USAFA-TR-77-12

AD A046105

A GENERAL TECHNIQUE FOR R & D COST FORECASTING

MAJOR WILLIAM J. WEIDA

DEPT OF ECONOMICS, GEOGRAPHY
AND MANAGEMENT
USAF ACADEMY, COLORADO 80840

DDC
NOV 7 1977
F.

SEPTEMBER 1977

FINAL REPORT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

AD No.
DDC FILE COPY

DEAN OF THE FACULTY

UNITED STATES AIR FORCE ACADEMY

COLORADO 80840

Editorial Review by Lt Col Elser
Department of English
USAF Academy, Colorado 80840

This research report is presented as a competent treatment of the subject, worthy of publication. The United States Air Force Academy vouches for the quality of the research, without necessarily endorsing the opinions and conclusions of the author.

This report has been cleared for open publication and/or public release by the appropriate Office of Information in accordance with AFR 190-17 and DODD 5230.9. There is no objection to unlimited distribution of this report to the public at large, or by DDC to the National Technical Information Service.

This research report has been reviewed and is approved for publication.

PHILIP G. ERDLE, Colonel, USAF
Vice Dean of the Faculty

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: USAFA-TR-77-12

2. GOVT ACCESSION NO.:

3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): A GENERAL TECHNIQUE FOR R & D COST FORECASTING.

5. TYPE OF REPORT & PERIOD COVERED: FINAL REPORT.

6. PERFORMING ORG. REPORT NUMBER:

7. AUTHOR(s): Major William J. Weida

8. CONTRACT OR GRANT NUMBER(s):

9. PERFORMING ORGANIZATION NAME AND ADDRESS: Department of Economics, Geography and Management, United States Air Force Academy, Colorado 80840

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:

11. CONTROLLING OFFICE NAME AND ADDRESS: DFEGM, USAF Academy, CO 80840

12. REPORT DATE: Sep[redacted]77

13. NUMBER OF PAGES: 64

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):

15. SECURITY CLASS. (of this report): 65p.

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report): Approved for Public Release; Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):

18. SUPPLEMENTARY NOTES:

19. KEY WORDS (Continue on reverse side if necessary and identify by block number): Cost Forecasting, R & D Cost, Research and Development Cost, Growth Curves, Expenditure Pattern Analysis

20. ABSTRACT (Continue on reverse side if necessary and identify by block number):

A general model for R & D cost forecasting was developed based on an expenditure pattern analysis of twenty-one current weapon systems. This model which was validated on an additional twelve weapons systems, shows that R & D expenditures follow a certain, well-defined pattern, regardless of the type of system involved.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

011 550

TABLE OF CONTENTS

		Page
INTRODUCTION.		1
CHAPTER I.	Fitting the Curve	5
CHAPTER II.	The Rationale Behind the Use of the S-Shaped Curve	10
CHAPTER III.	Developing and Using the S-Shaped Curve	16
CHAPTER IV.	The S-Shaped Curve as a Forecasting Tool.	20
CHAPTER V.	An Application.	35
BIBLIOGRAPHY.		46
APPENDIX I:	Logistics and Gompertz Curve Equations.	47
APPENDIX II:	Removing Autocorrelation Through GLS.	48
TABLES I through IV:	Weapon Systems R & D Curve Data	52-55
TABLE V:	The Relationship Between Milestones and Budget Data.	55
TABLE VI:	Inflection Point Locations.	56
TABLE VII:	Project B Budget Data	57
TABLE VIII:	Project B Actuals	58
EXHIBIT I:	General Curve-Bottom Half	59
EXHIBIT II:	General Curve-Top Half.	60
EXHIBIT III:	Budget Curve-Bottom Half & Top Half	61-62
EXHIBIT IV:	Actuals-Bottom Half	63

ACCESSION for
NTIS
DDC B ff Section
UNANNOUNCED
JUSTIFICATION
BY
DISTRIBUTION/AVAILABILITY CODES
Di CIAL
A

A GENERAL TECHNIQUE FOR

R & D COST FORECASTING

INTRODUCTION

During the 1920's, two statisticians, Raymond Pearl and L.J. Reed, were extensively involved in research dealing with forecasting rates of population and biological growth. They discovered that the S-shaped curve form of Figure 1 accurately represented growth of these types which are characterized by three distinct phases: (1) a slow period of development, (2) a rapid period of expansion, and (3) a tapering off at maturity.

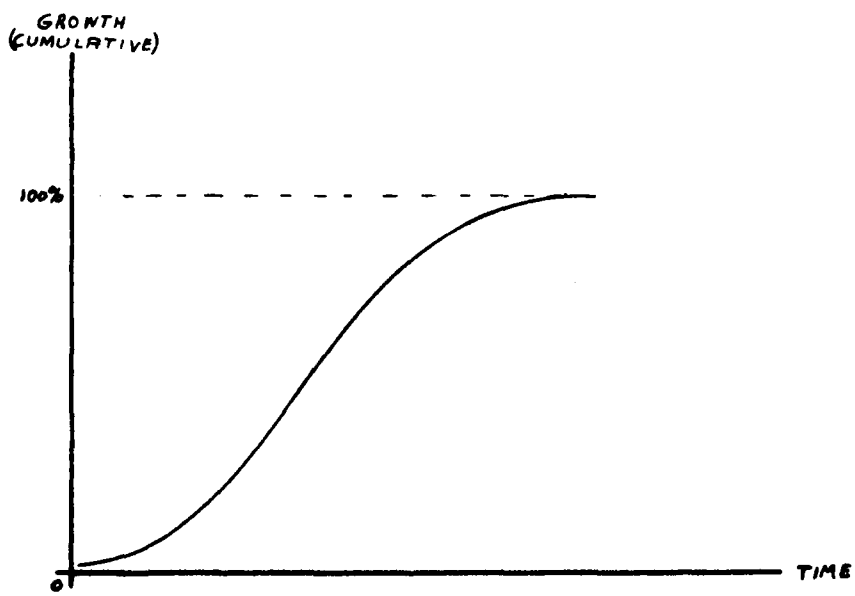

Figure 1. The Growth Curve

The growth rates of a number of developing industries were also found to follow this general format, but most attempts to use this curve form for forecasting were non-productive. The failure of the two forms of the S-shaped curve (the Logistic and the Gompertz)(See Appendix 1) to provide accurate forecasts can be attributed to two principle causes:

(1) The equation for the Gompertz curve assumes symmetry around an inflection point which is the geometric mean of the Y values, while the equation for the Logistics curve makes the same assumption for the true mean of the Y values. Neither of these assumptions is usually justified in real-world situations.[1]

(2) Because of the order of the polynomials needed to express these curves, slight perturbations of the data early in the growth process can quickly force the equation for either of the curves to forecast an unrealistically high or low figure for total growth.

This failure to forecast accurately becomes important in the context of Research and Development within the Department of Defense because R & D programs also follow the S-shaped curve form in terms of both cumulative dollar expenditures and cumulative milestone completion over time. This fact should be no surprise since the R & D process is intuitively a biological growth process in which there is slow initial <u>development</u> of the specific project, followed by a fairly rapid <u>building</u> phase which, in turn, is followed by a <u>tapering-off</u> to completion.

---

[1] Table VI shows the actual inflection point for 21 weapon systems. Note that in no case does this point consistently fall anywhere close to either the true or the geometric mean.

The existence of these three phases in Department of Defense R & D programs was verified by taking the Cost Performance Reports (CPRs) for 22 different weapon and component systems and tracing the actual costs as they were incurred throughout the life of each R & D program. Without exception the S-shaped curve was the dominating factor in the system development as it actually occurred. The milestones associated with these programs also exhibited the same characteristics in every case for which data were available.

The results based on the original 22 weapon systems were then checked against new and very complete data which were provided by another source on 15 weapon systems. Again the connection between R & D expenditures, the S-shaped curve, and milestone completion were firmly established.

Conventional wisdom would suggest that since the Logistics and Gompertz curve forms are generally non-productive forecasting tools, one should proceed instead to construct an econometric model of the R & D process. This model, which would embody each of the endogenous variables in any given R & D project, could then use the historical cost data for each of these inputs (labor, raw materials, etc.) to establish trends. The trends could then be combined to provide an overall forecast of the final cost of the R & D project.

Unfortunately, the existing data concerning the prices of many raw materials such as aluminum, stainless steel, titanium, steel alloys, etc., are either misleading, erroneous, or simply do not reflect the true prices. In addition, labor data are often not available for the narrow categories of labor employed in a given R & D project, nor do they properly reflect local or regional conditions which may significantly affect wage rates.

These deficiencies rule out the use of conventional econometric modeling techniques, and one is faced with the problem of finding some known element common to all R & D programs which could be adapted as a forecasting tool.

Since the S-shaped curve is a common element in all R & D projects, this paper will deal with an effort to modify the curve in such a way that accurate forecasting can be achieved.

CHAPTER I

FITTING THE CURVE

The major problem in handling an S-shaped growth curve is one of keeping under control the higher-order polynomials necessary to express this complicated curve form. If one realizes that any S-shaped curve is merely the cumulative form of a bell curve (which may or may not be skewed) as shown in Figure 2, a solution to this particular problem becomes apparent.

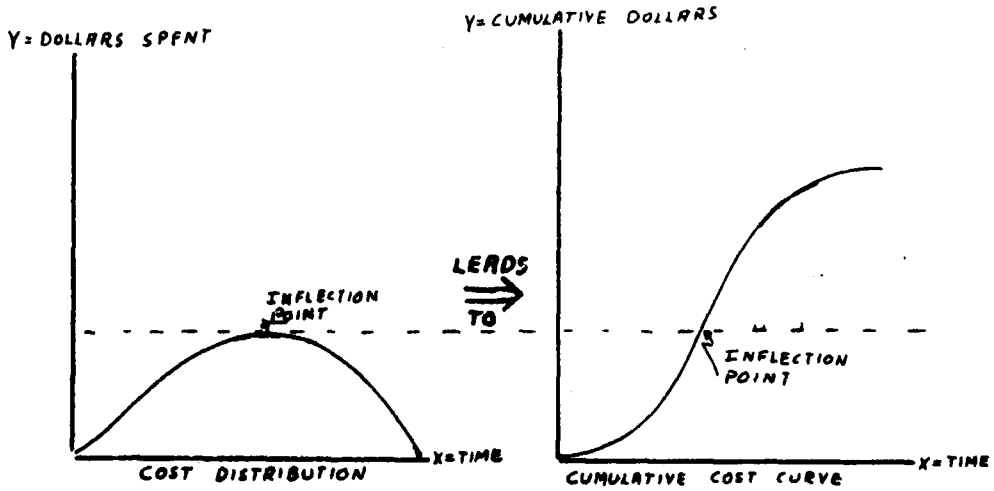

Figure 2. The Derivation of the S-shaped Curve

The S-shaped curve may be separated (or "broken") at the inflection point to yield two simpler curves of a form which may be expressed by either a logrithmic ($y = ax^b$) or a quadratic ($y = a + b_1 x_1 + b_2 x_1^2$) equation. Standard econometric fitting techniques may then be utilized to determine which of these two curve forms is the best fit.

