earlier, many companies buy shares on the stock-exchange to provide extra financial backing for their enterprises.

A short historical survey

In the recent past, three works stand out from the rest in presenting new ideas; those of Markowitz (1959), Clarkson (1962), and Weaver and Hall (1967). These represent three widely differing approaches. Markowitz used mathematical analysis and optimizing techniques; Clarkson used simulation of the human approach and Weaver and Hall employed commonly-used statistical analysis methods. However, all three approaches can be thought of as starting from a basic model of investment. Markowitz's model is the simplest in order to fit his strict mathematical methods. Weaver and Hall's model is more complex although limited by the choice of statistical analysis (linear regression). Clarkson, however, concentrated on modelling tile decision procedures which parallel the numerical techniques outlined above as much as the static model of investment similar to that of Markowitz and Weaver and Hall. In other words the investment process can be split into two parts. The first is static and it states the factors on which share-prices depend as well as their relationship. This function can be manipulated in a number of ways to yield valid predictions of future share-prices. The techniques employed by the three authors mentioned are three ways of manipulating the function, but many others exist and some are still untried.

(1) *Markowitz*

Nearly all large investment is done in a portfolio of several individual investments. This not only spreads risk over the whole of the portfolio (which could contain 100 or more different securities) but in theory a higher return can be achieved over a long period of time by constantly changing the composition of the portfolio.

Markowitz attempted to answer the question 'which securities should be invested in and how much of an investment should be made in each?' He did this by assuming that security return (increase or decrease in share price) is a random variable subject to the laws of probability. The simple formulation of the problem led to one of maximizing return subject to various constraints and Markowitz devised an algorithm to solve the resultant quadratic equations. A typical solution to select 60 securities from a list of 300 takes less than 1.5 min on a modern high-speed computer.

The mathematical techniques are extremely elegant and highly efficient but tend to obscure the issue somewhat. The whole basis of the method is that probabilities have to be assigned to growth and yield forecasts and there are largely subjective or at least are until an optimal method is devised for obtaining them. In other words, although accuracy is obtained in the selection process the inadequacies of prediction and data collection methods are highlighted even more strongly.

(2) *Clarkson*

In many ways Clarkson's approach to investment is entirely opposite to that of Markowitz. Ile was concerned solely with reproducing the actual behaviour of one trust investment officer. This he did by observing him at work, extracting the basic procedures and simulating them in a computer program. The whole thing is extremely enlightening and fie succeeds in apparently refuting several theories of investment current at the time. An interesting point here is that any theory of investment is also a theory of human problem-solving and decision -making. Clarkson's success in simulating one particular piece of behaviour had implications in these fields and was and is a supremely sensible way of investigating theories of human activities in general. His main conclusion in this area is summed up neatly by one of his own sentences:

> It is our suggestion that the weighted decision functions and the maximizing hypothesis of current theory be replaced by processes that employ binary tests as the basis upon which decisions are made and patterns recognised.

As far as investment analysis is concerned Clarkson's work is only interesting in that it brings out some of the procedures actually used; however, it provides no direct answer to the question of optimal portfolio. Presumably the investment officer learned his craft by trial and error over many years of practical experience and in this sense the portfolio produced is optimal, but is limited by the human processes outlined in the quote above.

(3*) Weaver and Hall*

Weaver and Hall start their work to a derivation of a relationship between share prices and share data and to this end they use the well-tried method of multiple regression analysis. The predicted variable is the dividend yield of the share and the dependent variables (the 'predictor' variables as they call them) are:

(1) Dividend payout ratio.
(2) Forecast earnings growth rate -short-term.
(3) Forecast dividend growth rate-long-term.
(4) Historical earnings variability.
(5) Historical earnings growth rate.

It can be seen straight away that variables 2 and 3, being themselves forecasts, have to be obtained subjectively whereas the others can be obtained directly from past and present figures.

The exact formulae used for calculating these factors are not really relevant; suffice to say that they have been carefully chosen over a long period of time and are not based on any strict theory of the stock-market beyond that of the regression model. However, they are, say Weaver and Hall:

> ... probably the main ones uppermost in investors' minds when ordinary shares are evaluated ...

This is of vital importance since the statistical techniques, such as regression, used in much similar work, mathematically satisfying as they arc, rely heavily on having well-organized, vvell-thought-out data to provide meaningful results. Many erroneous conclusions have been reached based on false statistical arguments. I feet that a larg
proportion of this is due to placing too much reliance on the limited semantics which much of statistical processes.

