AD-A080 453 TEXAS UNIV AT AUSTIN CENTER FOR CYBERNETIC STUDIES F/G 5/1

AUDITING AND ACCOUNTING FOR PROGRAM EFFICIENCY AND MANAGEMENT E--ETC(U)

AUG 79 A CHARNES, W W COOPER N00014-75-C-0616

UNCLASSIFIED CCS-343 NL

1 OF 1

AD A080453

END

DATE FILMED

3 - 80

DDC

LEVEL

CENTER FOR CYBERNETIC STUDIES

The University of Texas
Austin, Texas 78712

Research Report CCS-343

AUDITING AND ACCOUNTING FOR PROGRAM EFFICIENCY AND MANAGEMENT EFFICIENCY IN NOT-FOR-PROFIT ENTITIES.

by

A. Charnes*
W.W. Cooper**

53

August 1979

*The University of Texas at Austin
**Harvard University Graduate School of Business

This paper represents the draft of a paper to be presented at the August 24, 1979, plenary session of the American Accounting Association Honolulu Convention. It was partly supported by NSF Grant No. SOC76-15876, "Collaborative Research on the Analytical Capabilities of a Goals Accounting System." It is also part of the program of the Navy Personnel Research and Development Center and is partly supported by Projects NR047-071 and 021, ONR Contracts N00014-75-C-0616 and 0569 with the Center for Cybernetic Studies, The University of Texas at Austin, and ONR Contract N00014-75-C-0932 at Carnegie-Mellon University School of Urban and Public Affairs. Reproduction in whole or in part is permitted for any purpose of the U.S. Government.

CENTER FOR CYBERNETIC STUDIES

A. Charnes, Director
Business-Economics Building, 203E
The University of Texas at Austin
Austin, TX 78712
(512) 471-1821

406 199 LB

Auditing and Accounting for Program Efficiency and Management Efficiency in Not-for-Profit Entities

by

A. Charnes and W.W. Cooper

ABSTRACT

A measure of efficiency for not-for-profit entities--developed by the authors in association with Edward Rhodes--is explained and illustrated by data from Program Follow Through, a large scale social experiment in U.S. public school education. A division into Follow Through and Non-Follow Through participants facilitates a distinction between "program efficiency" and "managerial efficiency" which is also illustrated and examined for its use in evaluating such programs. Relations to comprehensive audits and other possible uses are explored.

COMPREHENSIVE AUDIT, AN INTRODUCTION

A plenary address in the same session with U.S. Comptroller General Elmer Staats provides an opportunity to relate this presentation to the new kinds of audits which his organization, the U.S. General Accounting Office, has pioneered. This type of audit, which we have elsewhere referred to as "comprehensive audit," is also being developed in other governments (both foreign and domestic)[1/] and in the internal audit practices of some of the large multinational corporations.[2/] Nevertheless the U.S. General Accounting Office, under Dr. Staats, must be accorded the position of leader in these important developments[3/] and it is their terminology and concepts which we shall adapt as a guide in the discussion that follows.

The term "audit" is often viewed, especially by accountants, as synonymous with the CPA function of attesting to financial statements and related representations by management. The types of audit we are considering, although broader in scope, do continue the CPA tradition of servicing "third party" needs for information as one of its major functions -- along with corollary responsibilities for objectivity and professional canons of care and validation. In operational terms this means that conditions such as auditor independence and attention to conditions of fairness (or balance) and evidentiary documentation must attend all aspects of the appraisals that are undertaken in these comprehensive audits.

The term "comprehensive" seems appropriate because this type of audit extends the process of objective appraisal to <u>all</u> aspects of management. By this we mean that the auditor assumes responsibility for the aspect of management that is designated for audit as well as the way the

---

[1] *Vide* [32] and [33] by Dr. Staats and the report of the Auditor General of Canada [17]. Other developments are also examined in [20] and [21]. Se also [22].

[2] See the discussion in [11]. See also [10].

[3] Cf. [12.1]-[12.3]. See also [18] and [33].

audit is conducted. Under comprehensive audit, the auditor does not stop short with the attest function. Instead, he also assumes responsibility for the report that is rendered, as well as the third party groups to be serviced by the audit process. Finally, instead of restricting audit examinations to financial transactions (and their representations in financial reports), such comprehensive audits extend to examination and appraisal of any of the following aspects of management and organization behavior:

Table 1

1. Propriety of
   a. Objectives pursued
   b. Methods used
2. Effectiveness in
   a. Stating objectives
   b. Attaining objectives
3. Efficiency of performance as measured by
   a. Benefits received
   b. Resources utilized.

Some or all of these aspects of "comprehensive audit," as we have already observed, are to be found in well-developed forms in certain types of governmental and internal corporate audit activities. They are also to be found in areas as widely dispersed as the audit of CPA firms by other CPAs[1] in what are called "peer reviews" and extend also to peer review processes that are now being used to audit and appraise the practices of physicians in various U.S. hospitals,[2] and include the efficiency audits that are now being ordered by public utility commissions for the utilities under their jurisdiction.[3]

---

[1] See [25]. See also pp. 34ff. in Appendix E to [34] for disccussion of the need for public disclosure of the results of such peer reviews in order to service 3rd party needs for information and evaluation.

[2] E.g., by so-called PSRO's Professional Standards Review Organizations). See Govindarajan [16] for further discussion in a study that explicitly relates these PSRO procedures to the audit practices we are describing.

[3] See San Miguel [29] and [30].

In search of an enhanced accountability, these extensions of the audit function are also being accompanied by alterations in corporate and government agency governance structures. Witness, for instance, the current attention to reforms such as the new Inspector General legislation[1/] and the attention to outside directors and audit committees in private corporations.[2/] These all represent attempts at evolution toward new forms of social and management controls which can (and should) be aided by the kinds of research that academic accountants (and others) can supply.

---

[1] See the discussion in the paper [32] by Dr. Staats.

[2/] A report of the SEC's views on the current state of these developments may be found on pp. 6ff. in Appendix D of [34].

2. DECISION-MAKING EFFICIENCY

It is one purpose of the present paper to try to effect such a contribution. In particular we want to introduce a new way to evaluate the efficiency of management in the not-for-profit entities that constitute such a large (and growing) part of our economy. We want to do this, moreover, in a manner that opens the possibility of distinguishing between programs, policies, organization designs, and related assignments of decision-making responsibilities, etc., in order to help ascertain how these may serve to enhance or impede decision-making efficiency.

Referring to Table 1, above, we might specifically say that we are directing our attention to the synthesis of a single number summary to help measure the efficiency of not-for-profit entities. This is to be done in a way that takes account of the fact that the activities of most such entities involve multiple outputs and multiple inputs which are often remote from any markets. We therefore want to avoid the need for *a priori* weights such as imputed market prices, etc., in our measure of efficiency. We seek instead to synthesize a method for objectively determining the weights we shall be using. Then we shall supply interpretations that will lend meaning to our results and also open avenues for uses that extend beyond the particular applications that we shall be exploring.

Ways in which our approach might be related to audits and other aspects of accounting will also be indicated at appropriate places in the sections that follow. To conclude the present section, however, we might first observe that the private enterprise sector already has a measure, *viz.*, net profit, which simultaneously accommodates both the effectiveness and efficiency components which are listed in Table 1.[1] The measure we shall be exploring is more limited in that it is restricted to the efficiency dimension only. For public programs and not-for-profit entities, this seems to be the best way to proceed. Note, for instance, that it is acceptable private enterprise practice for management to invest in new undertakings which are far removed from past and present activities. Such a diversion of funds to programs which are far removed from present and past activities would not, in general, be available to the management of a government agency, say, without first securing legislative and/or other authorizations - perhaps including even public referenda -- and such freedom to specify directions as well as the magnitude of the efforts to be undertaken are components that would need to be evaluated in terms of "effectiveness" as well as "efficiency."

[1]Note that "propriety," in the sense of Table 1, is not necessarily accommodated either by profit or the efficiency measures we shall be examining.

3. A MODEL FOR EFFICIENCY: HYPOTHETICAL EXAMPLE[1/]

The data in Table 2, below, will help to fix ideas in the developments pointed toward the efficiency measure we shall be examining. Here we are supposing a hypothetical situation in which three Decision-Making Units[2/] -- DMU_1, DMU_2, and DMU_3 -- each use two inputs in amounts x_1 and x_2 to produce a single output in amount y. Indeed, we may imagine the data in the table as referring to the amounts of the two inputs utilized per unit output so that we can then have the same one unit output reference for evaluating the efficiency of each DMU.

Table 2

An Illustration of DMU Efficiency

Input \ DMU No.	1	2	3
x_1	2	3	4
x_2	2	2	1

Assuming that all inputs as well as the one output have some "value," we may note that DMU_2 is not as efficient as DMU_1. Observe, for example, that DMU_1 has produced the same amount of output (one unit) with one unit less of the first input (i.e., we have $x_{11} = 2$ vs. $x_{12} = 3$) and no more of the second input.

Of course, the above characterization is only qualitative. That is, it produces only a classification of "efficient vs. non-efficient" whereas what we are seeking is a scalar measure of efficiency which can be accorded operational significance.

1/ This and subsequent parts of the present paper draw heavily on work done in collaboration with E. Rhodes as reported in [8]. See also [6] and [9].

2/ We use the term "Decision Making Units" by virtue of the fact that we are concerned with such units as "Schools," "Hospitals," Government Agencies" etc. and hence want to avoid the connotation of mure such terms as "firms," "corporate bodies," etc. See [9] and [26] for further discussions.

To obtain the latter result we now utilize all of the data in Table 2 in order to obtain the rating that is implied in the following model:

$$\max.\ h_o = \frac{1u}{3v_1 + 2v_2}$$

subject to:

$$1 \geq \frac{1u}{2v_1 + 2v_2}$$

$$(1) \qquad 1 \geq \frac{1u}{3v_1 + 2v_2}$$

$$1 \geq \frac{1u}{4v_1 + 1v_2}$$

$$u,\ v_1,\ v_2 > 0.$$

Observe that all three DMUs are represented in the constraints in a manner which guarantees that all will have efficiency ratio values that lie between 0 and 1. Hence DMU_2 , which is singled out for evaluation in the functional to maximized in (1), must also have this same characteristic.

