OMEGA Int. J. of Mgmt Sci., Vol. 15, No. 6, pp. 489–494, 1987
Printed in Great Britain.

0305-0483/87 $3.00 + 0.00

Application of the Knapsack Model for Budgeting

SAMUEL EILON

Imperial College of Science and Technology, UK

(*Received April 1987; in revised form May 1987*)

The budget problem of selecting projects (or activities) with known values (or payoffs) and associated costs, subject to a prescribed maximum budget, is akin to the knapsack problem, which is well documented in the literature. The optimal solution to maximise the total value of selected projects for a given budget constraint can readily be obtained. In practice, budgets are often somewhat flexible, or subject to possible changes, so that an optimal solution for a given budget value may not remain optimal when the budget is modified. It is, therefore, sensible in many situations to consider a *budget range*, instead of a single budget value. In addition to their original objective of maximising the total value of selected projects, decision makers are often concerned to get 'value for money', indicated by the ratio of payoff to cost. This paper examines how these questions can be tackled through the introduction of a *stability index*, to guide project selection within a defined budget range, and the use of a portfolio diagram, to help in the ranking of projects with respect to the stated twin objectives.

INTRODUCTION

THE KNAPSACK MODEL is well documented in the literature. It concerns the problem of selecting objects from a given array, each object i having a capacity requirement C_i and a value Z_i (or C and Z for short) to be loaded into a container (or a 'knapsack') of known capacity B. The purpose of the selection is to maximise the total value of the selected objects, provided that their total capacity requirement does not exceed the available capacity B of the container. The mathematical formulation of this model can be found in most textbooks on operational research and it is not intended to repeat it here. It is classed as a 0–1 programming problem, in that each object is associated with a variable that can assume the value of either 0 (when the object is not selected) or 1 (when it is), and efficient codes are now available to solve even problems of a moderate size on personal computers. The solution method generally relies heavily on the so-called 'greedy algorithm', which gives preference to objects with relatively high Z/C ratios (i.e. value per unit space requirement), and selection in descending order of Z/C is the basis of many developed heuristic methods.

The knapsack model has many applications, and a useful example occurs in the field of budgeting. Suppose that a selection needs to be made from n possible projects, each having a project value of payoff Z_i and involving a known cost C_i, subject to an overall budget constraint B. If a project can either be selected or rejected (i.e. partial projects are not allowed), then this budgeting problem may be formulated as a knapsack model, as described above. This type of budgeting problem is commonly found in practice: selection of production activities, R&D projects, investment portfolio, computer applications, advertising campaigns, and so on. In all these cases selection is necessary because of the limitation on available budgets, and the executive's purpose is to derive the maximum value, or payoff, from the selected projects.

SENSITIVITY OF THE SOLUTION

The knapsack problem is usually presented as a deterministic model, where the values Z, the costs C and the capacity constraint B are strictly defined. Sensitivity analyses can be carried out to ascertain the possible effect of deviations from the given data. In problems of budget allocation it is often the case that the budget is not rigidly stated but is subject to some flexibility, for example by being defined as a range. Alternatively, the budget may be fixed for one period with the prospect of being subsequently enlarged, so that if projects (or activities) cannot be easily switched, the optimal solution in the first period may well be incompatible with the optimal solution that would apply later. Thus, while a solution can easily be found for each point in a given budget range, it is not clear how the many possible solutions can be reconciled.

The following example will illustrate these points. A selection needs to be made from 12 projects, the values and cost of which are given in Table 1, where—for convenience—the projects are arranged in descending order of Z/C.

The number of possible combinations of projects, even within a fairly narrow budget range, can be very large. For example, if a budget range of 600–650 is considered, there are 391 solutions, the values of which are shown as a widely varied mass in Fig. 1. As suspected, the total value of the selected projects tends to increase with budget availability, and this can be perceived from Fig. 1, where all the solutions are grouped into five sub-ranges for the budget, namely: 600–610, 611–620, . . . , 641–650. If for each sub-range, the maximum, the average and the minimum values of all the possible project selections are determined (the results are presented in Fig. 2), they all show an upward trend, as the budget increases.