However, selecting the best curve form is greatly complicated by the fact that the time dependent nature of these curves, and indeed the entire R & D process, tends to cause a great deal of autocorrelation in the resulting data. This autocorrelation, which is a violation of the necessary assumption that regression residuals are not related, must be dealt with and eliminated from the data before a valid selection of equations can be made to express the lower and upper halves of the S-shaped curve. For this particular research, autocorrelation was removed through the use of Generalized Least Squares following an autoregressive transformation which used a two-step procedure for the estimation of Rho. This process is explained in its entirety in Appendix 2.

It should be noted that the treatment of each additional order of autocorrelation by this method removes one additional data point from the lower end of whichever half of the S-shaped curve is being considered. This occurs because each succeeding transformation steps back in time one data point further in an attempt to remove any past influences on the current regression

residual. This process has an inherent benefit in that it tends to weight recent observations more heavily than past observations. The result is a rather sophisticated smoothing technique which the user may vary to assure that the fitted curves will converge rapidly on the proper values.

Tables I through IV show the results of this curve fitting technique when it was applied to both actual and budgeted cost figures for 22 major weapon systems or R & D projects. In almost every case, the quadratic equation provided the best fit of both the upper and lower segments of the S-shaped curve. The statistics accompanying each of the curve equations indicate that the upper and lower curve segments, when rejoined, do provide an excellent proxy for the original S-shaped curve for each R & D project, and the mean square error figures accompanying each equation assure that these reconstructed S-shaped curves are useful forecasting tools.

With the curves for each of the specific programs developed, the same technique of calculating the upper and lower halves of the S-shaped curve and then mating these segments at the inflection point can be used to derive a general curve based on all weapon systems for which data are available. This general curve can then serve as the forecasting vehicle for new weapon systems.

The actual calculations for the general curve are accomplished by pooling all of the data for 17 systems. Table VI shows that the

mean inflection point for these systems occurs at the 56.2% of expenditures and 46.2% of time points with the width of one standard deviation around this point being 5.4% for expenditures and 7.3% for time. Considering that these figures were derived from a collection of systems which were so diverse as to include an early 1950's era weapon system (the F-105), cargo and bomber aircraft (C-5, C-141, B-1, etc.), engines and guns (A-10), and future projects such as space tugs, these figures represent a surprisingly narrow confidence interval.

Exhibits I and II show the results of fitting the curve segments using a quadratic equation for both the lower and upper halves. A certain degree of heteroskedasticity was introduced into these data sets by a normalization of the data which forced the S-shaped curve to begin at the 0% time and expenditures point and end at the 100% time and expenditures point. Because of the obvious significance of the data provided in Exhibits I and II, the heteroskedasticity was not deemed to be a significant problem.

Figure 3 shows the general curve form and a 1σ confidence interval around this curve. The calculations for this confidence interval may be found in Table VI. This type of curve derivation has the obvious advantages of generating a result which should have wide application across the broad spectrum of weapon systems, and as such it will form the basis for the forecasting efforts which will be explored in the next chapters.

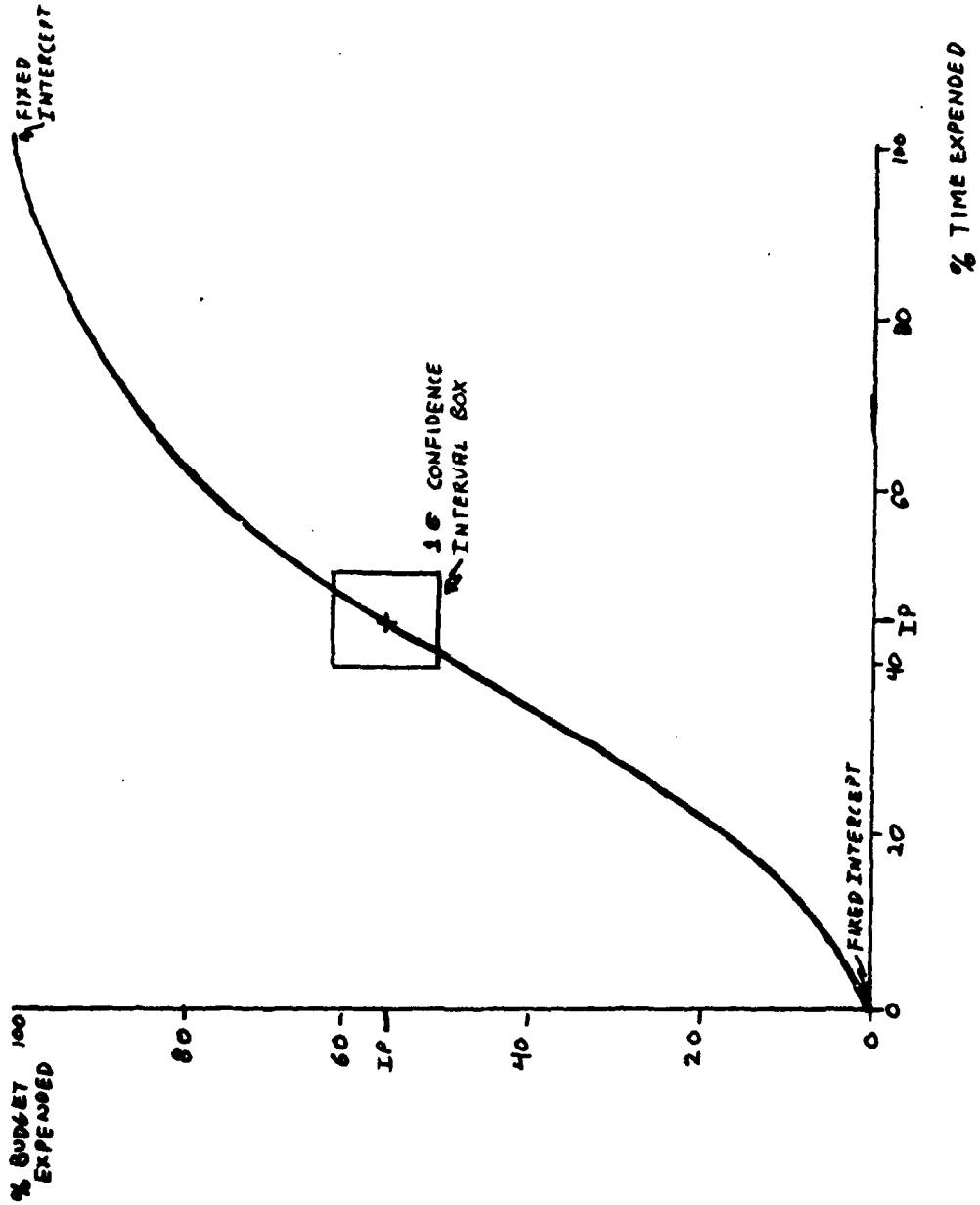

Figure 3. The General R & D Curve

CHAPTER II

THE RATIONALE BEHIND THE USE OF THE S-SHAPED CURVE

Chapter I showed that the S-shaped curve described every military R & D project which was investigated, and Chapter II will demonstrate a practical method of finding the curve for either a specific project or a general model. However, the rationale behind the use of both a two-part curve and the initial budget data in the forecasting process deserves further explanation.

Since no one can estimate the cost of developing a new weapon system without an effort which would virtually duplicate the original contract-letting process, it is obvious that the rational approach is to accept the R & D budget proposed by the contractor as the best initial estimate of the project cost.

However, once this approach is decided upon and it is realized that a two-part S-shaped growth curve can be chosen to express the program as a whole, three very interesting approaches may be adopted:

(1) The proposed project budget may be compared to all past R & D budgets to see if the expenditure pattern is <u>generally</u> rational in light of R & D experience with other weapon systems. Figure 3

shows the general curve with a 1σ confidence band based on the R & D projects which were studied for this paper. Considering the diversity of the weapon systems which were investigated, this band represents a fairly narrow range within which the expenditure pattern for any new R & D project should be expected to fall. If the proposed expenditure pattern falls outside of this range, the contractor should be expected to explain why his program is unique in this respect.

(2) The proposed budget expenditure pattern can be compared to the proposed project milestones to check the specific rationality of the budget proposal. Table V shows a reperesentative selection of projects for which milestone data are available. There is a very high order of correlation between the budget expenditure pattern and the pattern in which the milestones are accomplished. This suggests that the manner in which the milestones are completed over time provides an excellent cross-check on the way in which the budget should be expended. And it also means that slips or changes in the order of milestone completion will invariably be reflected in the expenditures for the entire project.

(3) When the contractor or procurement officer considers uncertainty in any R & D program, he can generally resolve this uncertainty into different types of unknowns. Drake states that

> these unknowns are of two kinds:
> (1) The unknowns that he is aware of and believes he can resolve when he accepts a contract--for example, the configuration of

> an aircraft's slats, flaps, thrust reversers, speed brakes, and other devices needed to meet the specified performance factors (take-off, landing distance, etc.).
>
> (2) The unknowns which are bound to crop up unexpectedly and for which he is not prepared--the 'unanticipated unknowns.' Strictly speaking, they are the unknowns that *cannot* be foreseen.[1]

In other words, type 1 unknowns are known unknowns while type 2 unknowns are unknown unknowns.

If one then pictures himself standing at the beginning of an R & D project and looking down the time line toward the completion of the project, he will see the two different types of uncertainty shown in Figures 4 and 5.