The program based on Weaver and Hall's method has been operating successfully in a commercial (i.e., stockbroking) environment for a number of years and is easily the most sophisticated of its kind in Britain, if only because of the amount of work put into the whole project. According to their published results, the computer analysis out-performs other much-used investment analysis procedures and is consistently much better than the average, as embodied in the F.T. index.

(4) *Other work*

More recent work can be split into main groups one of which contains the other.

The first and by far the simpler is called technicalism. This assumes that share prices can be predicted by analysis of two stock-market variables: historical share-price and historical turnover of shares. Working on the hypothesis that history repeats itself it is almost a problem in pattern recognition-the patterns being trends, cycles, etc., in the two variables mentioned. A rneasure of the reliability of this apparently naive method is that one investment company in America bases its portfolio management on a 'technical' computer program to which the method is readily adaptable.

The other main approach is called fundamentalism and is the attempt to take into account every relevant factor from micro-economic sources. It is essentially a model-building exercise which yields predictions of share prices from the interactionof several models. The macro-economic models take into account national and international, accounts, working down through industry to individual industries. Micro-economic models analyse individual companies on the basis of annual and quarterly reports. Needless to say fully automated fundamental analysis is almost inconceivable given the problems of collection and organization of data. However, several programs of this type are in existence and despite the problems involved valid predictions must be forthcoming.

One thing is absolutely clear in all this-worthwhile investment analysis is only possible with the computer as aid. The full scope of fundamental analysis is formidable in comparison with the process described by Clarkson as typical of trust investment by one man. A strong point in the computer's favour is that a lot of investment analysis is routine work carried out, say, every week, i.e., collection and organization of raw data and this sort of task is ideally suited to our electronic slave. Furthermore, model-building--the computer's forte--is increasingly coming into use as a necessary and useful technique for decision-making of all kinds.

I think it should be stressed however, that computer techniques, enormously efficient though they are, should be considered purely as aids-electronic slide-rules if you like-until the computer methods become a lot more sophisticated. The final decision must always rest with a human being-even if sometimes he is wrong.

The new approach--program COIN

COIN is a COmputerized Investment program which predicts the relative future performance of a number of shares on the Stock Exchange, according to some pre-chosen criterion.

Its input is in the form of quantitative or qualitative data on 10 variables which affect the share's performance, for each of 10 discrete time periods in the share's history; the eleventh variable is, of course, the share prices, again for the same 10 periods.

The output consists, essentially, of an ordered list of the shares for which data was input, the order being that of relative performance over the next time-period. The top 10 or so could then form the basis of a portfolio which can be reviewed by repeating the same procedure (with new data) at the end of each period.

What follows is a mathematical formalism of the working of the program, and then a discussion of the choice of variables and parameters.

Procedure formalism

Since the basis of the whole program is comparison we need only consider two shares X and Y. The predicted variable is P and its value p. p is assumed to be a function of a number of variables, numbered I to n which can be considered as forming a vector.

Thus: $p_x = f(\mathbf{X})$ where $\mathbf{X}$=(x1,x2.... Xn)

$p_y = f(\mathbf{Y})$ where $\mathbf{Y}$ = (y1, y2, . - - Yn)

The functional form is assumed to be linear, not only for the trivial reason of simplicity, but since binary comparison of variables is used throughout, any higher order function would seem to be unnecessary.

Thus we can assign a set of weights:

$\mathbf{W}=(w_1, w_2, ---, w_n)$ to the *comparison of* **X** and **Y**.

(For comparison purposes the weights are normalized $\sum_{1}^{n} w_i = N$

The comparison of **X** and **Y** is initiated by forming the vector **C** = **X** - **Y**. Again for comparison purposes w_i must be independent of x_i and y_i thus the vector $\mathbf{C}^1$ is formed by means of an operator *O*.

$\mathbf{C}^1 = O\mathrm{C}$ where *O* is such that $s_i^1 = \dfrac{s_i}{|s_i|}$

In other words $\mathbf{C}^1$ is a vector whose components only have values 0, + 1, -1.

The prediction is $\dfrac{px - py}{|px - py|} = \dfrac{p^1}{|p^1|}$ where $p^1 = \mathbf{W} \cdot \mathbf{S}^1$

Therefore prediction of relative performance,

$$R = \frac{W \cdot O(X - Y)}{|W \cdot O(X - Y)|} = +1, -1, 0$$

depending as X will do better than, worse than or as well as Y.

The programming problem is that of calculating the weights W so that predictions can be made.