In conformance with the measures that are customary in both the natural and social sciences, we have synthesized our model so that the efficiency ratings are in the form of ratios that are limited to a maximum value of 100%. Furthermore, the resultant rating can be shown to be scale independent,[1/] which is to say that each input and each output can be measured in any units that are convenient, without altering the results, provided these same units are utilized for every DMU.

[1] See [6] and [26] for further discussion of these and other invariance properties of these efficiency measures.

It should be observed that our efficiency rating is a relative one. That is, the rating of DMU_2 in the above case is obtained by reference to the attainments only of the other DMUs - viz., DMU_1 and DMU_2 - with which it is being compared. This is the way we shall proceed although, in passing, we might also remark that comparisons obtained from engineering or other estimates of efficiency may also be included when they are available. Indeed, it can be shown that all of the customary engineering (and natural science) definitions of efficiency[1] can be accommodated (and extended) by reference to the general form of the model that is represented in (1).

Although the model represented in (1) is non-linear, it can also be shown to have a wealth of properties which include a "fractional programming" equivalent that can, in turn, be replaced by an ordinary linear programming problem.[2] Thus, by using this route we may apply the simplex method of ordinary linear programming to the data of Table 2 in order to obtain

(2.1) $$\max.\ h_o = h_o^* = 6/7$$

along with the optimal weights

(2.2) $$u^* = 1,\ v_1^* - 1/6 \text{ and } v_2^* = 1/3.$$

Substituting these values in (1) we obtain

---

[1] See [9] and [26] for further discussion of the underlying definitions and the way these relate to ideas like Pareto efficiency in economics. See [6] for further discussion in relation to engineering definitions of efficiency and a demonstration that the measure we are using does not depend on the measurement units used.

[2] See [9] and [26]. This means that we also have the duality theory of fractional programming available for use in a variety of contexts, as well as the duality theory of ordinary linear programming.

$$\frac{1u^*}{2v_1^* + 2v_2^*} = 1$$

(3) $$\frac{1u^*}{3v_1^* + 2v_2^*} = 6/7$$

$$\frac{1u^*}{4v_1^* + 1v_2^*} = 1 ,$$

and so, as a byproduct of our calculations, we may observe that both DMU_1 and DMU_3 are characterized as efficient. I.e., they have values of unity for their h_o^*. The $h_o^* = 6/7$ for DMU_2 (which will also appear as the optimal value of the functional) means that it is inefficient. This value of $h_o^* = 6/7$ is also accorded the operational significance of meaning that under efficient production only 6/7 of the amount of each input utilized by DMU_2 should have been required. Alternatively, the reciprocal value could have been employed to mean that 7/6 units of output could have been secured in place of the one unit actually realized from the input amounts utlized by DMU_2.

Reference to Figure 1, below, will provide some easy extensions that will also further help us to interpret and evaluate the results we have now achieved. Here DMU_1, DMU_2 and DMU_3 are accorded interpretations as points in a two-dimensional space of the inputs associated with x_1 and x_2, respectively. The solid line connecting P_1 (= DMU_1) and P_3 (= DMU_3) represents the so-called "unit isoquant" along which efficient production should occur when only one unit of output is produced.

Evidently P_2 (= DMU_2) does not lie on this isoquant. The point P_2' obtained from

(4.1) $$\frac{6}{7} P_2 = \frac{6}{7}\begin{pmatrix}3\\2\end{pmatrix} = \begin{pmatrix}18/7\\12/7\end{pmatrix} = P_2'$$

does, however, lie on this line. This is interpreted to mean that under efficient production DMU_2 should have been able to cut the amounts of each input to 6/7 of its observed value in achieving one unit of output. Alternatively, we can use the v^* values of (2.2) in order to substitute in the denominator of the second expression -- i.e., the expression for DMU_2 -- in (3) to obtain

$$(4.2) \qquad 3v_1^* + 2v_2^* = 7/6 \; .$$

In this form we are saying that DMU_2 should have obtained 7/6 units of output (or output rate) from the amounts of the inputs utilized.

To appreciate the sense of what is being said, we may observe that these v_1^* and v_2^* values represent the "marginal productivities" associated with the inputs to which they are assigned in this single input case.[1] These are efficient marginal productivities, however, obtained by reference to the data of P_1 (= DMU_1) and P_3 (= DMU_3).[2] In short these v_1^* and v_2^* values are the <u>efficient</u> marginal productivities.

The production surface from which the isoquant in Figure 1 is derived is piecewise linear. It follows that for each such segment we have the result that marginal and average productivities are equal. Thus the v_1^* and v_2^* values obtained from (1) provide an estimate of the total product that <u>should</u> have been obtained if the <u>efficient</u> productivities were applied to the input data of DMU_2 as in (4.2).

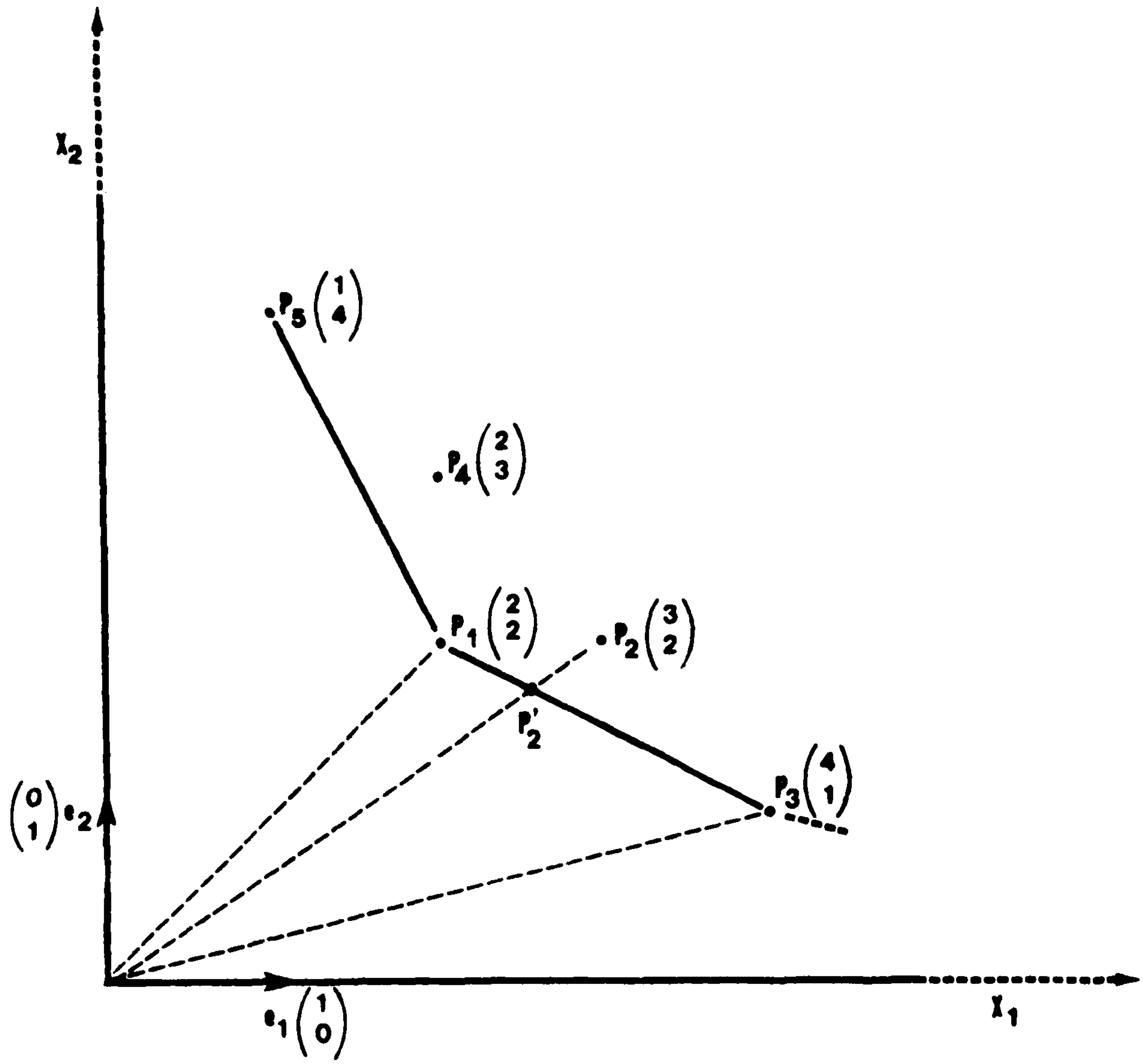
---

[1] Proofs of this proposition are provided in [9] and [26].

[2] Observe, in particular, that these values are <u>not</u> the marginal productivities observed for DMU_2. For further discussion and accompanying proofs, see [[illegible]] and [[illegible]].

Figure 1.

Efficiency Points and Isoquant



Turning again to Figure 1 we may imagine that P_5 and P_4 represent data for two additional DMU's. That is, continuing with the same inputs and same outputs as the DMUs in Table 2 we imagine that the data set is expanded to include DMU_4 and DMU_5 while retaining the data for DMU_1, DMU_2 and DMU_3.

Singling out DMU_4 with values for its input coordinates as represented in Figure 1, we use <u>all</u> of the data shown in this Figure to rate the unit's efficiency via

$$\max.\ h_o = \frac{1u}{2v_1 + 3v_2}$$

subject to:

$$
(5)\qquad
\begin{aligned}
1 &\geq \frac{1u}{2v_1 + 2v_2} \\
1 &\geq \frac{1u}{3v_1 + 2v_2} \\
1 &\geq \frac{1u}{4v_1 + 1v_2} \\
1 &\geq \frac{1u}{2v_1 + 3v_2} \\
1 &\geq \frac{1u}{1v_1 + 4v_2} \\
& u,\ v_1,\ v_2 > 0\ .
\end{aligned}
$$

This, we may observe, simply extends the formulation in (1) by (a) adjoining the two new constraints associated with the data for DMU_4 and DMU_5 and (b) replacing the previous functional to be maximized for rating DMU_2 with the new one for rating DMU_4.