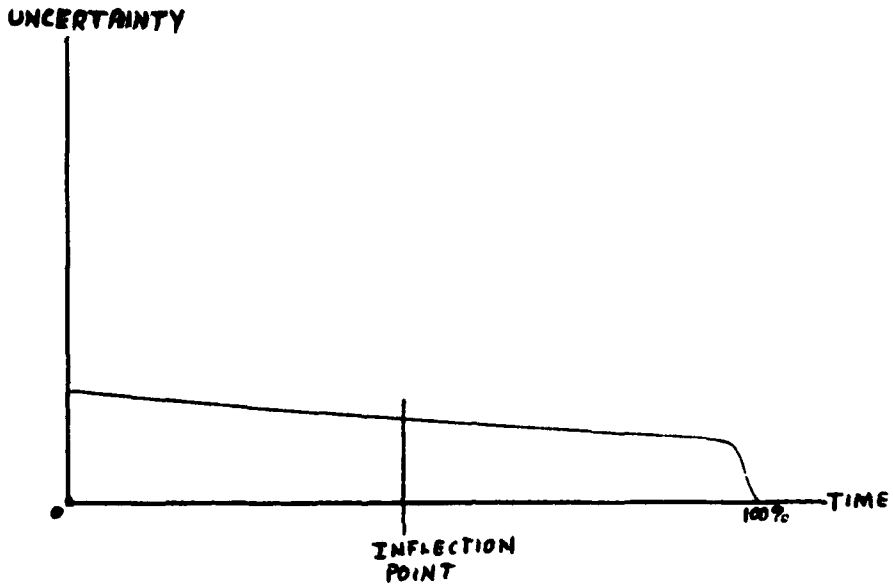

Figure 4. Uncertainty Due to Known Unknowns

[1] Drake, Hudson B., "Major DoD Procurements at War with Reality," The Harvard Business Review, January-February 1970, p. 124.

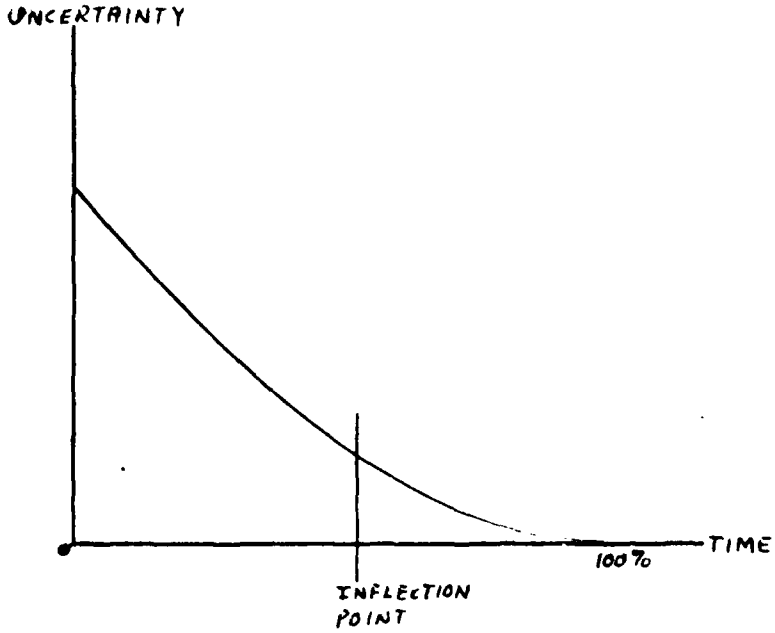

Figure 5. Uncertainty Due to Unknown Unknowns

These two figures can be combined to show that the total program uncertainty is really a kinked curve as shown in Figure 6.

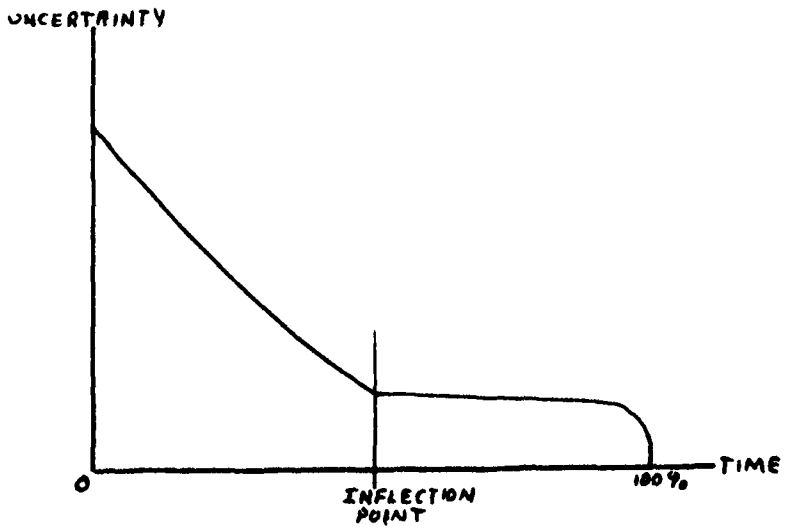

Figure 6. Total Program Uncertainty

The critical point is the inflection point of the program since it is at this point that most of the uncertainty has been dispelled. Thus, accurate forecasting is most essential and also most difficult in the early part of the R & D program. The two-part curve development overcomes the problem of changing uncertainty by allowing the researcher to forecast a new lower curve and inflection point from the actual expenditures in the program being investigated. This forecast may then be constrained by using the top half of either the budget or the general curve in the area where uncertainty is lower and hence, these curves should be the most accurate estimators of program performance.

It should be assumed that since the contractor and the government have both attempted to plan the R & D project in the best possible manner given the financial and political constraints on the program, this planning effort should not be wholly disregarded as soon as the first actual expenditures begin to appear. Instead, it is more rational to let the general framework of the expenditure pattern, as expressed in the budget curve, continue to represent this planning as an inherent constraint during any forecasting activity. Used in this manner, the original budget expenditure curve can be viewed as a storehouse of subjective or judgmental information which may be used to constrain a forecast and thus allow for future imputs by the managers in control of the R & D program.

These three points indicate why the use of the two part S-shape curve is particularly appealing. Not only is the contractor subjected to intense scrutiny during the initial planning phase (points one and two), but once he passes these checks his planning may then be used as a forecasting constraint during the rest of the program.

CHAPTER III

DEVELOPING AND USING THE S-SHAPED CURVE

The general method for the development and use of the S-shaped curve for a specific R & D program follows these steps:

I. The budget figures for the R & D program are gathered and the monthly, quarterly, or other incremental expenditures are recorded as a cumulative percentage of the total expenditure. Similarly, the amount of time over which the program is to run is determined and each succeeding increment is recorded as a cumulative percent of the total program time. This step has the effect of normalizing the program so that it can be compared with all past R & D programs when plotted on the axis of Figure 7.

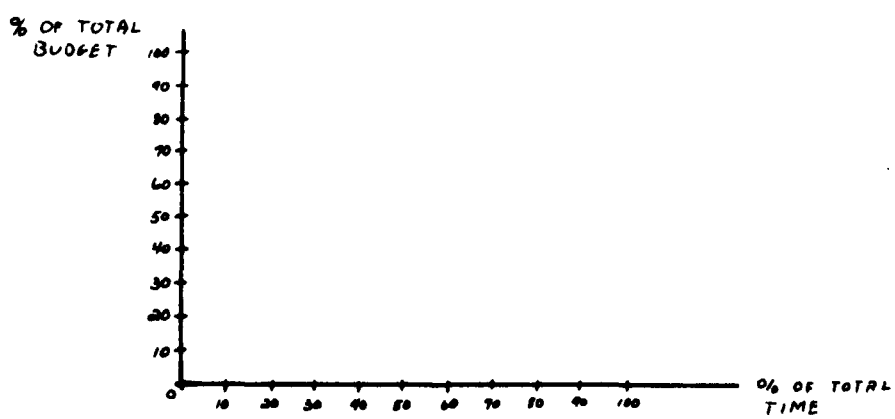

Figure 7. Expenditures vs. Time

II. With the data arrayed in this normalized format and plotted on the axis of Figure 7, the budget expenditure pattern may be immediately checked for general rationality. This is accomplished by determining whether or not the proposed budget curve falls within the confidence band which has been constructed for all past R & D projects (Figure 3). If the proposed budget expenditure pattern falls within this band, the R & D project planning is in accordance with past R & D experience. However, if the expenditure pattern falls outside the confidence band it should be determined why this particular program is planned to proceed differently from all past programs have proceeded.

III. The total number of milestones connected with all phases of the R & D project is determined and the month in which the end or completion point of each milestone occurs is noted. The cumulative percentage of completed milestones is then calculated for each period for which budget data are available. This step has the effect of putting milestone accomplishment in the same normalized form as the financial data.[1] These normalized milestone accomplishment data

---

[1] At first it may seem counterintuitive to sum up milestones in the same manner in which one sums up dollars. However, even though it is obvious that no two milestones represent identical amounts of accomplishment, deriving cumulative sums of milestone completions by financial period has invariably provided an excellent picture of the manner in which the program is expected to proceed. It should be noted that the best results have been obtained when the milestones were expressed on the lowest, or most specific level available. Every attempt should be made in this process to avoid double counting between specific and general program milestones, although no adverse affects have been noted when <u>different</u> specific and general milestones were summed together using this technique.

are then regressed with the normalized budget data. The results of this regression should show an extremely significant relationship between the budget expenditures and the milestone completion. Table V shows the results for five sample programs to give a general idea of the significance of the relationship. If this relationship does not appear, the proposed expenditure pattern for the specific program in question is highly dubious and indicates an unrealistic program plan which could easily lead to financial problems.

IV. Next locate the largest incremental change in budgeted expenditures which is followed by two periods of decreasing expenditures. This increment is designated as the inflection point. The S-shaped curve is broken at this point and the inflection point becomes the last data point in the first (or lower) curve and the first data point on the second (upper) curve. This common point allows the curves to be spliced again after curve fitting. Table VI shows the location of the inflection points for the R & D programs included in the initial part of this study. Note the high degree of variability in the point locations.

V. Equations for the lower and upper portion of the budget curve are developed by the regression scheme discussed in Appendix 2. Due to the shape of the curves, a quadratic ($y = a + b_1 x_1 + b_2 x_1^2$) or logrithmic ($y = ax^b$) curve equation is usually appropriate. Particular care must be taken in this step to assure that the curve equations which are developed have dealt with the problems inherent in the use of time series data. Failure to correct the problem of

autocorrelation will result in curve equations which are of little value and which will adversely affect the performance of the completed model. To avoid this situation, some techniques of the type explained in the appendix must be used at this juncture, and indeed, should be used in <u>any</u> R & D forecasting.