If the optimal solution is sought for every budget value in the range 600–650, i.e. if the knapsack model is solved 51 times (only integer values are considered), the resultant use of the budget (i.e. total cost) is shown in Fig. 3. The step-wise curve indicates that, throughout the range, the budget is often under-utilised. For

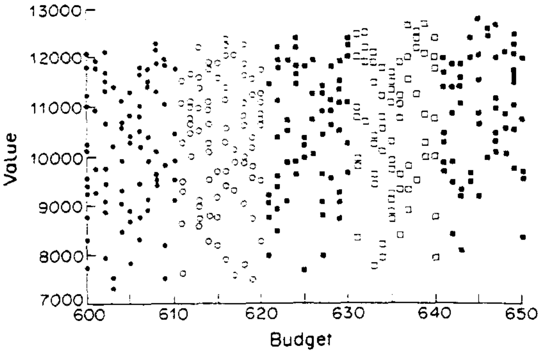

Fig. 1. A scatter diagram of solutions.

example, if the available budget is 616, it is fully used (the cost is also 616), but when the budget increases to 617, 618, etc., the amount used remains 616, until the available budget reaches 624, when the cost jumps to 624 and the budget is again fully used. This is further shown in Fig. 4, where the utilisation is calculated as the ratio of the budget used to the budget available. Out of 51 budget values in the range, only ten result in full utilisation. The optimal value attained also assumes a step-wise function, as shown in Fig. 5.

The change in the level of the budget used is reflected in the project selection, and Fig. 6

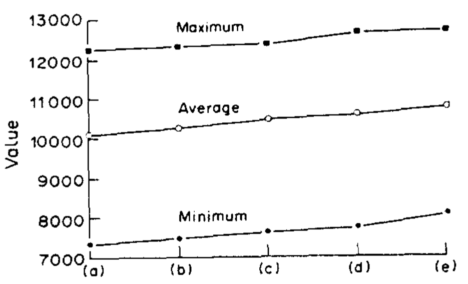

Fig. 2. Trends of values within the budget range. (a) 600–610; (b) 611–620; (c) 621–630; (d) 631–640; (e) 641–650.

Table 1. An array of 12 projects

No.	Value Z	Cost C	Ratio Z/C
1	4113	131	31.40
2	2890	119	24.29
3	577	37	15.59
4	1780	117	15.21
5	2096	140	14.97
6	2184	148	14.76
7	1170	93	12.58
8	780	64	12.19
9	739	78	9.47
10	147	16	9.19
11	136	22	6.18
12	211	58	3.64

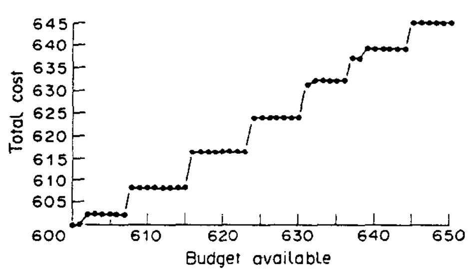

Fig. 3. Total cost vs budget.

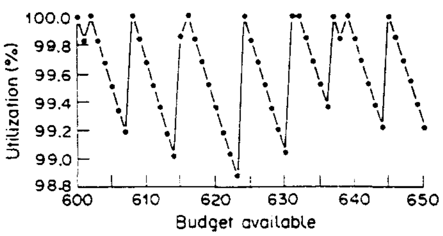

Fig. 4. Budget utilisation.

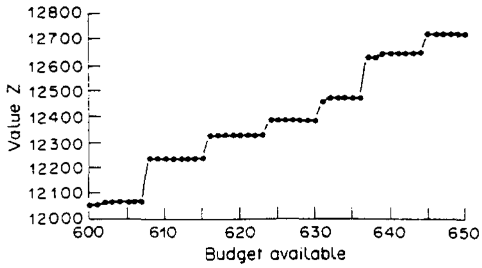

Fig. 5. Maximum attainable value for a given budget.

shows which projects are included in the optimal solution for each value of the budget range. For example, if $B = 600$, the selected projects are 1, 2, 4, 5 and 7, giving the total value of 12,049. This selection remains valid for $B = 601$, but at 602 the optimal portfolio changes to 1, 2, 5, 6 and 8, and this selection remains valid up to 607, but changes at 608, and so on. As the budget increases, projects get selected and deselected, as shown in Fig. 6. Throughout the range 600–650 projects 1 and 2 are always included in the optimal solutions, whereas projects 9, 11 and 12 are never chosen; the selection of the others depends on the prevailing budget value.