Proceeding as before to obtain a solution to (5) we have:

(6.1) $$h_o^* = 6/7$$

along with the optimal weights

(6.2) $$u^* = 1,\ v_1^* = 1/3,\ v_2^* = 1/6.$$

Substituting in (5) we see that all constraints are satisfied - viz.,

(7)
$$\frac{1u^*}{2v_1^* + 2v_2^*} = 1$$
$$\frac{1u^*}{3v_1^* + 2v_2^*} = 3/4$$
$$\frac{1u^*}{4v_1^* + 1v_2^*} = 2/3$$
$$\frac{1u^*}{2v_1^* + 3v_2^*} = 6/7$$
$$\frac{1u^*}{1v_1^* + 4v_2^*} = 1\ .$$

We now make the important observation that the efficient reference set consists of P_1 (= DMU_1) and P_5 (= DMU_5) which are the only DMUs achieving a value of unity as a result of the optimization in (5). Comparing the v^* values in (6.2) and (2.2) we see that these have altered. Thus, although DMU_4 has a value $h_o^* = 6/7$ the same as DMU_2 in (2.1), we may also observe that DMU_2 has a value of 3/4 in the solution provided by (7). The former value, e.g., $h_o^* = 6/7$, is the correct rating, however, for DMU and is obtained from its reference set consisting of P_1 and P_3. The points P_1 and P_5 in (7) are the correct reference set for DMU_4 but not DMU_2. Thus, we must be sure to obtain the efficient marginal productivities from the currect reference set.[1] Simply averaging or regressing across the entire efficient set will not do!

[1] The way in which this extends the Marshallian "representative firm" concept is discussed in [9].

We may now summarize the results we have achieved by formally writing

(8.1) $\{(x_1,x_2) : \frac{1}{6}x_1 + \frac{1}{3}x_2 = 1;\ 2 \le x_1 \le 4;\ 1 \le x_2 \le 2\}$,

for the efficient isoquant segment that extends between P_1 and P_2. This, in turn, differs from

(8.2) $\{(x_1,x_2) : \frac{1}{3}x_1 + \frac{1}{6}x_2 = 1;\ 1 \le x_1 \le 2;\ 2 \le x_2 \le 4\}$,

for the segment that extends between P_1 and P_5 in the "piecewise linear" fashion that is exhibited in Figure 1.

Observe that we have now also provided a new way for effecting estimates of economic relations from empirical data. These differ from the ordinary statistical approaches such as "regression" and "simultaneous equation" estimation in various ways. The latter, for instance, focus on averages and other measures that are in the interior of the observations from which estimates are to be effected. Our measures are directed to extremals, or extremal relations, that lie rather on the boundaries rather than in the interior of these same observations.

It must be expected that uses of these new methods of estimation may also differ in important ways from usages that are associated with the more customary techniques of statistical regression. Note, for instance, that the slope conditions in (8.1) and (8.2) assume that the efficient boundary is, or can be made, applicable to DMUs which lie within the regions defined by the inequalities in each set. This is to say that two possible sources of economizing are present. One source consists of economizing on inputs or expanding the outputs by bringing each DMU onto the efficient surface in accordance with its h_o^* value.[1/]

[1] Because it permits an expansion of output without requiring an increase in any input, this is technically identified as "waste"-- in order to distinguish it from other types of economizing that may also be possible.

The other source consists of effecting substitutions between various inputs in accordance with prices or other conditions for the applicable ranges, as exhibited in each of (8.1) and (8.2) after the associated DMU has been brought into efficient operations.

We have elsewhere distinguished this new method for estimating extremal relations by referring to it as a "control prediction."[1] That is, in contrast to the "pure predictions" which we associate with ordinary social science usages of statistical regressions, etc., such control predictions assume that mechanisms exist for bringing about the wanted states of efficient production.

Here we would like to associate these kinds of "control prediction" results with procedures such as the "analytic reviews," etc., which are often used to guide the planning of audit strategies in both financial and comprehensive-type audits. Thus we might, for instance, assemble data on a set of public schools -- along lines such as we shall soon be discussing -- in order to direct audit efforts to more detailed attention at specified efficient or inefficient sites.

Such "control predictions" are, of course, not entirely new to accountants. They may be found, for instance, in areas like engineered standard costs. Indeed the analogy with "efficiency standards" may be considered to be a close one when the latter is associated with predicted outcomes that should occur when prescribed (efficient) procedures are followed. Just as in the case of our measure, too, an efficiency variance does not include price variations explicitly although, of course, relevant substitution of materials and procedural alterations may be effected when prices warrant this.

---

[1] See [6].

Further, the appearance of a black or red efficiency variance serves as an attention directing device for further inquiry-- perhaps by an auditor -- rather than by automatically providing the specific correctives that may be required.

Finally, we may observe that, as is also true for standard costs, our efficiency measure is intended to supply new options that were previously not available for policy guidance and it is also intended to open still further possibilities such as the distinction between "program efficiency" and "managerial efficiency" that we shall now begin to explore.

4. DATA ENVELOPMENT ANALYSIS

Figure 2, below, will help us to introduce our next topic. We may think of this as a portrayal in which amounts x of a single input have resulted in amounts y of a single output (all else held constant) for a collection of DMUs with the resulting observations portrayed as in Figure 2. There is a further distinction in that these DMUs operate under two different technologies, or programs, so that the efficient frontiers for one set are at B while the efficient frontiers of the other set are at A. Evidently even efficient DMUs in B cannot achieve the levels of those in A and even some of the less efficient DMUs operating under A exceed what is attainable under B.

Customary approaches as in a use of statistical regressions might have concealed some of the alternatives for choice that are present because the observations generally contain both different technological possibilities and different degrees of efficiency in their utilization. This is the situation we would now like to confront by extending our previous analysis in order to bring it to bear in ways that might help us distinguish between what we shall refer to as "managerial efficiency" and "program efficiency."[1/] The latter may be thought of in terms of the kinds of frontiers depicted in Figure 2 while the former may contain inefficiencies resulting from managerial decisions that fail to utilize these opportunities to the full.

"Data Envelopment Analysis" is the name that we shall use for the approach we will be suggesting because (a) it first tries to locate the boundaries that envelop the observations as in A and B of Figure 2 and then (b)

[1]There may be valid reasons for failure to utilize efficiency possibilities to the full, as R. Lehto has observed to us, but even then the approach we are suggesting can provide measures of the "cost" of the actual vs. potential efficiencies that are then foregone.

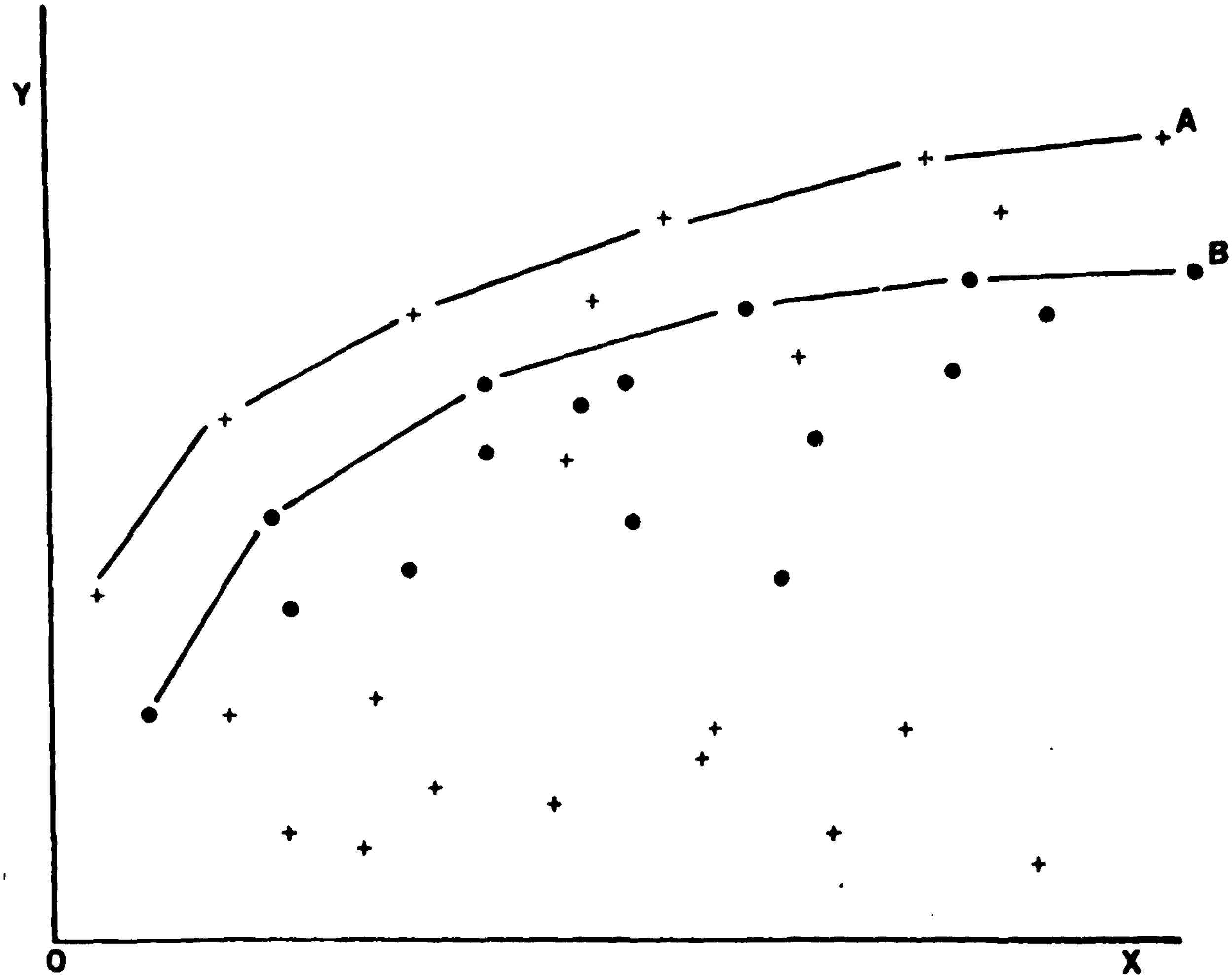
it brings the observations all up to the envelope that is pertinent in each case. The degree of "management inefficiency," if any, is first measured relative to the pertinent A envelope or B envelope[1] in the manner indicated in the preceding section for each observation. The remaining efficiency differences are then imputed to the respective programs (or technologies) so that in the situation of Figure 2, say, the program associated with A would be characterized as more efficient than the program associated with B.