To do this the program acts as a learning machine, changing weights when a wrong prediction has been made, until the whole of a training sequence of discrete time periods has been accurately described.

Assume that using the weights **W** a prediction R_T has been made on data $\mathbf{X}_T$ and $\mathbf{Y}_T$ i.e.

$$R_T = \frac{W \cdot O(X_T - Y_T)}{|W \cdot O(X_T - Y_T)|}$$

Assume further that the prediction is wrong (R_T is known since the data is all historical in the training sequence)

$$W \cdot O(X_T - Y_T) = \pm|U| = W \cdot S^1$$

Let

New weights $\mathbf{W}^1$ are to be calculated such that

$$W^1 \cdot S^1 = \mp|A| \quad \text{where A is a constant parameter.}$$

Clearly the prediction is made right so long as $\mathbf{W}^1$ can be calculated. We have

$$\sum_1^n w_i s_i' = \pm|U|$$

$$\sum_1^n w_i s_i' = \mp|A|$$

and

$$\sum_1^n w_i' = \sum_1^n w_i' \quad \text{since the weights are normalized}$$

We can split up each sum into three parts according as

$$s_i' \; is +1, -1 \; or \; 0 \; i.e.$$

$$\sum w_{i_1} s_{i_1}' + \sum w_{i_{-1}} s_{i_{-1}}' + \sum w_{i_0} s_{i_0}' = \pm|U| \qquad (1)$$

$$\sum w_{i_1} s_{i_1}' + \sum w_{i_{-1}} s_{i_{-1}}' + \sum w_{i_0} s_{i_0}' = \mp|A| \qquad (2)$$

and

$$\sum w_{i_1} + \sum w_{i_{-1}} + \sum w_{i_0} = \sum w_{i_1}' + \sum w_{i_{-1}}' + \sum w_{i_0}' \qquad (3)$$

since

$$s_{i_1}' = +1 \cdot s_{i_{-1}}' = -1 \cdot s_{i_0}' \quad \text{we have from (1) and (2)}$$

$$\sum w_{i_1} - \sum w_{i_{-1}} = \pm|U|$$

$$\sum w_{i_1} - \sum w_{i_{-1}} = \mp|A|$$

From these two we get, by subtraction

$$\sum (w_{i_1} - w_{i_1}') - \sum (w_{i_{-1}} - w_{i_{-1}}') = \mp(|U| + |A|)$$

let

$$w_{i_1} - w_{i_{-1}}' = K \; and \; w_{i_{-1}} - w_{i_{-1}}' = L \qquad \text{(4) and (5)}$$

Then $K \sum s_{i_1}' - L \sum s_{i_{-1}}' = \pm(|U| + |A|)$ since $\sum s_{i_1}'$ is the number of terms w_{i_1} and the same for $\sum s_{i_{-1}}'$

From (3) we get $\sum(w_{i_1} - w'_{i_1}) + \sum(w_{i_{-1}} - w'_{i_{-1}}) + \sum(w_{i_0} - w'_{i_0}) = 0$

i.e. $K\sum s_{i_1} + L\sum s_{i_{-1}} + \sum(w_{i_0} - w'_{i_0}) = 0$

if we put $w_{i_0} = w'_{i_0}$ i.e., the weights for which the variables in **X** and **Y** are the same remain unchanged, then we can solve for *K* and *L* from:

$$Kn_1 - Ln_{-1} = \pm(|U| + |A|) \quad n_1 = \text{ number of +1's in a'}$$

$$Kn_1 + Ln_{-1} = 0 \quad n_{-1} = \text{ number of -1's in a'}$$

$$K = \pm\frac{(|U| + |A|)}{2n_1}$$

$$L = \mp\frac{(|U| + |A|)}{2n_{-1}}$$

Thus from (4) and (5)

$$w'_{i_1} = w_{i_1} \mp \frac{(|U| + |A|)}{2n_1}$$

and

$$w'_{i_{-1}} = w_{i_{-1}} \pm \frac{(|U| + |A|)}{2n_{-1}}$$

The choice of sign depends on whether X is better or worse than Y over the last period.

This then gives an algorithm for changing the weights when a wrong prediction has been made. Clearly it only puts right one prediction at a time and may well (and often does) throw some of the previous predictions out, i.e., some predictions which were correct with the old weights are Wrong with the new weights. The solution to this is to cycle through the history, putting right any wrong predictions as they occur until the whole of the history can be predicted with one set of weights. Thus the method is an iterative one and. so merits the term learning machine. The training method (the weight-changing procedures outlined above) can be shown to converge (i.e., any history can be so described with one set of weights) and a formal proof can be found in Nillson, Chapter 5 (1965).