VI. Once the curve equations have been developed from the budget data, two specific types of knowledge have been gained:

A. The equation form which best fits the R & D budget data has now been determined. This is usually a quadratic form for both the upper and lower halves of the S-shaped curve and this specific curve form should be used with any actuals when later attempts are made to forecast the end price of the R & D project.

B. Equations expressing the subjective planning inherent in the R & D program are now available for the upper and lower parts of the S-shaped curve. These original equations can be used as constraints during forecasting, thus providing a method of incorporating this subjective information into the final price forecast. The specific methods by which the two part S-shaped curve may be used for price forecasting are the subjects of the next chapter.

CHAPTER IV

THE S-SHAPED CURVE AS A FORECASTING TOOL

The objective of the methodology advocated in the first three chapters of this paper is obviously the production of a forecast cost for the system being developed. In pursuing this objective it is well to remember Scrooge's question to the Spirit of Christmas Future: "Are these the shadows of things that will be, or are they the shadows of things that may be, only? Men's courses will foreshadow certain ends, to which, if persevered in, they must lead. . . . But if the courses be departed from the ends will change."

This quote conveys the proper manner in which one should employ a forecast to make it an effective management tool. The program manager should view the forecast as a non-threatening means of alerting managers to possible program difficulties and it should be presented not as a point estimate, but rather as a range of values within which the end cost of the program is likely to fall <u>if the present courses of action are continued</u>. For the purpose of this paper three points along this possible range of cost will be identified:

(a) the best possible program cost, (b) the most likely program cost, and (c) the worst possible program cost.

The best possible program cost is the cost of the R & D program if it is assumed that the second half of the program will follow exactly the proposed budget curve irrespective of the performance record established in the first half of the program.

The most likely program cost is that figure obtained if the second half of the program follws the course indicated by the general R & D curve. This general curve being either one developed for all weapon systems, such as that curve displayed in Figure 3, or a curve developed for weapon systems of the specific type of the R & D project being investigated. In other words, a curve based on general missile system data would be used when the R & D project is for a missile. Intuitively, the latter approach should provide a tighter confidence interval for the forecast.

The worst possible program cost would be the figure indicated by the upper limit of the confidence interval around the forecast. The second half of the general curve would again be used, and the confidence interval around this curve would provide the upper limit (within a certain probability) of the R & D cost.

These three types of forecasts are shown in Figure 8.

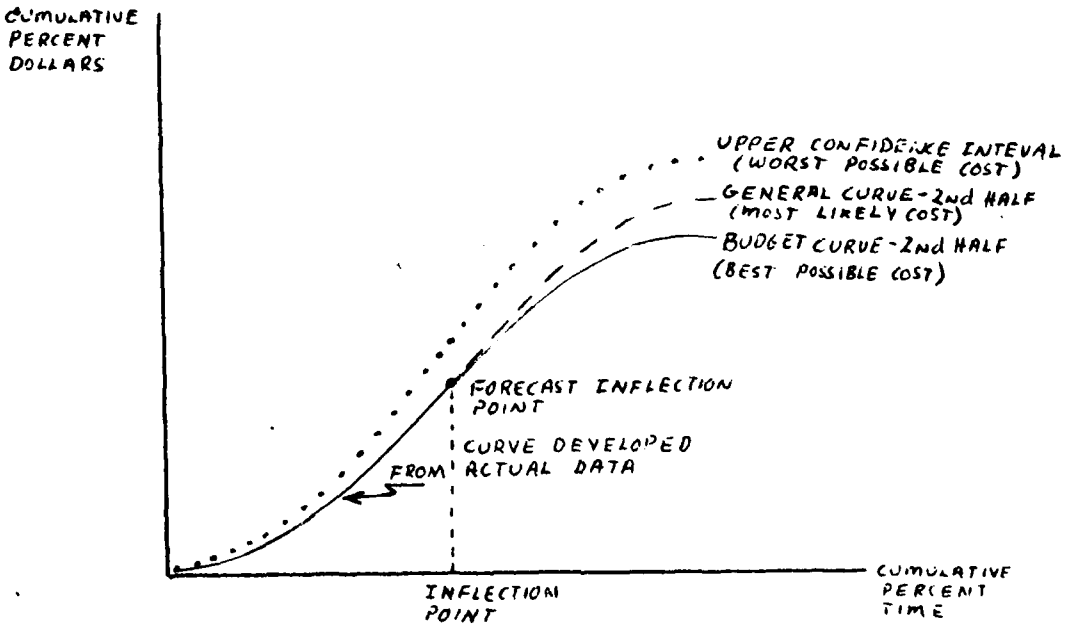

Figure 8. Three Possible Forecasts

from Program Actuals

The specifics involved in forming each of these forecasts will now be covered.

THE BEST POSSIBLE COST

(1) First derive the two halves of the equation for the S-shaped curve in the manner previously outlined. This gives curve 1 of Figure 9, the <u>budget curve</u>.

(2) Assume now that the first data points concerning actual expenditure information have become available. These data points are first deflated by dividing the dollar figures by an appropriate index. Studies have shown that the GNP Deflator is

usually a good choice for this index.[1] The deflated figures are then converted to percentage figures by dividing by the latest deflated total program cost, and these percentage figures are plotted on the axis of Figure 7. This leads to the beginning of an "actuals" curve, shown as curve 2 of Figure 9. These actuals may be used to forecast a new end cost for the program as follows:

(a) Derive a new lower half of the S-shaped curve by fitting the actuals to an equation of the form found to be appropriate for the budget data--in general, this will be a quadratic curve.

(b) Using this quadratic curve equation, insert the percent of total time figure for the budget curve inflection point (35% on Figure 9) to forecast a new inflection point, (Point 5 on Figure 9) and then use other points on the X (time) axis to derive a new lower half for the S-shaped curve.

(c) Now take the equation which was developed for the top half of the budget curve and substitute the percent time and percent budget figures for the forecast inflection point into this equation to calculate a new intercept for the upper curve. This new intercept, along with the original slope figures from the budget curve, has the effect of "splicing" the equation developed

---

[1] Brush, John S., "Study of Possible Improvements in the Accuracy of Aeronautical Economic Escalation Indices," unpublished paper, USAF Academy, Colorado, February 1976.

from the first half actuals to the budget equation for the second half of the curve, all of which yields the new S-shaped curve 4 of Figure 9. In addition, this procedure allows the development of a forecast for the end cost of the project which is constrained by the planning and other subjective information inherent in the original budget curve.

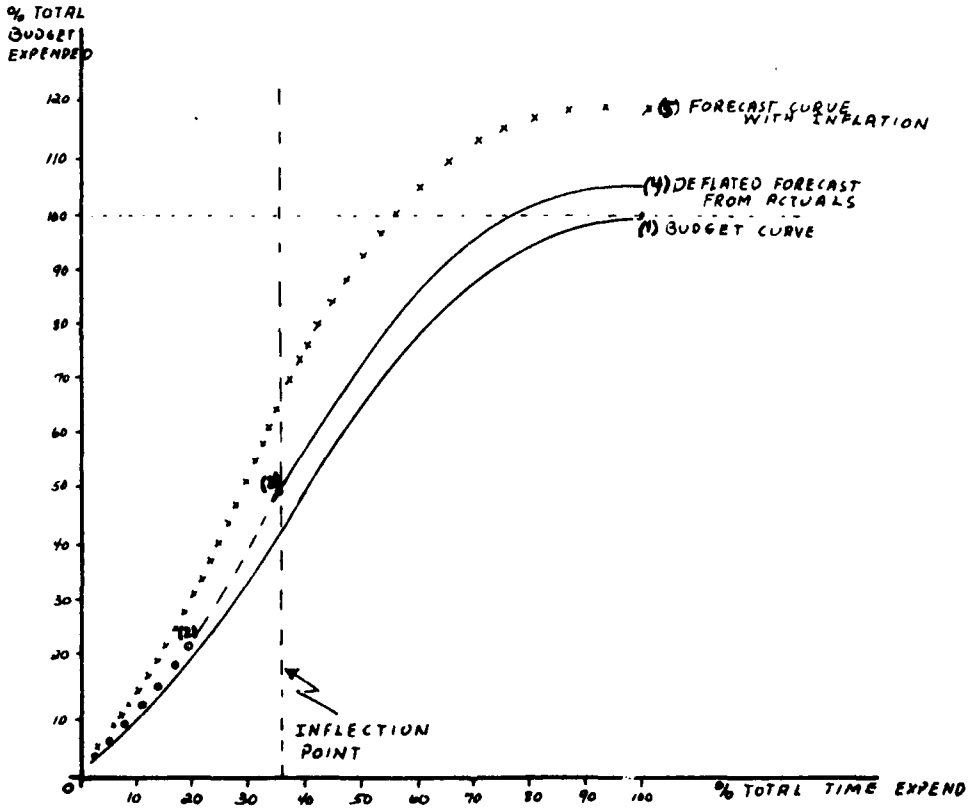

Figure 9. The Forecasting Process

(d) At this point, one may take several different approaches. First, if he wants to learn the absolute figure for the final cost of the project, curve 4 may be modified by

inclusion of inflation data. In this case, the forecast expenditure data of curve 4 would be multiplied by an inflation index to get a new curve which is labeled 5 on Figure 9. However, in doing this one should have in mind a concept of the errors inherent in any process such as the one just described.

Up to this point we have not mentioned, for the sake of simplicity, that there is an error involved in forecasting which should be expressed as a confidence interval around curve 4. The confidence band indicates that, with some given probability, one may expect the real value for any point on the line to fall somewhere within this particular interval. When the budget curve is compared with the forecast curve, only one error, the standard error of the forecast, must be considered. This leads to the situation shown in Figure 10.