In the preceding section we provided a measure of efficiency that had operational significance when the managers were all operating under a common program. It is evidently also desirable to try to do this in our assessment of program efficiencies as well. That is, we would like to have a means of taking account of the efficiency gains (or losses) that can be secured by moving DMUs from one program to the other.

---

[1] We refer to these more generally as "α-envelopes" in order to allow for comparisons between α=1, 2, ..., n different technologies or programs. See [8].

Figure 2



Legend:

Pluses(+) = Observations for DMU's in A

Dots (●) = Observations for DMU's in B

To clarify what we are about, we may again refer to our earlier contrast between "control predictions" and "pure predictions." A variety of possibilities then confront us. For instance, we might fit two separate statistical regressions of the ordinary least squares variety to the observations shown in Figure 2. This could lead us to conclude that the B program was better than A after which we might employ the resulting B regression to estimate how the DMUs previously operating under A would perform if the latter were eliminated in favor of B.

However, if new "controls" could be instituted to alter past patterns of behavior then a different prediction (of the "control" variety) might better be employed. In particular, if these controls could ensure that all DMUs would operate at the program boundaries then A would be preferred to B.[1]

DEA, Data Envelopment Analysis, the approach we shall now begin to examine, is intended for use in the latter situations since, without such controls, the resulting "efficient behavior" predictions will generally be invalid. We may, in fact, think of DEA as providing guidance for program audits which will help to locate sources of inefficiency along lines such as we have indicated in the preceding sections of this paper - e.g., directing such audit examinations to particular DMUs brought into view by a DEA analysis. A separate decision may then be made concerning whether the projected efficiencies can be realized and perhaps thereby justify a choice of A rather than B.

---

[1]Still other possibilities would allow for situations in which only some of the DMUs might be brought up to their boundaries while others are only partially improved, etc. The point is, in any case, that our DEA approach opens new alternatives for consideration besides those available only from the "pure predictions" approach.

5. THE PROGRAM FOLLOW-THROUGH EXPERIMENT

In order to illustrate what is involved we turn to data from "Program Follow-Through," a large-scale experiment in U.S. public school education designed to help disadvantaged children. This program was intended to further the objectives of a pre-school program known as Project Head Start, by carrying on specially designed programs for disadvantaged children from kindergarten through grade 3. This was done at the 70 selected "school sites" in various parts of the U.S. which are listed in Table A-1. (See first two columns of Table A-1 for the numbering that will hereafter be identified with these sites.)

An assessment of Program Follow-Through was sought in what was intended as a designed experiment[1] by selecting a matched pair of Program Follow-Through and Non-Follow-Through sites from which observations were to be secured. These are identified as PFT and NFT, respectively, in Columns 1 and 2 of Table A-1. Although we need to observe that the experiment was not carried out successfully in all respects, it nevertheless produced a wealth of data from which we effect a selection as follows.

We restrict ourselves only to educational aspects of the program while ignoring other parts of its activities such as supplying medical and dental services, nutritional programs, etc. We also omit any explicit consideration of costs on the supposition that a cost/benefit analysis only becomes pertinent if statistically significant effects are obtained from program elements such as we are considering.

---

[1] Or at least as close to such an experiment as one is likely to achieve in public school situations.

Table A-1

Site Level Distribution of DEA Study Sample

PFT Site #	NFT Site #	Model and Site Name	Region[1]	City Size[2]	PFT[3] Student Pop.	NFT[4] Student Pop.
		Responsive Education Model				
1	50	Berkeley,CA	W	Medium City	99	71
2	51	Buffalo,NY	NE	Large City	77	27
3	52	Duluth,MN	NC	Medium City	77	79
4	53	Fresno,CA	W	Medium City	48	54
5	54	Lebanon,NH	NE	Rural Area	14	97
6	55	Salt Lake,UT	W	Medium City	36	51
7	56	Tacoma,WA	W	Medium City	51	42
		TEEM Model				
8		Baltimore,MD	S	Large City	99	-
9		Lakewood,NJ	NE	Small City	80	-
10	57	Lincoln,NB	NC	Medium City	96	55
11	58	Wichita,KS	NC	Large City	84	36
		Bank Street Model				
12	59	New York,NY	NE	Large City	72	245*
13	60	Philadelphia,PA	NE	Large City	80	37
14		Brattleboro,VT	NE	Small City	20	-
15	61	Fall River,MA	NE	Medium City	39	18
16		Wilmington,DE	S	Medium City	109	-
		DIM Model				
17		New York,NY	NE	Large City	31	-
18	62	E.St. Louis,IL	NC	Large City	56	21
19		Grand Rapids,MI	NC	Medium City	103	-
20	63	Racine,WI	NC	Medium City	62	27
21	64	Flint,MI	NC	Medium City	77	66
		BA Model				
22		New York,NY	NE	Large City	43	-
23	65	Philadelphia,PA	NE	Large City	208	27
24		Portageville,MO	NC	Rural Area	47	-
25		Kansas City,MO	NC	Large City	61	-
26		Louisville,KY	S	Large City	90	-
27		Meridian,IL	NC	Rural Area	68	-
		Cognitive Curriculum Model				
28		New York,NY	NE	Large City	32	-
29		Chicago,IL	NC	Large City	18	-
30		Okaloosa Co.,FL	S	Small City	48	-
		Parent Education Model				
31		Philadelphia,PA	NE	Large City	46	-
32	66	Jacksonville,FL	S	Large City	15	53
33	67	Richmond,VA	S	Large City	111	69
34	68	Houston,TX	S	Large City	95	78
		EDC Model				
35		Philadelphia,PA	NE	Large City	152	-
36		Paterson,NJ	NE	Medium City	42	-
		Self-Sponsored Model				
37		Detroit,MI	NC	Large City	43	-
38	69	New York,NY	NE	Large City	20	13
39		Philadelphia,PA	NE	Large City	86	-
40		Portland,OR	W	Large City	45	-
41		San Diego,CA	W	Large City	71	-
		ILM Model				
42		New York,NY	NE	Large City	53	-
		SEDL Model				
43	70	Philadelphia,PA	NE	Large City	86	36
44		Tulare,CA	W	Small City	173	-
		Home-School Partnership Model				
45		New York,NY	NE	Large City	26	-
		California Process Model				
46		Los Angeles,CA	W	Large City	98	-
47		Ravenswood,CA	W	Small City	74	-
48		Lamont,CA	W	Rural Area	27	-
49		San Jose,CA	W	Large City	42	-
				Total Student Pop.	3210	1202

* Pooled Citywide NFT Population Figure for New York City

(Footnotes continued on next page)

Table A-1
(continued)

[1]NE = North Eastern United States
S = Southern United States
NC = North Central United States
W = Western United States

[2]Large City	=	200,000 or more
Medium City	=	50,000 to 199,999
Small City	=	10,000 to 49,999
Rural Area	=	Less than 10,000

[3]All Data Envelopment Analysis study information refers to the Cohort II-K student population. II-K indicates that this group of students began their Program Follow Through experience in kindergarten. (This was also the only one of three Cohorts which had completed all of the grades from kindergarten through third grade at the time of our study for which site level information was available.) However, due to incomplete statistics along some DEA variable dimensions, some of the Cohort II-K PFT sites were not included in the DEA study. Specifically, Bank Street Model: Rochester,NJ site; EDC Model: Chicago,IL site; and SEDL Model: St. Martin Parish,LA site were excluded from the DEA study student population. The actual Cohort II-K PFT population was 3,367 of which, as noted above, a set of 3,210 students were used in the DEA study. This exclusion of sites also extended to the NFT groups which were similarly reduced to 1,202 students.

[4]Two sets of NFT students groups were created in the original Program Follow Through study. One group was a local student set, usually in the same school system as the subject PFT site. The second group, and the one selected for the DEA study, was a "best matched" group, which may or may not have been located in the same school system or even the same geographical region. The NFT group which most nearly matched the PFT students of a given site along a number of demographic and initial performance dimensions was considered the "best match" for the latter. For several PFT sites the same "best matched" NFT group was used. The much smaller NFT student population total of 1,202 as compared to the PFT student total of 3,210 resulted. See also preceding footnote.

We shall focus on only one of several cohorts from the subjects comprehended in the study. In addition, we shall utilize only the terminal (grade 3) results for this cohort to avoid the additional complications needed to deal with dynamic or transient behavior en route to this terminus.[1] From a set of 11 output measures we select only the following 3 as sufficiently indicative for our purpose:

y_1: Total Reading Score as measured by the Metropolitan Achievement Test.

(9) y_2: Total Mathematics Score as measured by the Metropolitan Achievement Test.

y_3: Coopersmith Self-Esteem Inventory, intended as a measure of self-esteem.

This y_3 measure, we may note, is directed to non-cognitive growth (or affective behavior) in a dimension that is deemed pertinent to the objectives of this program. Together with y_1 and y_2, this y_3 variable provides a good indication of what is involved in assessing such programs. Note, in particular, that no easily available scheme for weighting the relative importance of these outputs is at hand. Nevertheless, some "overall" measure of program efficiency is wanted in order to enable us to evaluate PFT vs. its NFT alternative and this is to be achieved from data such as are exhibited in Tables A-2 and A-3.

---

[1] More detailed developments of the approach we are examining may be found in [8] and [26].