Stability index

The results in Fig. 6 lead us to introduce the concept of a *stability index*, to indicate the degree to which the inclusion of a project in the selected portfolio remains valid throughout a given budget range. The index simply reflects the frequency with which the project appears in the optimal solution.

For example, for the budget range 600–610 (11 points) the frequencies of projects 3, 4 and 5 are (see Fig. 6) 3, 5 and 11 respectively, or have a stability index of 27%, 45% and 100% respectively. Various likely budget ranges may be considered. For example, if in the budget in our example an initial budget of 600 is likely to be extended up to 610, 620, etc., then the budget ranges 600–610, 600–620, etc., need to be considered, and the stability index for each project for the alternative ranges can easily be computed.

The results are presented in Fig. 7, so that for any known budget range the descending order of stability index provides a first order of priority for selecting projects within the range. For example, if the budget range to be considered is 600–650 (range E in Fig. 7) then the projects may be ordered (in descending frequency) as 1, 2, 3, 8, 4, 5, 6, 10 and 7 (the others are never selected), as a basis for the final selection decision. If this order is followed, then the initial portfolio selection would consist of projects 1, 2,

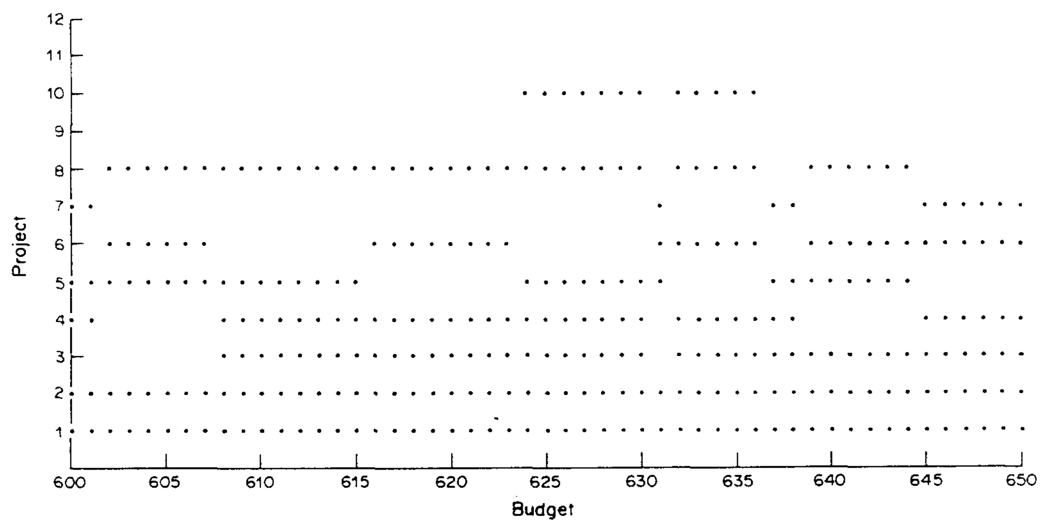

Fig. 6. The optimal solutions in the budget range 600–650.

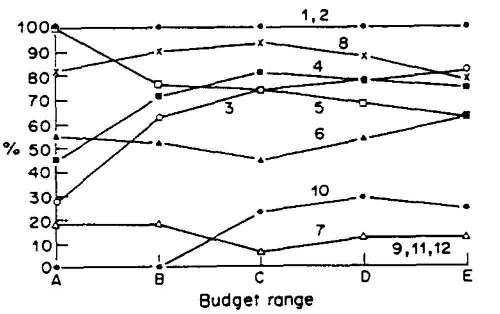

Fig. 7. Project frequency for various budget ranges. A, 600–610; B, 600–620; C, 600–630; D, 600–640; E, 600–650.