The above is the basic computational unit, i.e., the comparison of any two shares. The end result is a prediction of their relative performance over the next period. This unit is then utilized in the prediction of the performance of a whole set of shares.

Having discussed the formalism it remains to state how and why values were chosen for all the parameters, variables and constants which are essential for the running of the program on a computer. These can be listed as follows:

(1) Performance criterion.
(2) Number of valuables per share, and which variables.
(3) Number of periods and length of each period in history.
(4) Normalization constant for weights.
(5) Value of constant A.

(1) The performance criterion, i.e., the method of comparing two shares' performance, is not a point oil which to delay long. The object in buying shares is to make money by selling them when their price rises enough to make a satisfactory profit. Whether the dividend is included in the profits depends on the period of investment and a number of other factors but no one likes it if the price drops. Therefore as a sound basis for any more complex criterion increase in share-price over the last period considered was chosen. In other words the critical variable which is compared for different pairs of shares is:

$P_T - P_{T-1}$ where P_T is the share-price at time T.

(2) Which variables to choose and how many should be the subject of much experimentation once the program is running. Neither question is easily answered without seeing the program's performance at first hand, but to start with ally of the existing investment analysis models could be used. The only direct comparison available is with Weaver and Hall's (1967) successful program which uses the following six variables:

(1) Dividend yield.
(2) Dividend payout ratio.
(3) Forecast earnings growth rate-short term.
(4) Forecast dividend growth rate-long term.
(5) Historical earnings variability.
(6) Historical earnings growth rate.

James Morell (1968) suggests the following I I variables:

(1) Dividend.
(2) Earnings record.
(3) Dividend cover.
(4) Volatility of earnings.
(5) Dividend yield.
(6) Earnings yield.
(7) Management.
(8) Politics.
(9) Assert value/share.
(10) Expected growth in dividend.
(11) Expected growth in earnings.

It can be seen that Weaver and Hall concentrate on the two 'technical' variables, dividend and earnings which are quantitative and readily available for every share. Morell adds, besides the asset value per share, the two qualitative variables management and politics. Clearly these are of supreme importance (especially in, say, a take-over situation) but are equally, difficult to quantify. However, as stressed right from the start, there is no need to quantify when only comparisons are used, or at least quantification is made much easier. Comparisons involving management and politics can be made by the time-honoured method (subjective opinion).

Since the program will throw out any useless variables (by assigning a low or zero weight) the correctness of the subjective comparison (or at least its usefulness) can soon be discovered.

Allowance has been made in the program for 10 variables but this call be extended with little difficulty. An obvious point which it is, however, worth while to state, is that the useful variables are going to be those which are directly comparable, irrespective of which companies are involved. There may be cases where this is not the case, e.g., it may well be that having a large board of directors is better than a small one. This clearly puts a premium on the overall size of the company, but in general variables like these should be made directly comparable by suitable means, e.g., expressing the number of directors as a percentage of the number of employees.

The final point concerning the number of variables conics from Ivakhnenko and Lapa (1965). They suggest that 'secondary' feedback be applied (as well as the learning feedback) to select the most useful set of variables from a large set initially presented. However, as previously mentioned, this exists in minor form already. To explain, once, as a consequence of weigh t-chan ging, a variable weight reaches zero its stays there for the rest of the learning period; in other words, that variable is rejected. This process is only possible because of the insistence on the normalization of the weights (see (4) below).

(3) The form and amount of the historical data must be seen as a compromise between the optimal requirements of the program and the practicalities of data collection and client requirements. The best length of the history (i.e., number of periods considered) depends, in theory, on whether the process is stationary or non-stationary.

In fact, all processes are non-stationary but some are more stationary than others. This means that the average of the prediction variable over an ensemble is always time-dependent, but in some processes the rate of change is sufficiently slow to be neglected. These are stationary processes. It turns out that the more stationary the process the longer is the optimal (see Ivakhnenko and Lapa).

Since comparison is the basis of this method we have the ideal stationary process arid so a long history is better. The practical problems of using a long history are obvious, the main one being storage space which is expensive. The program handles 10 periods in the history but, again, this can be extended easily if necessary.