Table A-2

Unadjusted PFT Output Observations

Site #	Total Reading Scores, PHS* Y_1	Total Math Scores, PHS* Y_2	Total Coopersmith Scores, PHS* Y_3
1	54.53	58.98	38.16
2	24.69	33.89	26.02
3	36.41	40.62	28.51
4	14.94	17.58	16.19
5	7.81	6.94	5.37
6	12.59	16.85	12.84
7	17.06	16.99	17.82
8	20.29	30.64	33.16
9	26.13	29.80	26.29
10	46.42	51.59	35.20
11	39.80	37.73	30.29
12	37.84	47.85	25.35
13	26.48	31.36	26.54
14	10.31	10.86	7.47
15	14.39	18.30	14.33
16	32.94	36.03	38.19
17	17.25	20.80	12.07
18	27.55	38.19	20.44
19	41.12	43.80	36.54
20	29.43	42.63	23.34
21	37.46	51.02	27.44
22	19.40	25.18	16.52
23	39.88	47.72	38.97
24	25.72	30.81	16.54
25	24.88	25.27	22.43
26	31.62	40.78	31.16
27	31.31	38.32	25.03
28	21.00	21.30	18.30
29	6.51	7.02	6.16
30	11.64	15.26	15.68
31	12.58	15.90	14.42
32	4.59	6.16	4.99
33	43.76	46.64	39.10
34	32.38	38.55	31.05
35	34.64	45.46	39.22
36	11.52	15.14	13.91
37	15.96	19.21	15.30
38	9.91	12.30	7.22
39	30.44	33.53	29.80
40	22.63	25.24	17.15
41	24.41	27.16	25.30
42	23.11	22.67	17.56
43	21.82	31.45	27.54
44	63.92	79.67	63.11
45	9.47	11.92	8.85
46	33.94	39.18	34.61
47	29.42	35.10	28.42
48	7.70	11.02	9.02
49	12.17	16.03	15.82

* PHS = Per Hundred Students

Table A-3

Unadjusted PFT Input Observations

Site #	Education Level of Mother, PHS* X_1	Occupation Index, PHS* X_2	Parental Visit Index, PHS* X_3	Counseling Index, PHS X_4	Number of Teachers X_5
1	86.13	16.24	48.21	49.69	9
2	29.26	10.24	41.96	40.65	5
3	43.12	11.31	38.19	35.03	9
4	24.96	6.14	24.81	25.15	7
5	11.62	2.21	6.85	6.37	4
6	11.88	4.97	18.73	18.04	4
7	32.64	6.88	28.10	25.45	7
8	20.79	12.97	54.85	52.07	8
9	34.40	11.04	38.16	42.40	8
10	61.74	14.50	49.09	42.92	9
11	52.92	11.67	39.48	39.64	5
12	36.00	10.15	37.80	39.52	5
13	39.20	10.80	41.04	41.12	7
14	14.6	2.88	9.64	11.14	3
15	4.29	5.42	21.45	17.27	5
16	27.25	14.17	56.46	55.26	9
17.	22.63	4.43	15.40	15.00	2
18	28.00	7.61	28.73	27.04	9
19	53.56	13.70	53.04	49.85	7
20	25.42	9.05	29.69	31.74	4
21	31.57	10.08	39.34	40.57	6
22	16.34	5.84	20.89	22.10	4
23	44.28	14.14	56.70	52.27	11
24	19.74	6.43	24.20	25.66	3
25	24.40	8.05	33.42	31.29	7
26	41.40	11.70	44.01	46.35	7
27	27.20	9.38	37.80	31.55	4
28	23.92	7.12	25.58	29.01	3
29	10.62	2.55	10.10	9.09	4
30	12.48	6.14	23.13	22.46	6
31	19.32	5.89	24.01	24.74	6
32	6.30	1.93	7.11	7.68	4
33	46.62	14.65	65.71	57.49	10
34	38.95	12.82	47.02	48.92	9
35	61.60	15.56	53.98	50.29	6
36	31.08	6.26	22.18	21.96	4
37	19.35	6.68	22.61	23.31	4
38	11.20	3.08	9.90	10.06	2
39	34.40	11.61	41.79	41.79	5
40	35.55	6.48	21.69	21.69	6
41	30.53	9.30	35.50	35.14	8
42	25.44	7.10	26.81	26.23	3
43	26.66	11.43	41.36	44.63	6
44	39.79	22.49	84.77	76.12	11
45	8.32	3.64	12.92	13.13	2
46	59.78	13.52	48.80	49.69	15
47	39.22	10.06	37.00	38.33	4
48	3.28	3.18	13.12	12.71	5
49	7.14	5.29	23.10	19.06	8

*PHS = Per Hundred Students

The latter table, i.e., A-3, contains the input data for those same PFT sites in terms of the following variables:

(10)

x_1: Education level of mother as measured in terms of percentage of high school graduates among female parents.

x_2: Highest occupation of a family member according to a pre-arranged rating scale.

x_3: Parental visit index representing the number of visits to the school site.

x_4: Parent counseling index calculated from data on time spent with child on school-related topics such as reading together, etc.

x_5: Number of teachers at a given site.

These variables, like the output variables, are all measured in terms of scores per hundred students with the exception of x_5, the number of teachers, which is measured without further adjustment as given at each site.[1]

---

[1]The value of h_o^* as a measure of relative efficiency is not affected by such scale choices provided different scales are not used at different sites for any particular input or output. See [6] and [26].

The above variables represent a selection from 25 input variables which together with the 11 output variables define the effectiveness sought for the educational components of Program Follow-Through. Most of the inputs are given so, in fact, the measures of eefectiveness are largely in terms of the output scores that should be attainable with the input mixes available at each of the sites. This, in turn, is dependent on the efficiency with which these resources are utilized -- which we shall now measure along the lines of the more restricted data sets that we have already indicated.

Turning to the data in Tables A-2 and A-3 we might proceed to rate the efficiency of site 1, say, relative to the data for all 49 PFT sites by extending our previous analysis. This would yield the following expression:

$$\max.\ h_o = \frac{54.53u_1 + 58.98u_2 + 38.16u_3}{86.13v_1 + 16.24v_2 + 48.21v_3 + 49.69v_4 + 9v_5}$$

subject to:

$$1 \geq \frac{54.53u_1 + 58.98u_2 + 38.16u_3}{86.13v_1 + 16.24v_2 + 48.21v_3 + 49.69v_4 + 9v_5}$$

(11.1) $$1 \geq \frac{24.69u_1 + 33.89u_2 + 26.02u_3}{29.26v_1 + 10.24v_2 + 41.96v_3 + 40.65v_4 + 5v_5}$$

. . . . [1]

$$1 \geq \frac{12.17u_1 + 16.03u_2 + 15.82u_3}{7.14v_1 + 5.29v_2 + 23.10v_3 + 19.06v_4 + 8v_5}$$

$$u_1,\ u_2,\ u_3,\ v_1,\ v_2,\ v_3,\ v_4,\ v_5 > 0$$

Now we may observe that we are confronted with a situation involving multiple outputs and hence the simple concept of a production function with its related efficient marginal productivities is no longer

[1] The ellipsis "....." means that the 46 omitted expressions are to be formulated in the manner indicated by the expressions that are written explicitly.

directly available. We may instead think of the numerator as defining a new output and the denominator as defining a new input which are both scalar (single valued) numbers, that we shall refer to as the "virtual output" and the virtual input," respectively. In any case the value of h_o is limited to lie between 0 and 1, as before.

In a similar way we refer to Tables A-4 and A-5 to achieve efficiency ratings for the NFT sites. Thus, for instance, for site number 70 we would utilize the following to obtain its h_o^* efficiency rating:[1]

$$\max.\ h_o = \frac{13.69u_1 + 14.19u_2 + 12.99u_3}{10.44v_1 + 4.82v_2 + 17.13v_3 + 18.21v_4 + 9v_5}$$

subject to:

$$1 \geq \frac{39.07u_1 + 42.71u_2 + 27.67u_3}{68.16v_1 + 12.28v_2 + 33.58v_3 + 34.64v_4 + 15v_3}$$

(11.2) $$1 \geq \frac{9.96u_1 + 14.34u_2 + 9.33u_3}{11.88v_1 + 3.59v_2 + 13.41v_3 + 13.82v_4 + 8v_5}$$

. . . .

$$1 \geq \frac{13.69u_1 + 14.19u_2 + 12.99u_3}{10.44v_1 + 4.82v_2 + 17.13v_3 + 18.21v_4 + 9v_5}$$

$$u_1, u_2, u_3, v_1, v_2, v_3, v_4, v_5 > 0$$

[1] Note that the constraints remain the same in all cases and only the expression in the function being maximized changes. This feature can be utilized moreover, in the computing routines by allowing for an alteration between the "simplex" and "dual methods" as described in [26].

Table A- 4

Unadjusted NFT Output Observations

Site #	Total Reading Scores, PHS* Y_1	Total Math Scores, PHS* Y_2	Total Coopersmith Scores, PHS* Y_3
50	39.07	42.71	27.67
51	9.96	14.34	9.33
52	45.37	51.38	31.61
53	18.23	22.05	17.56
54	59.63	64.41	35.89
55	24.20	28.21	18.74
56	13.53	17.09	15.61
57	28.39	27.65	20.79
58	21.67	26.22	13.66
59	120.17	144.67	88.59
60	15.15	18.04	13.58
61	6.92	7.10	6.35
62	9.35	9.85	7.70
63	13.03	13.40	10.29
64	18.63	24.48	23.13
65	12.28	13.01	9.89
66	16.81	19.72	18.70
67	26.36	28.22	24.46
68	22.85	26.21	28.14
69	8.17	8.70	5.12
70	13.69	14.19	12.99

*PHS = Per Hundred Students

Table A-5

Unadjusted NFT Input Observations

Site #	Education Level of Mother, PHS* X_1	Occupation Index, PHS* X_2	Parental Visit Index, PHS* X_3	Counseling Index, PHS* X_4	Number of Teachers X_5
50	68.16	12.28	33.58	34.64	15
51	11.88	3.59	13.41	13.82	8
52	55.30	11.53	36.73	35.78	6
53	16.20	7.02	26.94	26.30	9
54	82.45	15.52	45.00	44.23	13
55	15.81	6.93	23.91	23.61	7
56	4.65	5.50	20.91	23.39	5
57	41.25	8.41	26.23	25.24	10
58	10.44	5.22	17.10	18.93	3
59	139.65	35.03	119.56	130.83	22
60	16.28	4.81	18.20	18.98	5
61	12.06	2.59	8.74	8.17	5
62	4.20	2.64	9.89	11.25	2
63	19.44	3.83	12.87	13.23	5
64	28.38	8.91	30.95	33.33	8
65	13.50	3.61	15.60	12.39	4
66	23.32	7.10	24.96	28.56	22
67	27.60	9.38	32.29	34.01	20
68	11.70	10.53	37.67	43.60	8
69	4.68	1.85	6.22	5.46	5
70	10.44	4.82	17.13	18.21	9

*PHS = Per Hundred Students

Repeated applications of the above procedures yields the results listed in Table A-6. Note that the efficiency ratings are with respect to each of the programs, PFT and NFT, respectively. Thus the h_o^* = 1 for site 1 is obtained relative to PFT sites only, while the h_o^* = 0.94 for site 70 is obtained relative to NFT sites only.