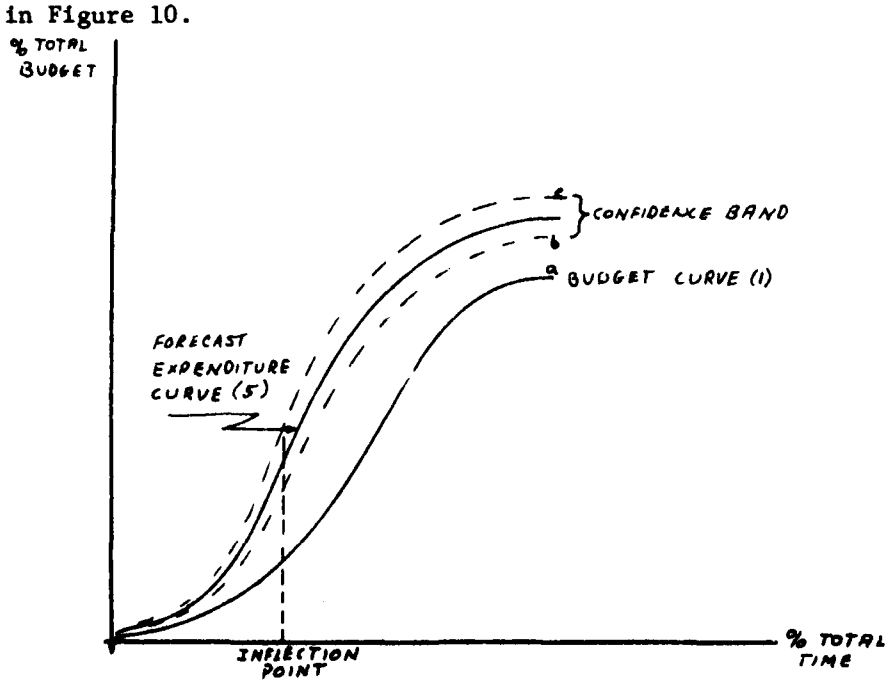

Figure 10. The Error of the Forecast

Here the confidence band indicates the possible range of values (from b to c) in which the true, deflated cost of the program is expected to fall, and similarly, the range of the size [from ab to ac] of the potencial program overrun.

However, if one desires to compare the full cost, with inflation, of the project (5, Figure 9) with the full inflated cost of the budget, both the error of the forecast and the error involved in developing the inflation figures must be considered. This has the effect of greatly increasing the size of the confidence bands as is shown in Figure 11. The end result is that the ability to compare the final cost of the project with the budget cost is greatly impaired. As Figure 11 shows, in this case one could anticipate a tremendous overrun [a - d] and an underrun [c - b] from the same data.

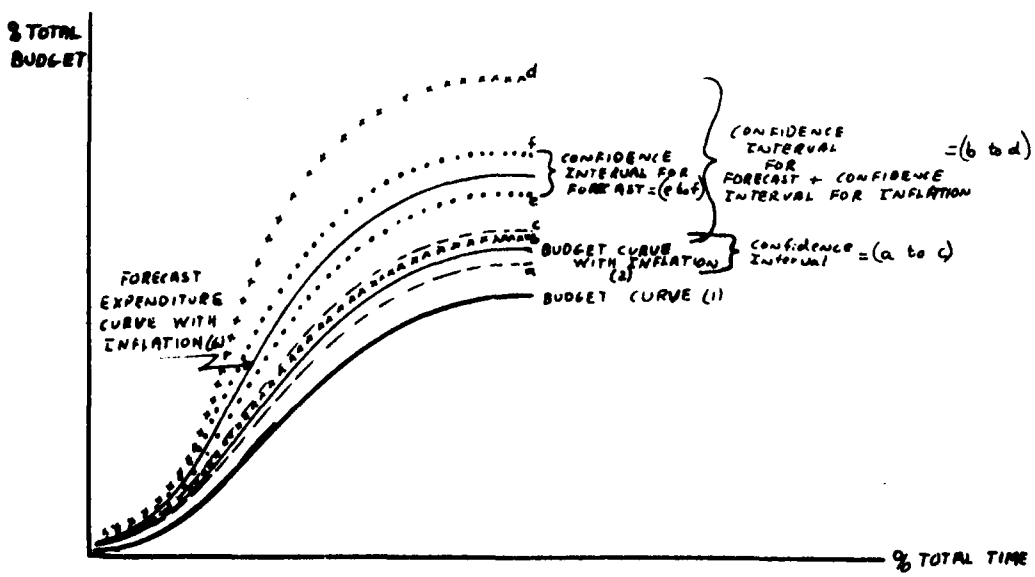

Figure 11. The Error of the Forecast and The Error of the Inflation Forecast

The lesson here is to compare figures in a manner which will minimize the errors involved in the comparison. In other words, the best picture of the status of a project may be gained by comparing the two deflated curves shown in Figure 10. This comparison provides all of the information required for day-to-day management of the program. If a full end cost of the program is desired, this can be developed quickly by simple multiplication utilizing whatever inflation forecast is deemed appropriate at the time that the information is required.

This does not mean, however, that one should not use the actual inflation data when it is available. In this case, no errors of forecast are present because the actuals in both program cost and inflation rates are known. This makes it very easy to remove the effects of inflation to see how much of an overrun is actually attributable to other causes.

Figure 12 shows a case in which the deflated budget curve 1 is modified by the actual experienced inflation to derive curve 2. One may readily compare this curve with the contractor's inflated actuals (curve 3) to determine the actual extent of the overrun [a - b].

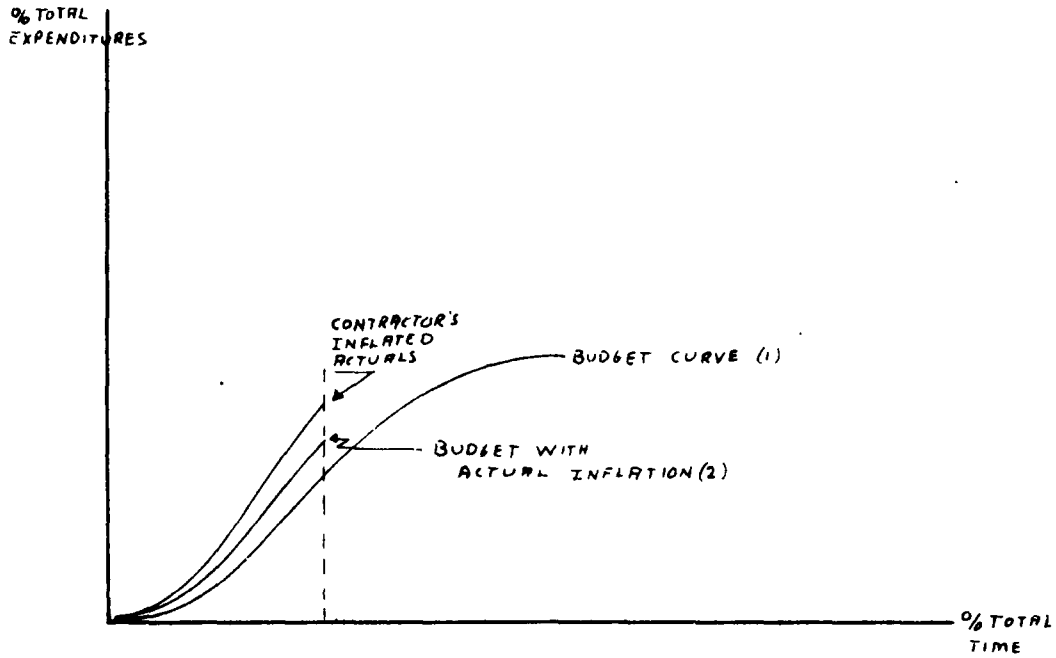

Figure 12. The Use of Actual Inflation Data

(e) Another situation which this method of program monitoring will easily handle is the case of the schedule slippage or program extension. Of the two, the slippage is the most severe because it often occurs early in the project where it has a profound effect on costs. Assume once again the basic deflated budget curve shown in Figure 13 with an actual deflated expenditure curve as shown.

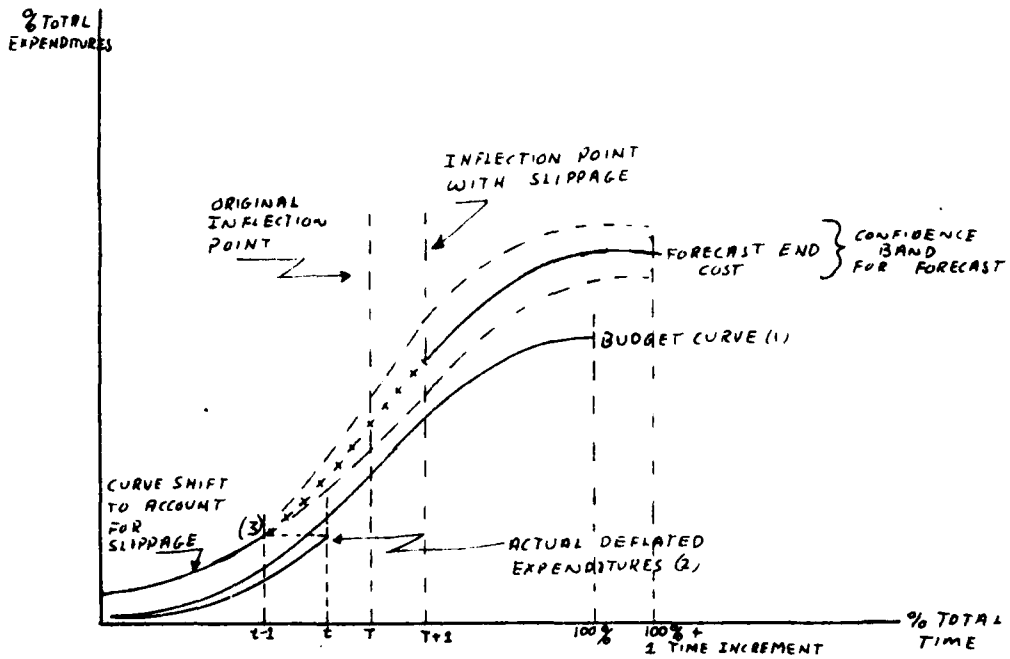

Figure 13. The Program Slippage Situation

It would appear upon initial inspection that the program is running slightly below the planned expenditures at time t. However, it is revealed that the R & D program is actually behind schedule, having only accomplished the number of milestones associated with time t-1. To compensate for this slippage, move curve 2 back one unit from t to t-1 so that the actual expenditures are now shown as curve 3 in their proper relationship with the budget curve. This is actually accomplished mathematically by calculating a new inflection point which will reflect the slippage in the schedule. This new point is derived from the equation for curve 2 by calculating the inflection

point not at time T, the location of the original point, but rather at time T + 1, the location of the inflection point after slippage has occurred. This new inflection point becomes the intercept of the equation for the top half of the budget curve, and the time values which are used to forecast from the top half of the budget curve now start at the T + 1 increment (instead of T) and continue to the 100% + 1 increment (instead of the 100% increment).