The actual time-period used depends purely on the practicalities of the client's requirements. If the client is prepared to buy and sell at intervals of two weeks then the time-period should be fortnightly and so on. If he has no idea of this buying/selling period then the smaller the time-period the better, although this means increased problems in data-collection and so on. No final decision has been made on this point, since it is not necessary for data-file organization or the program itself.

(4) and (5) The normalization constant (i.e., the sum of all the weights) is a compromise between accuracy and running time. The larger the number the better the accuracy, i.e., there is more resolution between individual weights, but there is also increased complexity in the arithmetic involved in the weight-changing processes. Another effect to note is that arithmetic rounding on division can lead to unintentional results (based on direct experience!) and so the constant should be large enough to get over this satisfactorily. It was decided to choose an initial value of 100 for each of the 10 weights, giving a constant sum of 1000.

The value of the parameter A clearly affects the time of the learning **period and it would** seem that the smaller the value the shorter the time in learning to describe the

history. In other word-, since A is some sort of 'search distance' over and above the minimum required it would seem unnecessary to search too far. However, this was found not to be the case and the inference is that the learning time depends more on the complexity of the history being described than the value of A.

A is fixed constantly at 100.

COIN-performance and problems

The success or failure of COIN depends very heavily on the form of the data with which it is presented. If there are any inconsistencies in the data then the program breaks down and produces no prediction (other than an arbitrary one-in other words it always produces *something).* This can be paralleled in any learning situation where a disturbance or inconsistency in the environment tends to degrade the learning process or even stop it altogether. This brings us to what I believe is the central problem in investment and portfolio analysis-the choice of descriptor variables. In other words, which variables explain the movement in share prices?

This, of course, brings us right back to the split between building the model and then manipulating it as referred to in the survey. Many of the approaches adopted previously have tended to concentrate on die manipulation phase and leave the first phase, the model building, as an implicit reference. This also reflects a lack of agreement on a workable theory of stock-market procedures and performance. However, much work has been done using statistical techniques on analysing correlations between share-prices and various pre-chosen variables (especially Rayner and Little (1966)).

All of - this work seems to leave the central problem untouched-, a useful model is as far away as ever.

The answer is, I feel, not in applying one approach rigidly but in using a combination of many techniques in use in other fields. Feature extraction and the cognitive approach in general seem to offer much more scope than other methods in providing an understanding of the subject of investment as a whole.

The problem of building a useful model has been approached with COIN in mind. It is hypothesized that the data for COIN can be manipulated using a combinatory logic method (a modification of Boolean algebra) in such a way that the performance of COIN becomes optimal-it will almost certainly never be perfect. Marked success has been achieved by this technique, the success rate going up from 35% to a maximum so far of 70%. Thus the quality of data has been improved from not very good to quite good with minimal change. In fact the only changes made in the data are to reverse some of the correlation signs of the 10 variables chosen (they are all assumed positive a priori). However, the technique clearly has a maximum success rate (it is [4] narrow-minded') and a lot more can be done with one set of data in the way of general combinatory functions. The approach here can be seen to fit Ivakhnenko's secondary feed-back'; I am in fact attempting to make the selection of variables and their relationships (and thus building the model) automatic, given a data-base to work on.

The future is very bright, since the preliminary results are encouraging. There is no doubt that the field of investment is ripe for examination using methods used in other fields (in a word-cybernetic methods) and interesting and useful results are bound to occur.

Acknowledgments

Thanks are due to Professor F. H. George for providing the initial interest, and the S.R.C. for providing the Research Studentship under which this paper was written.

References

Clarkson, G. P. E. (1962). *Portfolio Selection: A Simulation of Trust Investment.* Prentice-Hall.

Cohen and Zinbarg (1967). *Investment Analysis and Portfolio Management. Wiley.*

Ginsberg, L. (1968). *The Meaning of Investment in an Age of Plenty. Wiley.*

Ivakhnenko and Lapa. (1965). *Cybernetics and Forecasting Techniques.* Elsevier.

Little, 1. M. D. and Rayner, A. C. (1966). *Higgledy Piggeldy Growth Again.* Blackwell.

Markowitz, 11. M. (1959). *Portfolio Selection: Efficient Diversification of Investments.* Wiley.

Morell, J. (1968). *An Approach to Short-Term Business Forecasting.* Admin. Staff Coll.

Nillson. (1965). *Learning Machines.* McGraw-Hill.

Weaver, D. and Hall, M. G. (1967). Evaluation of ordinary shares using a computer. J. *Inst. Actu.*, **93** (reprinted in *Investment Analysis and Portfolio Management* (ed. B. Taylor) Elik. 1970.)