Table A-6

PFT and NFT Program Specific α-Envelope Efficiency Values

PFT Site #	h_0^{*1} Efficiency Value	NFT Site #	h_0^{*2} Efficiency Value
1*	1.00	50	0.95
2	0.90	51	0.92
3	0.98	52*	1.00
4	0.90	53	0.87
5*	1.00	54*	1.00
6	0.90	55*	1.00
7	0.89	56*	1.00
8	0.91	57	0.92
9	0.87	58*	1.00
10*	1.00	59	0.92
11	0.98	60	0.98
12	0.97	61	0.88
13	0.86	62*	1.00
14	0.98	63	0.96
15*	1.00	64	0.91
16	0.95	65	0.97
17*	1.00	66	0.92
18*	1.00	68*	1.00
19	0.95	69*	1.00
20*	1.00	70	0.94
21*	1.00		
22*	1.00		
23	0.96		
24*	1.00		
25	0.97		
26	0.93		
27*	1.00		
28	0.94		
29	0.84		
30	0.90		
31	0.83		
32	0.90		
33	0.94		
34	0.85		
35*	1.00		
36	0.60		
37	0.94		
38	0.94		
39	0.91		
40*	1.00		
41	0.94		
42	0.94		
43	0.87		
44*	1.00		
45	0.89		
46	0.90		
47*	1.00		
48*	1.00		
49*	1.00		

* Denotes a site with an efficiency value of "1"

6.0 PROGRAM AND MANAGEMENT EFFICIENCY UNDER DEA

As a first step in our analysis we might try to ascertain whether managerial efficiencies differ under PFT and NFT. With DEA this is done, we may recall, by separately obtaining the efficiency ratings for the DMUs under each of PFT and NFT by reference to their respective envelopes, and this is the way the data in Table A-6 were obtained.

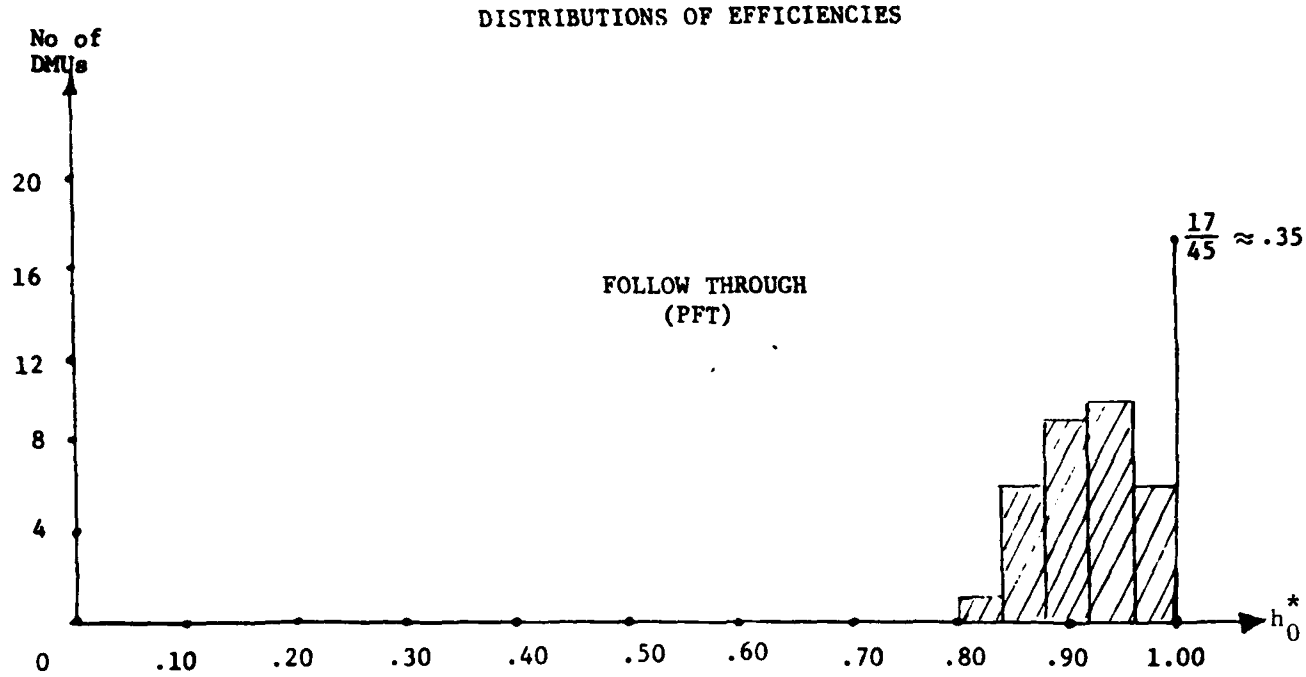
For convenience of reference we now depict the data of Table A-6 as in Figure 3. The latter can be interpreted in terms of frequency diagrams from which a statistical comparison between PFT and NFT might be effected -- e.g., to ascertain whether the two differ sufficiently so that we might reasonably reject the hypothesis that there is no difference in the managerial efficiency in these two distributions. We shall not undertake such an analysis in detail, however, since this would divert us into technical statistical considerations and away from our main objectives.[1/] Instead we call attention to the ratios associated with the spikes on the right of each distribution and observe that these ratios do not differ significantly. That is, there is no statistically significant difference between the proportion of DMU's that achieve an efficient rating in PFT and NFT, respectively.

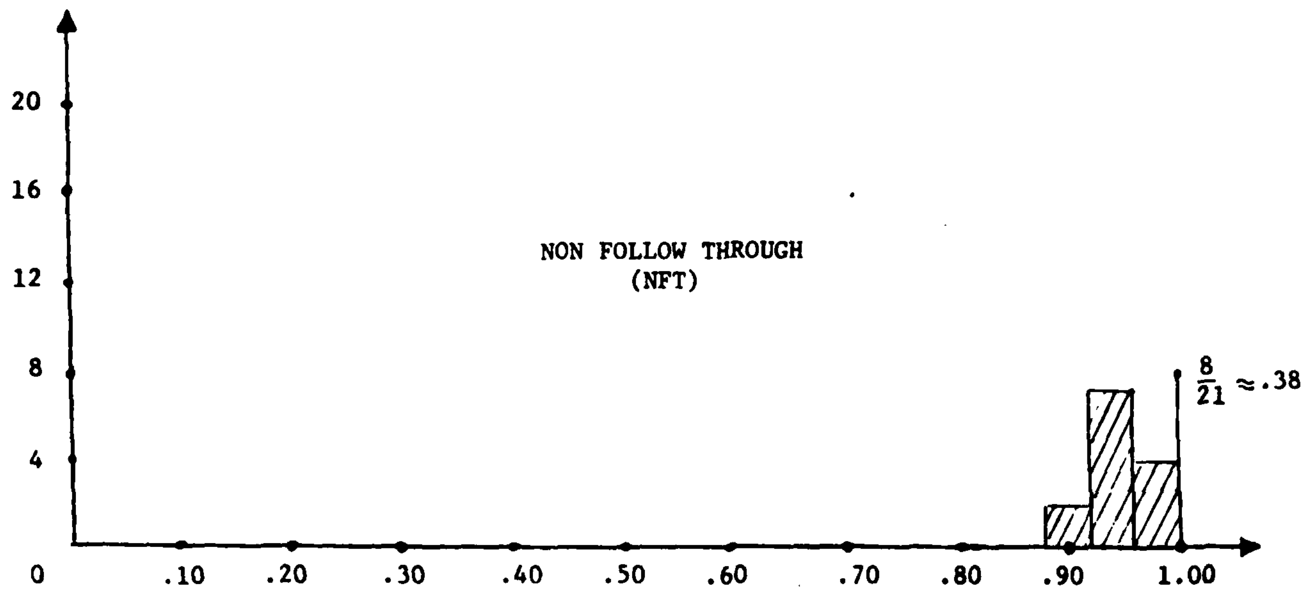
---

[1] A formulation which may be regarded as canonical for the unusual types of density functions involved in these distributions may be found in the Appendix to [6]. which is coauthored with E. Rhodes.

Figure 3

DISTRIBUTIONS OF EFFICIENCIES





This, of course, does not end the analysis, since we are also interested in comparing the programs *per se*. That is we also want to effect a comparative assessment of PFT and NFT on the assumption that the DMUs in each are all efficiently operated.

This brings into play the next part of our DEA approach in a way that we can illustrate with the aid of Figure 4. This diagram, which is purely hypothetical, may be interpreted in a manner analogous to Figure 1, except that we are here considering PFT and NFT isoquants in the same diagram. Referring therefore to observations such as A, we may characterize it as inefficient relative to the PFT envelope while a point such as C is characterized as inefficient relative to the NFT envelope.

The latter, i.e., the PFT and NFT envelopes are derived in the manner we have already indicated. To the left of the point where they intersect the solid line associated with the PFT envelope is evidently less efficient than the solid line associated with NFT[1] and the opposite situation holds thereafter. In such a situation, one may conclude that with efficient operations the NFT program may best be used under one set of conditions and PFT under another set of conditions.[2]

Something more than this may be wanted, however, in the form of an "overall" evaluation of PFT and NFT. For this purpose we arrange an "Inter-Program Envelope" which we abbreviate to "Inter-Envelope" such as the one portrayed by the connected series of broken line segments that are also shown in Figure 4.[3] Observe that this broken line isoquant is coincident with NFT in some sections and that it is coincident with PFT in other sections. In no case is it ever less efficient than either. In other