THE MOST LIKELY COST

(FORECASTING USING THE GENERAL CURVE)

Forecasting the most likely cost proceeds in the same manner listed in the previous section up to the point at which a new inflection point is forecast. The actuals are deflated, converted to percentages and plotted in the same manner, and the curve form to plot these actuals is the same equation type selected to describe the bottom half of the general curve. At this point, however, the method of forecasting changes considerably.

Instead of merely splicing the top half of the general curve onto the new bottom curve, the bottom curve is actually mapped into the general curve framework. This is accomplished as follows:

(1) Using the deflated actuals from the program, fit whatever curve form is used in the bottom half of the general curve to these data and forecast a new value for cumulative expenditures at the inflection point. This new value is found by substituting the cumulative percent time figure which corresponds to the

general curve inflection point into the new equation which was derived from the actuals.

(2) Take the new value for cumulative expenditures and let this value be equal to the cumulative percent of budget figure which is associated with the inflection point on the general curve.

(3) Using the relationships established in 1 and 2, the top half of the general curve may now be converted from cumulative percentage figures to forecast cumulative expenditures for the program being investigated.

This forecasting method has several advantages:

(1) The time over which the program is planned to run is taken as a given unless evidence to the contrary is discovered.

(2) The lower curve forecast is mapped into the general curve format, thereby creating a smooth S-shaped curve for the entire program. Simply splicing the curves as is done with the budget curve in the previous section will often create discontinuities in the curve.

(3) The forecast which is created in this method is based strictly on the assumption that expenditures in this particular program are proceeding in the same manner that all past programs have proceeded.

THE WORST POSSIBLE COST

Developing the forecast for the worst possible cost is only a matter of slightly modifying the previous most likely cost forecast.

A confidence interval for the most likely cost is calculated by the methods covered in Chapter V. The upper limit of this confidence band, based on whatever level of confidence was selected by the analyst, will give the cost figure that one can be X% certain will not be exceeded. Coupled with the most likely cost, this is an excellent management tool.

In summary, one may develop three possible forecasts from the S-shaped curve. The "spliced" curve forecast using the program budget curve reflects an expenditure of the lowest possible magnitude. For this expenditure to be realized, the program must run exactly as planned from the inflection point onward. This is a highly unlikely situation if any increased expenditures have been incurred early in the program. The most likely cost and its confidence band which extends to the upper confidence limit (or the worst possible cost) for the program are clearly the most realistic forecasts. This is because the method of mapping the new forecast for the bottom of the curve into the general curve format places the entire program in a more legitimate, historical perspective.

FORECASTING IN THE AREA OF THE INFLECTION POINT

With the general methods for handling data early in program established, the next area of interest is the forecast which is made when the string of actual expenditures stretches all the way to the assumed inflection point in the program. When this takes place two courses of action are called for as the actuals approach and cross the inflection point:

(1) Until the inflection point has been reached, the best method of forecasting is to continue to develop a new lower half of the curve from the actuals, and then to map this curve into the general curve as was done in the previous sections of this chapter.

(2) After the actuals appear to have crossed the inflection point (*i.e.*, a large expenditure in a given period has been followed by two periods of decreasing expenditures) one can proceed in two different ways:

(a) If the actual data points continue to fit the top half of the general curve within an appropriate confidence interval, continue to forecast by using the top half of the general curve.

(b) If the actual data points are diverging from the general curve, and if enough actual data points exist beyond the inflection point for regression analysis to be used, then calculate a new equation for the top half of the curve using whichever equation form was appropriate in deriving the general curve. This method presumes that <u>at least</u> 5 data points have occurred past the inflection point. It should be noted that good accuracy in generating a new curve will not occur until at least 10 points have been identified.

CONCLUSION

The general techniques concerning the use of inflation in the model and the method of handling time slippages are applied to the forecasts based on the general curve in exactly the same manner as

they were used with the budget curve. The forecasting method selected should be the one which best serves the decision maker's needs.

Since the general curve would be representative of either a mix of weapon systems or a specific type of weapon system, either of which have been completed at some time in the past, the data around which this curve is structured should be closely examined. It will be readily apparent in most cases that these completed programs all contained a certain number of changes during their R & D phases. Thus, it should be assumed that the final cost figures which are generated from forecasts using the general curve also include a like number of anticipated program changes, even if those changes are not visualized at the early stage of the program when the forecast is made.

CHAPTER V

AN APPLICATION

Project B is a weapon system whose R & D phase started in year 1. The level 6 budget which was provided by the contractor prior to the start of the R & D program contained both a total cost figure and individual figures for the categories of engineering, manufacturing and program management labor. These deflated dollar figures were extracted from the budget and normalized to a percent of total time-percent of total budget basis in Table VII. The inflection point for each of the four budget curves was also identified at this time.

The total expenditure curve was plotted on the axis of Figure 7 and compared for general rationality with the confidence band for all prior R & D programs shown in Figure 3. The proposed budget for project B was contained within the confidence band and the contractor's proposal was therefore judged to be sound.

As a second test, the total expenditure data was regressed against the project B milestones. The results of this test were as follows:

adj. R^2 = .98

F = 165 Durbin Watson Statistic: 1.97

t = 41

These statistics show a significant relationship between the milestones and the expenditure pattern, and, therefore, confirm the specific rationality of this particular expenditure pattern for this particular project.

Having passed these two important tests, the budget data can be assumed to be valid and equations for the top and bottom halves of both the total budget curve and each of the labor curves can be developed. Exhibit III shows the output of the computer regression runs for the total budget curve, and from this output the equations for the bottom and top halves of the curve can be found to be

bottom half: $Y = -.193572 + 1.88402X_1 + .470083X_1^2$

top half: $Y = .149350 + 1.75834X_1 - .91088X_1^2$.

All of the above research would be accomplished prior to the start of the program itself, and the program manager, who has justified in his own mind the validity of the budget in both a general and a specific sense, may now sit back and await the arrival of the actuals.

By the second quarter of year three, the actuals shown in Table VIII have been accumulated. Obviously, the program manager would have been forecasting the future program costs each time a new set of actuals arrived, but let us assume that this task is now to be accomplished based on the latest figures which have just arrived. First, the program manager fits a quadratic curve form to each of the strings of actuals representing the total budget and the engineering, manufacturing and program managing labor. A

quadratic is chosen for the equation form because, in each case, it was this type of equation which provided the best fit for the budget curves. For simplicity, we will only deal with the total budget curve in this example, even though the program manager would be highly interested in forecasting each of the other three curves in an actual situation.

The equation for the total budget actuals is developed by a computer regression technique, the results of which are shown in Exhibit IV. The equation is

$$Y = .00275322 - .057272X_1 + 4.84246X_1^2.$$

Figure 14 shows the actual expenditures and the budget curve for comparison.

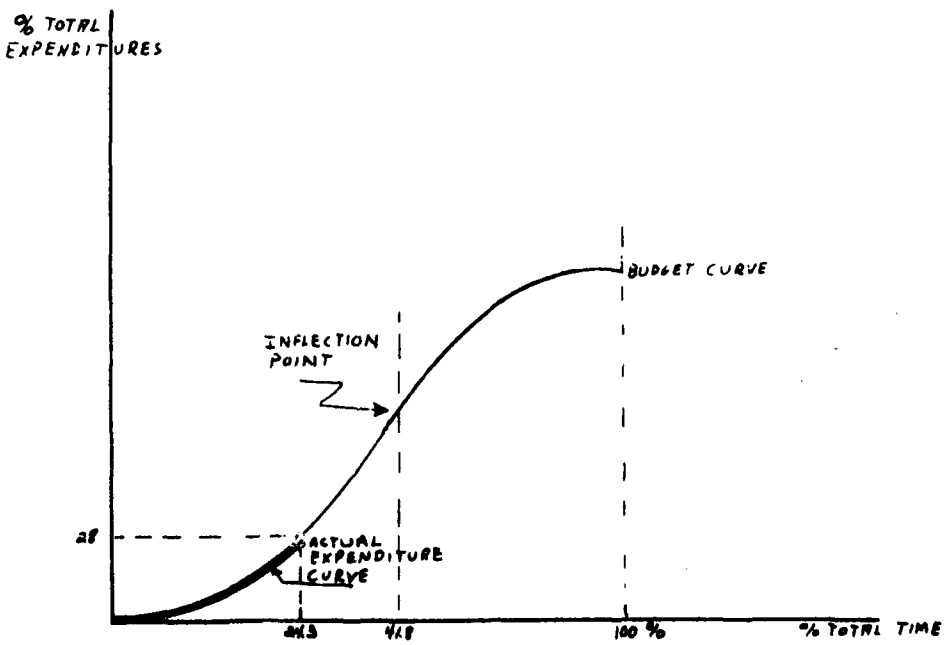

Figure 14. Actual Expenditures and the Proposed Budget

It appears from Figure 14 that the actual total expenditures are slightly underrunning the budgeted total expenditures. However, upon comparison of Tables VII and VIII, it becomes obvious that something is amiss. The engineering, program management, and budget expenditures are all on schedule or overrunning while the manufacturing labor is well below the budgeted expenditure level. The program is obviously not proceeding as planned, and more specifically, the engineering and program managing efforts are not producing the necessary results for manufacturing to take place.

A further comparison of the manufacturing figures in Tables VII and VIII indicates that the program is actually about 1 period behind schedule (2 quarters).[1] This means that a slippage has occurred, and the actuals curve is shifted to reflect this condition as shown in Figure 15.