3, 8, 4 and 5 (or 6, since 5 and 6 have equal rankings in Fig. 7) with a total cost of 608 and total value of 12,236.

'*Value for money*'

Further projects with minor rankings can be added to this initial portfolio, and this will increase the total value by increasing the total cost. It is interesting to note, though, that while the total value of the optimal portfolio increases with the budget available, as shown in Fig. 5, the ratio of the global Z/C (which may be interpreted as 'value for money', namely the average value derived per unit of budget available) is not a monotonic function, as shown in Fig. 8. The general trend is for the overall Z/C ratio to decline with increased budget availability, but occasionally the ratio increases, suggesting that local preferences need to be considered (for example, at budget values of 608 and 637 in Fig. 8).

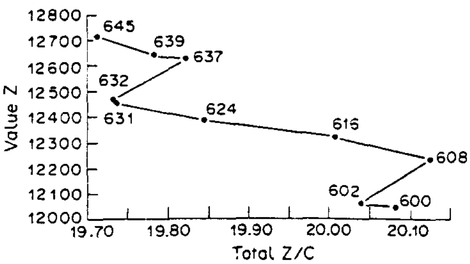

Fig. 10. Value Z vs Z/C.

This is further demonstrated as follows: if a base budget of 600 is postulated corresponding to a total value of $Z_0 = 12{,}049$, and if a budget increment of Δ is contemplated, then the incremental Z/C is computed as $(Z - Z_0)/\Delta$ (i.e. the ratio of the increment in Z to the increment in the budget). The results for the incremental Z/C (for values of $\Delta = 0$ to 50) are shown in Fig. 9 and indicate that the best incremental Z/C is obtained at a budget of 608.

If, in the budget range 600–650, we take the ten points at which the budget is fully utilised and compute for them the optimal values Z and

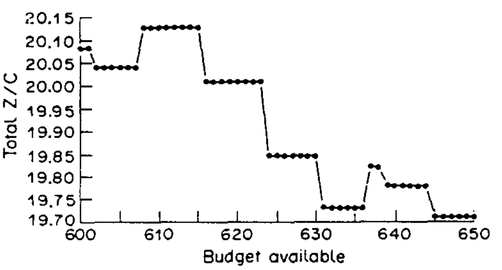

Fig. 8. 'Value for money'

Table 2. Ten optimal solutions with full budget utilisation

Budget	Total value Z	Total Z/C
600	12,049	20.08
602	12,063	20.04
608	12,236	20.12[a]
616	12,324	20.01[a]
624	12,383	19.84[a]
631	12,453	19.74
632	12,471	19.73
637	12,626	19.82[a]
639	12,640	19.78[a]
645	12,714	19.71[a]

[a] Points on the dominant frontier in Fig. 11.

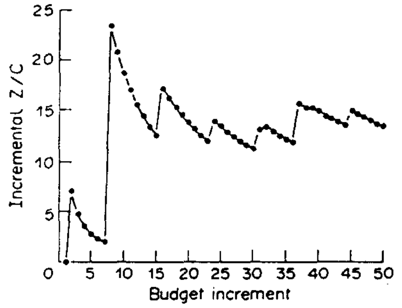

Fig. 9. Incremental Z/C for increased budget.

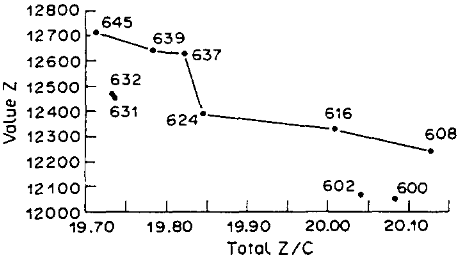

Fig. 11. The dominant frontier.