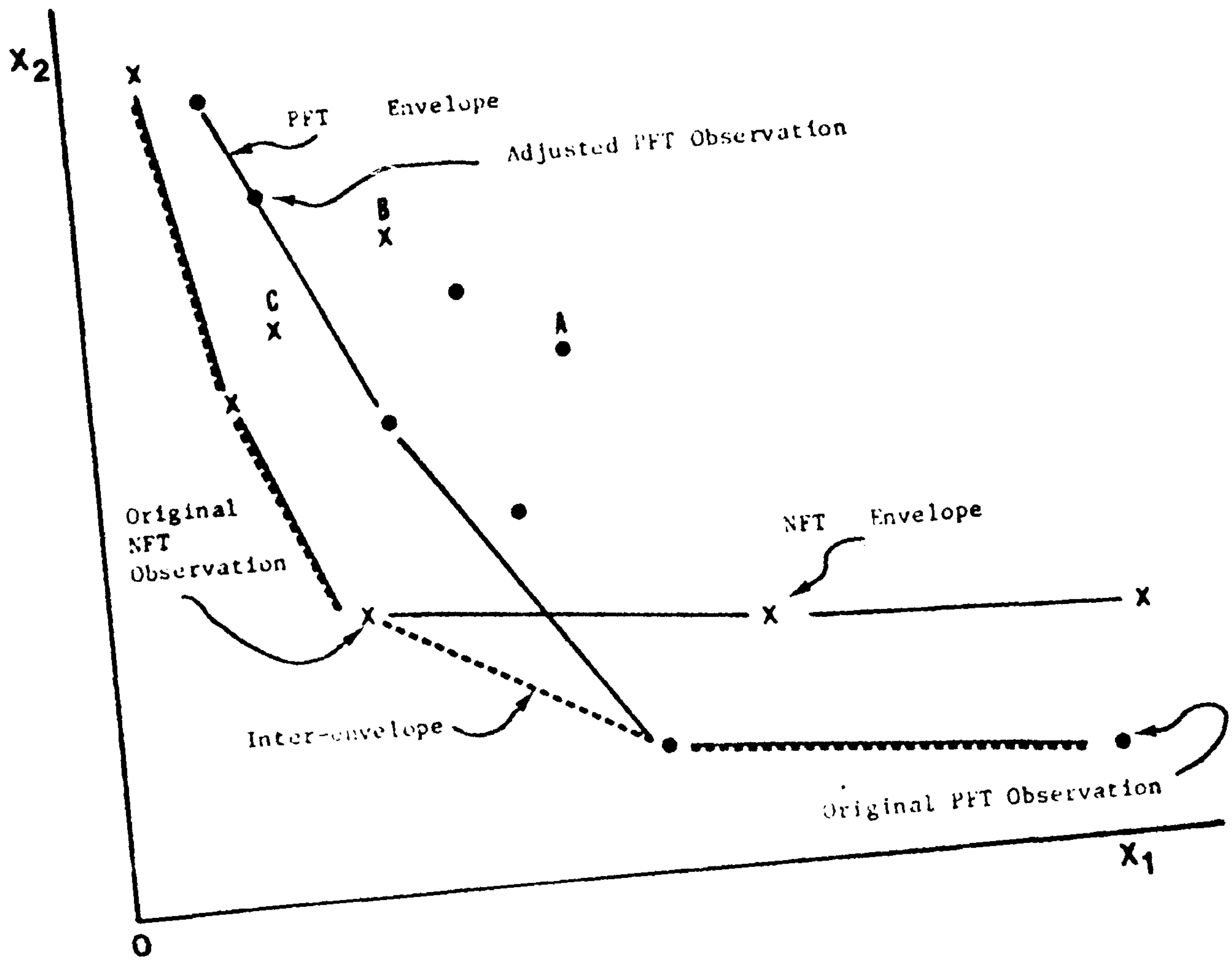
---

[1] Because to the left of the intersection the NFT envelope (= isoquant) indicates that the same level of output is achieved with a smaller amount of at least one input.

[2] For further discussion and references see [8].

[3] These are also distinguished as "each" and "every" evelopes in the terminology of [26]

Figure 4



Legend:

● = PFT Observations
X = NFT Observations
——●—● = Envelope for PFT
——X—X = Envelope for NFT
.... = Inter-envelope

words, the inter-envelope is intended to provide a standard of what is best attainable in the sense of a boundary to which efficient DMUs would push under either NFT or PFT as the efficiency situation might indicate.

Of course, we also need to take account of the number of DMUs that would thereby be improved as well as the amount of improvement in each case. For this we first enforce the assumption of efficiency by bringing every DMU onto its program boundary and then assess the further gain that may be effected by allowing the thus adjusted observations to move from the boundaries imposed by their respective programs to the (possibly) still further efficiencies associated with the inter-envelope.

The methods for effecting these adjustments have been set forth in detail in references such as [9] and [26].[1] Hence we do not need to do more than illustrate what is involved by reference to (12), below. Here we are supposing that site number 1 is to be rated by reference to the entire set of 70 constraints that are set forth under (11.1) and (11.2) in the preceding section after they have first been adjusted in the manner exhibited in Figure 4.[2]

$$\max.\ h_o = \frac{54.53u_1 + 58.98u_2 + 38.16u_3}{86.13v_1 + 16.24v_2 + 48.21v_3 + 49.69v_4 + 9v_5}$$

subject to:

(12)

$$1 \geq \frac{54.53u_1 + 58.98u_2 + 38.16u_3}{86.13v_1 + 16.24v_2 + 48.21v_3 + 49.69v_4 + 9v_5}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$1 \geq \frac{18.82u_1 + 20.04u_2 + 13.77u_3}{10.44v_1 + 4.82v_2 + 17.13v_3 + 17.72v_4 + 7v_5}$$

$$u_1, u_2, u_3, v_1, v_2, v_3, v_4, v_5 > 0.$$

To explain what is involved in such adjustments we might first observe that site 1, which appears in the functional and in the first constraint, is the same as in (11.2) since -- see Table A-6 -- it was rated as efficient and hence requires no adjustment. That is, site 1 is already on the PFT envelope and hence nothing further is required for it.

The same is not true for the site 70 coefficients which appear in the last constraint in (12). Comparison with (11.2) will show that all of the output coefficients (in the numerator) have been adjusted

---

[1] See also [8].

[2] I.e., after the observations have all been brought on to their respective program envelopes. We might also note in passing, however, that a still different envelope may be generated by affecting the efficiency ratings by reference to all 70 constraints without first effecting any such adjustments. See [26].

upwards and some of the input coefficients (in the denominator) have been decreased in order to bring site 70 onto the NFT envelope.

To see how this is accomplished we first observe that the efficiency of rating of site number 70 is $h_0^* = 0.94$. Thus, to adjust its $y_1 = 13.69$ component as given in the last row of Table A-4, we write

$$\frac{13.69}{.94} + 4.31 = 18.82$$

where the 4.31 increment represents slack secured by solving the linear programming equivalent of (11.2).[1]

Similarly, we adjust the originally observed $y_2 = 14.19$ component via

$$\frac{14.19}{.94} + 5 = 20.04$$

and, finally, the $y_3 = 12.99$ component via

$$\frac{12.99}{.94} = 13.77$$

since this output does not involve any positive slack in its optimal solution.

The above values represent the augmented outputs that site 70 would need to have achieved in order to be rated as efficient under NFT. This does not end the matter, however, since the program solution indicates that some of the inputs should also have been lower. In particular, the observed value $x_4 = 18.21$ for this site should be replaced by $\hat{x}_4 = 17.72$ and $x_5 = 9$ by $\hat{x}_5 = 7$. In other words parent counseling activities and number of teachers would also have had to be reduced (in addition to the indicated output augmentations) if site 70 is to be positioned on its portion of the NFT efficiency envelope.[2/]

---

[1] Further details and discussion may be found in [9] and [26].

[2] The point to bear in mind is that input dimensions as well as output dimensions formed part of the agreed upon measures of effectiveness by which the Program Follow Through experiment was to be evaluated.

Table A-7 presents the results of carrying out such calculations for the inter-envelope efficiency ratings. Statistical tests might now be conducted, but inspection of Table A-7 should suffice to indicate that PFT is less efficient than NFT.[1] This is evident both in the number of DMUs <u>on</u> the PFT envelope which are not efficient relative to the inter-envelope as well as the fact that the discrepancies from unity are generally larger than is true for the DMUs that have been brought onto the NFT envelope.[2]

---

[1] See [8] and [26] for a detailed discussion.

[2] It is of interest perhaps to note that the adjusted site 70 is now efficient relative to the inter-envelope whereas site 1 is not.

Table A-7

Inter-Envelope Efficiency Values

PFT Site #	h_0^* Efficiency Value	NFT Site #	h_0^* Efficiency Value
1	0.92	50*	1.00
2*	1.00	51*	1.00
3	0.94	52*	1.00
4*	1.00	53*	1.00
5	0.93	54*	1.00
6*	1.00	55	0.99
7	0.99	56*	1.00
8*	1.00	57*	1.00
9	0.98	58*	1.00
10	0.92	59*	1.00
11*	1.00	60	1.00
12*	1.00	61*	1.00
13	0.99	62*	1.00
14	0.95	63*	1.00
15*	1.00	64*	1.00
16*	1.00	65*	1.00
17*	1.00	66*	1.00
18*	1.00	67*	1.00
19	0.99	68	0.99
20*	1.00	69*	1.00
21*	1.00	70*	1.00
22*	1.00		
23	0.99		
24*	1.00		
25*	1.00		
26	0.99		
27*	1.00		
28*	1.00		
29	0.99		
30*	1.00		
31	0.99		
32*	1.00		
33	0.99		
34	0.98		
35*	1.00		
36*	1.00		
37	0.94		
38	0.99		
39*	1.00		
40	0.95		
41	0.99		
42*	1.00		
43	0.99		
44*	1.00		
45	0.99		
46*	1.00		
47*	1.00		
48*	1.00		
49*	1.00		

*Denotes a site with an efficiency value of "1"

We may now observe that the efficiency values presented in Table A-7 provide us with the wanted estimates of output augmentation and/or resource conservation. That is, these ratios have the same operational significance that we earlier accorded to DMU efficiency except, of course, they now refer to potential gains by removing the boundaries to efficiency that are associated with the program -- ie., the rules and requirements of PFT.

In the way of still further possibilities, however, we might also return to Figure 4 and observe that one portion of the inter-envelope is not coincident with either PFT or NFT. This suggests that there may be combinations of PFT and NFT which are better than either. Such possibilities also need to be considered as part of a PFT vs. NFT decision and, of course, as Figure 4 also shows, there may be situations in which PFT is better than NFT - and so on.

CONCLUSION

These inferences concerning the relative efficiencies of NFT and PFT could be developed by formalized statistical tests. We do not undertake these here, however, in order to avoid the technical developments that would be required to deal with properties such as are involved in the unusual statistical distributions for these measures. Cf., e.g., Figure 3. The procedure associated with DEA also involve certain novel features besides those already indicated. In particular the procedures of Data Envelopment Analysis reverse the relations between statistics and substantive theory that are customary in most of the social sciences. The latter -- i.e., customary usages --

generally apply statistical methods to the relations posited by substantive theory as a condition precedent to their potential use. In DEA, on the other hand, the substantive theory of economics -- as formalized in the definition of a production function, say -- is first used to adjust the data prior to any statistical testing (e.g., along lines such as were indicated in connection with Figure 4).