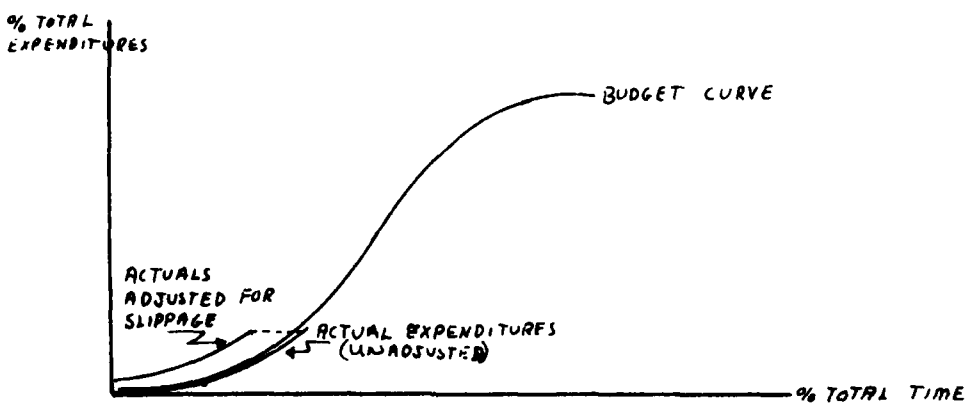

Figure 15. Shifting the Actuals Curve

[1]Note that this condition could have been detected at a much earlier time. Manufacturing begins to lag significantly by the 3rd reporting period.

FORECAST USING THE BUDGET CURVE

To determine the effect which this slippage has had on the total budget, we now forecast the end cost as follows:

(1) Forecast a new inflection point from the lower curve by using the derived equation for the actuals and the time period for quarter 2, year 5 instead of the original inflection point (quarter 4, year 4). Moving to the next data point in this manner mathematically adjusts the forecast for the 1 period slippage noted previously by adding 1 time period to the lower curve.

$\therefore$ new inflection point = Y_n where

$$Y_n = .00275322 - .057272(.476) + 4.84246(.476)^2$$

$$Y_n = 1.0726779$$

(2) Use the new inflection point to calculate a modified intercept for the equation for the top half of the budget data. From Exhibit I this curve equation is

$$Y_t = .149350 + 1.75834X_1 - .91088X_1^2$$

Substituting Y_n for Y_t yields

$$Y_n = 1.0726779 = \text{New intercept} + 1.75834(.476) - .91088(.476)^2$$

and $1.0726779 - 1.75834(.476) + .91088(.476)^2 = \text{New Intercept}$

$$.4420916 = \text{New Intercept}$$

Note that this value is calculated at the same inflection point used in step (1).

(3) This new equation for the upper half of the actuals

$$Y_t = .4420916 + 1.75834X_1 - .91088(X_1)^2$$

may now be used to calculate the forecast cost at any point from the inflection point on to the end of the program. To calculate the end cost of the R & D program we would proceed as follows:

(a) With no slippage the R & D program should end at the 100% of total time point. But with 1 period of slippage the program should now end at the:

(100 + 5.6)% point

where each time increment has been equal to

5.6% of the total time

$$\therefore \text{ end cost} = Y_t = .4420916 + 1.75834(1.056) - .91088(1.056)^2$$

$$= 1.283143$$

or, in other words, a 1 period slippage in the program has induced a 28% overrun in the final cost.

FORECAST USING THE GENERAL CURVE

Assuming once again that 1 period of slippage had been experienced in the program, the general curve may be used to forecast a new end cost as follows:

(1) The new inflection point is forecast in the same manner as before using the equation generated from the actuals in Exhibit IV:

$$Y_n = .00275332 - .057272(.476) + 4.84246(.476)^2$$

$$\text{or } Y_n = 1.0726779$$

(2) This point is now mapped into the framework of the general curve. Figure 3 shows that the inflection point of the general curve is located at 46% cumulative time and 56% cumulative expenditures. Thus, if

$Y_n = 1.0726779$, then

$$1.0726779 = 56\% \ (E)$$

where E = total program expenditures.

$\therefore$ E = (1.0726779) $\div$ .56 = 1.9153

and the most likely total program cost is forecast to be 191% of the budgeted cost.

(3) One can use the method outlined in (2) to fill in the rest of the points along the new forecast cost curve, but it is interesting to note that this forecast calls for a much higher figure than that produced by the budget curve "splicing" method.

THE CONFIDENCE INTERVAL

To use this forecast in its point estimate format would be an error. Rather there is a confidence interval around this forecast which must be considered. The calculation of this interval is a two-part process dealing first with the standard error of the forecast for the new inflection point. Exhibit IV shows that S^2_{yx}, the variance, is .000132813 for this particular line. The standard error of the forecast is calculated from the formula

$$S_f = S_{yx} \sqrt{1 + \frac{1}{n} + \frac{(X_o - \bar{X})^2}{\Sigma (X - \bar{X})^2}}$$

where X_o is, in this case, the inflection point.

$\therefore$ for this example

$$S_f = .011497 \sqrt{1 + \frac{1}{5} + \frac{.476}{.0339}}$$
$$= .044$$

and thus one can be 95% confident that the inflection point is within

an interval of

$$1.0726 \pm t_{n-1} S_f$$

$$\text{or } 1.0726 \pm 2.776\ (.044)$$

$$1.0726 \pm .1246$$

Figure 16 shows this confidence band around the lower half of the curve.

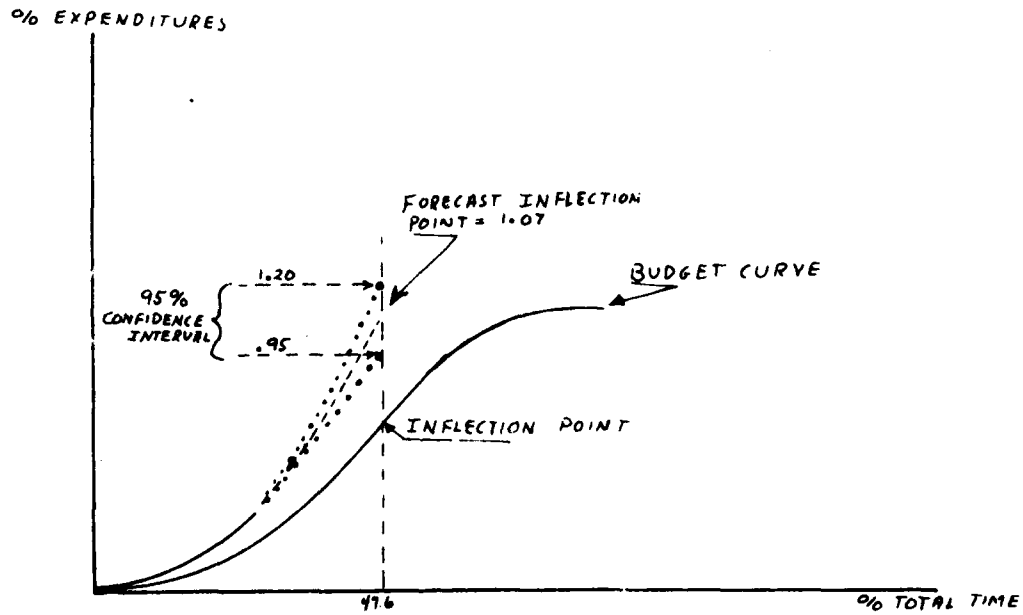

Figure 16. The Confidence Band for the Lower Half of the Actuals Curve

When the top half of the curve is spliced to the forecast inflection point an additional error, namely that of fitting the upper half of the curve to the budget data, is encountered. This fitting error can be read from Exhibit III as S^2_{yx} = .0000271942.

A conservative estimate for the total error involved in the forecast of the final cost can be gained by combining the forecast

error at the inflection point with the fitting error of the upper half of the curve as follows:

$$S_f + S_{yx} = \text{Total standard error} = S_t$$

$$.0440 + .0052 = .0492 = S_t$$

Thus the confidence band around the forecast for the end cost of the program using the budget curve becomes

$$1.283143 \pm t_{n-1}(S_t)$$

$$1.283143 \pm (2.776)(.0492)$$

$$1.28 \pm .1366$$

And, one can be 95% confident that the final cost of program B will be between 114.65 and 141.98 percent of the initially budgeted system cost. Figure 17 shows this completed forecast.

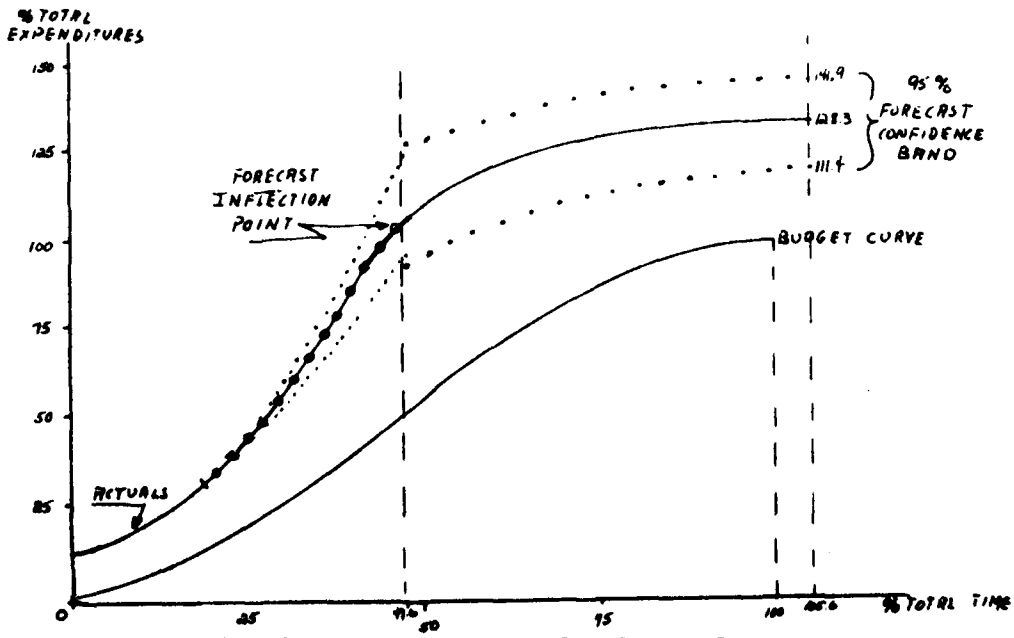

Figure 17. Completed Forecast for System B

Similarly, this process can be completed for the forecast made with the budget curve to derive the most likely cost. Here the standard error of the forecast remains

$$S_f = .044$$

as has previously been calculated. However, as might be expected, considerably more error is involved in fitting the top half of the general curve than was present in the budget curve. For this curve, composed of numerous system types, Exhibit II shows that the fitting error (S^2_{yx}) is

$$S^2_{yx} = .00746826$$

therefore $S_{yx} = .086$

and the total standard error (S_t) is

$$S_t = S_f + S_{yx}$$

$$S_t = .044 + .086 = .13$$

Now the confidence interval for the most likely cost is

$$1.9153 \pm t_{n-1}(S_t)$$

$$1.9153 \pm (2.776)(.13)$$

$$1.9153 \pm .36$$

and there is 95% confidence that the most likely cost of program B is between 228 and 156 percent of the initially budgeted system cost.