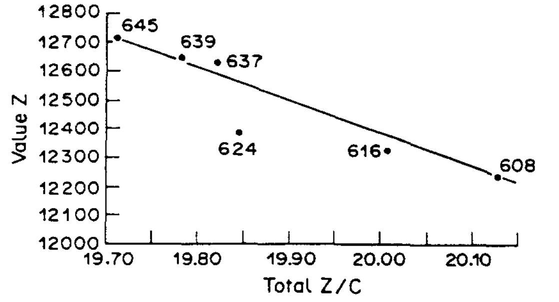

Fig. 12. Tradeoff between Z and Z/C on the dominant frontier.

the resultant overall Z/C, then the results are presented in Table 2 and Fig. 10, indicating that the total value Z increases with budget availability, while the overall ratio Z/C tends to decline (as already shown in Figs 4 and 8 respectively). Thus, if total Z and total Z/C are regarded as twin objectives, then an improvement in total Z can generally be achieved at the expense of a dilution in total Z/C, but (as Table 2 and Fig. 10 suggest) there are some notable exceptions. For example, an increase in the budget from 600 to 602 to 608 results in both measures (Z and total Z/C) improving, and a similar result is recorded when the budget is increased from 631 to 637.

The pursuit of twin objectives is reminiscent of the use of a portfolio diagram, where investments are projected against their expected financial returns and associated risks; the purpose of the portfolio is to maximise return and minimise risk, and the diagram helps to identify those projects that are superior to others with respect to both objectives. In our case, it is clear that a budget of 608 is preferable to 600 and 602, and similarly 637 (and 639) is preferable to 631 and 632, and this leads to the construction of the dominant frontier in Fig. 11, where every possible result under the line is dominated by a point on the line. This means that only the six budget points that determine the dominant frontier need to be considered. Any choice between the six points (for example, whether the point 608 is preferable to 616 or not) would depend on the tradeoff between the two objectives, i.e. the degree to which it is thought desirable to increase the total value Z at the expense of total Z/C. The general tradeoff trend is represented by the slope of the line in Fig. 12, indicating that sacrificing 0.1 in the total Z/C ratio is generally expected to improve the total Z value by about 115. However, only a few points lie on this general trend line, and budget points such as 616 and 614 results in poorer returns in terms of total value Z (they are below the line), whereas 637 is distinctly better in this respect (it lies above the line). All these considerations would tend to suggest that:

(a) There is a case to be made for increasing the initial budget of 600 to 608 and possibly to 637.

(b) Of all the budget points on the dominant frontier, the final choice should probably be between 608 and 637.

CONCLUSION

The budget problem can be formulated as a knapsack model, for which an optimal solution can easily be found. However, the optimal solution can be very sensitive to the value of the budget constraint, and it is quite possible for a very small change in the budget to invalidate this solution. This problem can become particularly serious when projects are selected and implemented within a given budget, without considering the possibility that the budget might subsequently be extended.

In practice, therefore, budget allocation often needs to be considered within a postulated budget range, to ensure that project selection is likely to be compatible with a range of budget values. Also, as decision makers are often concerned with 'value for money', which may be measured by the overall Z/C ratio, the problem of allocation within a budget range is associated with two objectives, which are rarely fully compatible, namely maximising the total value Z and the total ratio Z/C.

The methodology explored in this paper highlights the following:

(1) Within a defined budget range, the maximum total value Z rises, as a step-wise function, with budget availability (Fig. 5), while the budget utilisation assumes a spiky pattern (Fig. 4).

(2) The overall Z/C ratio tends to fall with budget availability, but there are notable local fluctuations, which indicate that certain budget points are preferable to others (Figs 8 and 9).

(3) The results of all the optimal solutions within a given range (Fig. 6) provide a useful *index of stability*, to indicate which projects are more likely to be included in the optimal selection compared with others.

(4) These considerations, coupled with the identification of dominant projects (i.e. those that are shown to be better than others with respect to both criteria) can greatly reduce the many optimal solutions (which correspond to the various points in the budget range) to a very small set, from which a final choice has to be made.

ADDRESS FOR CORRESPONDENCE: *Profesor S. Eilon, Imperial College of Science and Technology, Exhibition Road, London SW7 2BX, UK.*