Of course, the point is that we are here seeking policy evaluations rather than tests of the substantive theory (e.g., production theory) we have been employing. Furthermore, as we indicated in our discussion of "control predictions," our proposed usage of these methods depends on our being in a position to enforce the assumptions (such as efficient production) which we are employing. Given this situation, however, we can then say that our proposed DEA approach can help to supply part of the sustained pressure that is needed for efficiency in public organizations or not-for-profit entities when the market methods of competitive economics are not a really viable alternative.[1/]

This is not to deny that there are numerous activities in subdivisions of commercial enterprises where such alternatives to a direct application of market forces may be useful. Indeed, in the case of regulated industries such as public utilities, etc., the DEA approach provides an alternative that is superior to more standard econometric approaches for potential use in the kinds of "efficiency audits" that are now being ordered by certain public utility commissions.[2]

---

[1] For more detailed discussion along with other possible developments and substitutes, see . [7].

[2] See San Miguel [2?] for further discussion.

The distinction between program and management efficiency can also be adapted to the case where technological rather than program efficiencies provide the pertinent alternative dimensions. The comparison need not proceed by pairs such as NFT vs. PFT since multiple technologies or programs can also be accommodated.[1]

We might also say that these alternatives provide only a beginning for still further possibilities. Distinctions between organization and technological efficiencies might be studied, for instance, in order to evaluate various ways in which different decision makers might be organized under stipulated technologies and vice versa. See [6]. How extensions to such further new alternative uses of DEA might be effected for the still evolving activities involved in "comprehensive audit" certainly offers a variety of exciting possibilities. This, in turn, should help to produce an evolution in which subjects like managerial economics and managerial accounting will begin to include significant public sector as well as private sector decision-making problems on enhanced scales.[2] Such an alteration in educational courses and texts should contribute to improve decision making in public sector as well as private sector managements and thereby enhance the prospects for the expanded dimensions for accountability that are implied in presently evolving practices such as corporate social reporting and comprehensive audits.[3] Improved decision making and enhanced accountability could then go hand in hand toward ameliorating many of the perplexing problems that we seem to continually confront in maintaining a reasonable semblance of social order and consensus while improving the quality of life in our ever more rapidly changing technologically oriented society.

---

[1] See [8] and [26].

[2] See [7] for further discussion and compare with discussion in E. Richardson [27].

[3] See [31] and [14] and [15].

BIBLIOGRAPHY

[1] American Institute of Accountants Committee on Social Measurement, The Measurement of Corporate Social Performance, (New York: American Institute of Certified Public Accountants, 1977).

[2] Banker, R. "A Game Theortic Approach to Measuring Efficiency," European Journal of Operational Research (forthcoming).

[3] Burton, J. C. "Management Auditing," The Journal of Accountancy, May, 1968, pp. 41-46.

[4] Canadian Government, Crown Corporations, Direction Control and Accountability (Ottawa: Privy Concil Office, 1977)

[5] Charnes, A., C. Colantoni and W. W. Cooper, "A Futurological Justification for Historical Cost and Corporate Social Reporting," Accounting, Organizations and Society I, No. 4, 1976, pp. 315-337.

[6] ___________ and W. W. Cooper, "Management Science Relations for Evaluation and Management Accountability." Paper presented as the plenary address on June 20, 1979, for the Institute of Management Sciences XXIV Internations Meetings in Hawaii.

[7] ___________ ___________, "Managerial Economics -- Past, Present and Future," Journal of Enterprise Management I, No. 1, 1978, pp. 15ff.

[8] ___________, ___________, and E. Rhodes, "A Data Envelopment Analysis Approach to Evaluation of the Program Follow Through Experiment in U.S. Public School Education," Management Science (submitted).

[9] ___________ ___________ ___________, "Measuring the Efficiency of Decision Making Units," European Journal of Operational Research 2, No. 6, Nov. 1978, pp. 429-444.

[10] Churchill, N. C. and W. W. Cooper "Comprehensive Auditing," Paper presented to Conference of The Chartered Accountants of Ontario, Canada, June 12, 1978.

[11] ___________ , ___________, J. San Miguel, V. Govindarajan and J. Pond, "Comprehensive Audits and Some Suggestions for Research," Symposium on Auditing Research, II, (Dept. of Accountancy, University of Illinois at Urbana-Champaign, 1977).

[12] Cooper, W. W. and Y Ijiri, eds., Eric Louis Kohler -- Accounting's Man of Principles (Reston, VA: Reston Publishing Co., 1979).

[12.1] Brasfield, K. A. "Development of GAO Auditing Standards and Their Relevance to the Practice of Public Accounting."

[12.2] Cooper, W. W. and W. Frese, "Turnaround at the GAO."

[12.3] Pois, J., "Impact of the Corporation Audits Division upon the U.S. General Accounting Office."

[13] Deloitte, Haskins & Sells, "Audit Committees" (New York: Deloitte, Haskins & Sells, 1979).

[14] Dierkes, M., ed., Soziale Daten und Politische Planung (Frankfurt: Campus Verlag, 1975).

[15] _______________, "Was sind Sozialbilanzen? -- Vorschläge zur Lösing der terminologischen Sprachverwirrung" in E. Pieroth, ed., Sozialbilanzen in der Bundesrepublik Deutschland (Düsseldorf: Econ Verlag, 1978).

[16] Govindarajan, V., Behavioral Effects of Audits in Health Care Services -- A Study in the Field Using Exploratory Data Analysis and Classical Statistical Techniques, (Boston, Doctoral Dissertation, Harvard University Graduate School of Business Administration, 1978).

[17] Macdonell, J. C., "Opening Statement at a Meeting of the Standing Committee on Public Accounts" (Ottawa: Office of the Auditor General of Canada, 100th Annual Report, Nov. 23, 1978).

[18] Morse, E. H., Jr., "Operational and Program Auditing," Paper presented at the Accounting Colloquium, The University of Texas at Austin, Texas 1974-75.

[19] New York Stock Exchange "Recommendations and Comments on Financial Reporting to Shareholders and Related Matters" (New York: White Paper of the New York Stock Exchange, 1973).

[20] Normanton, E. L., "Accountability and Audits of Public Enterprise" (Brussels: International Board of Auditors for NATO, 1979).

[21] _______________ "Public Accountability and Audit: A Reconnaissance," in B. R. Smith and D. C. Hague, eds., The Dilemma of Accountability in Modern Government -- Independence vs. Control (London: MacMillan & Co., 1971).

[22] _______________ Reform in the Field of Public Accountability and Audit," Paper presented at the Harvard University Seminar on Planning and Accountability Systems, 1979.

[23] Peat, Marwick, Mitchell & Co. "A Peer Review Report, 1978" (New York: Peat, Marwick, Mitchell & Co., March 16, 1979).

[24] Pois, J., Watchdog on the Potomac (Washington, University Press of America, Inc., 1979).

[25] Public Oversight Board, Annual Report 1978-79 (New York: Public Oversight Board, American Institute of Certified Public Accountants, 1978).

[26] Rhodes, E., Data Envelopment Analysis and Related Approaches for Measuring the Efficiency of Decision Making Units with an Application to Program Follow Through in U.S. Public School Education (Pittsburgh: Carnegie-Mellon University, School of Urban and Public Affairs, Ph.D. Thesis, 1978).

[27] Richardson, E., The Creative Balance -- Government, Politics and the Individual in America's Third Century (New York: Holt Rinehart and Winston, 1976) pp. 138-142.

BIBLIOGRAPHY (Continued)

[29] San Miguel, J., "An Analysis of Management Audits in the Electric Utility Industry," unpublished draft, Harvard Graduate School of Business Administration.

[30] ____________, "Portland General Electric Company," Intercollegiate Case Clearing House, Boston, Massachusetts, 1978.

[31] Schoenfield, H. M., ed., The Status of Social Reporting in Selected Countries, Contemporary Issues in International Accounting, Occasional Paper No. 1 (Urbana: University of Illinois Center for International Research in Accounting, 1978).

[32] Staats, E. B., "Auditing in the Twenty-First Century -- New Challenges to be Met." Plenary Session Paper, American Accounting Association Conference in Honolulu, August 24, 1979.

[33] ____________, "Comprehensive Auditing -- Planning the New Approach for Century II," Proceedings of the Centennial Conference of the Auditor General of Canada, 1978 (to appear).

[34] U.S. Securities & Exchange Commission, Report to Congress on the Accounting Profession and the Commission's Oversight Role, Washington: U.S. Securities & Exchange Commission, Mimeo, July 1979).

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Center for Cybernetic Studies The University of Texas at Austin	2a. REPORT SECURITY CLASSIFICATION Unclassified
	2b. GROUP

3. REPORT TITLE

Auditing and Accounting for Program Efficiency and Management Efficiency in Not-for-Profit Entities

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

5. AUTHOR(S) *(First name, middle initial, last name)*

A. Charnes and W.W. Cooper

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
August 1979	50	34

8a. CONTRACT OR GRANT NO. N00014-75-C-0616 & 0569	9a. ORIGINATOR'S REPORT NUMBER(S) CCS 343
b. PROJECT NO. NR 047-071 & 021	
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office of Naval Research (Code 434) Washington, DC

13. ABSTRACT

A measure of efficiency for not-for-profit entities--developed by the authors in association with Edward Rhodes--is explained and illustrated by data from Program Follow Through, a large scale social experiment in U.S. public school education. A division into Follow Through and Non-Follow Through participants facilitates a distinction between "program efficiency" and "managerial efficiency" which is also illustrated and examined for its use in evaluating such programs. Relations to comprehensive audits and other possible uses are explored.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6811

Unclassified

Security Classification

A-31408

Unclassified
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Efficiency						
Program Efficiency						
Managerial Efficiency						
Comprehensive Audit						
Data Envelopment Analysis						
Accountability						
Ratio Programming						

DD FORM 1 NOV 65 1473 (BACK)
S/N 0102-014-6800

Unclassified
Security Classification

A-31409