Another output of this forecast is the worst possible cost. This cost is the upper limit of the confidence interval or, in this case, 228% of the amount initially budgeted for the project.

If these forecasts seem unreasonably high, it should be noted that program B was an actual program and that the final cost for the

R & D portion of program B was quite close to the worst possible cost (228%) figure. Note also that this entire process has been accomplished with deflated data. Incorporation of the forecast inflation rates for this period would have greatly increased the size of the confidence band in Figure 17 and hence, greatly decreased the possible forecasting accuracy of the overrun.

BIBLIOGRAPHY

Brush, John S. "Study of Possible Improvements in the Accuracy of Aeronautical Economic Escalation Indices," unpublished paper, U.S. Air Force Academy, Colorado, February, 1976.

Drake, Hudson B. "Major DoD Procurements at War with Reality," The Harvard Business Review, January-February, 1970, pp. 119-140.

Goldberger, Arthur S. Econometric Theory. New York: John Wiley and Sons, Inc., 1964, pp. 176-177.

Johnston, J. Econometric Methods. New York: McGraw-Hill, 1963, pp. 134-135, 196.

Johnston, J. Econometric Methods. 2nd ed. New York: McGraw-Hill, 1972, pp. 145-151.

Theil, Henri. Principles of Econometrics. New York: John Wiley and Sons, 1971, pp. 173, 174, 254.

APPENDIX I

LOGISTICS AND GOMPERTZ CURVE EQUATIONS

The equation for the Logistic Curve is

$$\frac{1}{Y_c} = k + ab^x \quad \text{where}$$

$$a = (\Sigma_2 Y - \Sigma_1 Y)\frac{b-1}{(b^n - 1)^2} \quad = \text{the difference between } Y_c \text{ at } X_o \text{ and } Y_c \text{ at } K$$

$$b = \sqrt[n]{\frac{\Sigma_3 Y - \Sigma_2 Y}{\Sigma_2 Y - \Sigma_1 Y}} \quad = \text{the ratio between successive increments of growth.}$$

$$k = \frac{1}{n}\left(\Sigma_1 Y - \frac{b^n - 1}{b-1}\, a\right) \quad = \text{the asymptote or limit}$$

and Σ_1, Σ_2, Σ_3 refer to the first, second and third parts of the Y values.

n is the number of observations in each of the three "parts" or sections of Y values.

The equation for the Gompertz curve is

$Y_c = k_a{}^{bx}$ using the same formulas for k, a and b as shown above.

APPENDIX II

REMOVING AUTOCORRELATION

Given that autocorrelation exists, the observed residual term e_t is actually a combination of two errors

$$e_t = \rho e_{t-1} + \nu_t$$

where ν_t is the true error term

and e_{t-1} is the previous residual term.

Since the true relationship between e_t and e_{t-1}, ρ, is unknown, it must be estimated. Several methods are available for doing this, and this study employs the following approach:

(a) compute all of the n LS residuals

$e_1, \ldots\ldots e_n$

(b) compute the ratio of the mean product of successive residuals to the LS variance estimator

$$\text{or} \quad \beta = \frac{\frac{1}{n-1} \sum_{\alpha=1}^{n-1} e_\alpha e_{\alpha+1}}{\frac{1}{n-k} \sum_{\alpha=1}^{n} e_\alpha^{\,2}} = \text{the estimated relationship}$$

$$\therefore \quad \beta = \frac{\sum_{\alpha=1}^{n-1} e_\alpha e_{\alpha+1}}{(n-1)\, S^2} \qquad (1)$$

Once this estimation has been made, $\hat{\rho}$ may be used with the Generalized Least Squares method to remove the autocorrelation from

the data. This method proceeds as follows:

Given: $Y = X\beta + e$ (2)

where e is the matrix of residuals in which autocorrelation is present.

Thus for any e_t

$$e_t = \rho e_{t-1} + \nu_t$$

substituting $\hat{\rho}$ for ρ, we multiply by a "differencing" matrix

$$D = \begin{bmatrix} \sqrt{1-\hat{\rho}^2} & & & & & \\ -\hat{\rho} & 1 & & & 0 & \\ & \ddots & \ddots & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & \ddots & \ddots \\ 0 & & & & -\hat{\rho} & 1 \end{bmatrix} \tag{3}$$

(2) may now be transformed by (3) to yield

$$(DY) = (DX)\beta + (De) \tag{4}$$

where

$$De = \begin{bmatrix} \sqrt{1-\hat{\rho}^2}\ e_1 \\ -\hat{\rho}e_1 + e_2 \\ -\hat{\rho}e_2 + e_3 \\ \vdots \\ -\hat{\rho}e_{n-1} + e_n \end{bmatrix} = \begin{bmatrix} \sqrt{1-\hat{\rho}^2}\ e_1 \\ \nu_2 \\ \nu_3 \\ \vdots \\ \nu_2 \end{bmatrix} \tag{5}$$

the set of true error terms.

Similarly

$$DY = \begin{bmatrix} \sqrt{1-\hat{\rho}^2}\ y_1 \\ -\hat{\rho}y_1 + y_2 \\ \vdots \\ -\hat{\rho}y_{n-1} + y_n \end{bmatrix} \quad \text{and } DX = \begin{bmatrix} \sqrt{1-\hat{\rho}^2}\ x_1 \\ -\hat{\rho}x_1 + x_2 \\ \vdots \\ -\hat{\rho}x_{n-1} + x_n \end{bmatrix} \tag{6}$$

Recalling that $\beta = (X'X)^{-1}X'Y$, (7)

the OLS solution of (4) is given by

substituting DY for Y and DX for X in (7) to yield

$$\hat{\beta} = [(DX)'(DX)]^{-1}\ (DX)'DY$$

$$= [X'(D'D)X]^{-1}\ X'(D'D)Y \tag{8}$$

From (3),

$$\Omega = D'D = \begin{bmatrix} 1-\hat{\rho} & & & & & 0 \\ -\hat{\rho} & 1+\hat{\rho}^2 & -\hat{\rho} & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \ddots \\ 0 & & & -\hat{\rho} & 1 & -\hat{\rho} \end{bmatrix} \tag{9}$$

and the GLS solution is

$$\hat{\beta} = (X'\Omega X)^{-1}\ X'\Omega Y \quad \text{which,}$$

by (5) eliminates the autocorrelation if $\hat{\rho} \equiv \rho$.

Since $\hat{\rho}$ is an estimate, and not an exact value for ρ, it is unlikely that this process will remove all autocorrelation on the first iteration. For this reason, multiple iterations are generally used with this technique and a new $\hat{\rho}$ is calculated by equation (1) after each iteration. For n iterations

$$\sum_{i=1}^{n} \hat{\rho}_i \simeq \rho.$$

TABLE I

QUADRATIC CURVE FIT--LOWER HALF

NAME	A	B_1	(t)	B_2	(t)	F	D.W.	adjR^2	MSE
A-10 System	-.0786	.699	(2.96)	.997	(3.09)	800.58	1.82	.992	.000059
Engine	-.0102	.0198	(.0717)	1.77	(5.23)	419.36	1.88	.981	.000155
Gun	-.0125	.712	(2.18)	1.17	(2.84)	537.95	1.96	.989	.000065
Milestones	-.0384	1.111	(15.49)	.498	(4.31)	3081.3	.82	.996	.00025
AGM 65 A	.0099	.5365	(1.21)	1.69	(2.31)	80.89	2.76	.975	.0029
AWACS	.0387	.9532	(8.69)	.733	(3.57)	1013.10	2.06	.998	.00015
B-1 Model 1 Cost Estimates	.7528	4.829	(12.38)	-3.498	(-8.19)	392.19	2.33	.989	.00021
B-1 Model 2 Cost Estimates*	218.729	-4842	(-8.03)	31,901	(15.68)	1328.74	1.90	.998	147.123
F 105 Milestones	.0534	.3839	(1.05)	3.526	(3.5)	137.74	2.62	.986	.000668
F 100 Milestones	.0258	.5122	(2.41)	.5411	(.922)	79.2	2.16	.963	.00026
C 141 Milestones	.0496	.2018	(1.55)	.3189	(3.93)	734.57	2.08	.997	.00022
C-5 Cost Initial Report Engineering	-.0252	1.514	(19.67)	5.594	(14.79)	13031.51	1.92	.999	.000011
C-5 Non-Recurring Cost	-.0342	.1035	(9.12)	.2478	(7.94)	3698.91	1.57	.999	.00002
XB-70	-.0306	.3080	(2.05)	1.732	(7.62)	1031.12	1.69	.996	.00021
Tug R	.1417	-1.523	(-3.71)	3.995	(6.46)	108.1	2.03	.997	.00094
Tug C	.0200	.3031	(5.81)	1.582	(17.62)	3763.39	2.64	.999	.000044
Tug E	.0232	.2418	(3.71)	1.964	(16.1)	2714.81	2.12	.999	.000057
Tug G	.0274	.2717	(4.80)	1.963	(18.55)	3769.95	2.49	.999	.000044
C-5 Quality Assurance Hours	.1719	-1.755	(-9.19)	5.04	(16.3)	848.45	2.17	.986	.000024
C-5 Production Hours	.1443	-1.686	(-10.49)	5.215	(20.42)	1576.09	2.07	.993	.000023
B-1 Engineering Labor (Budget)	-.0635	1.39	(5.66)	.6066	(1.22)	550.97	2.02	.995	.00024
B-1 P & M Labor (Budget)	-.0229	1.39	(15.63)	-.128	(-.632)	1631.24	2.13	.998	.000079
B-1 Total Labor (Budget)	-.194	1.88	(4.17)	.47	(.587)	490.26	2.13	.995	.00022
B-1 Manufacturing Labor (Budget)	-.075	.449	(.823)	3.63	(3.35)	215.29	1.78	.986	.00099

* Not Percentage Data

TABLE II

LOG FIT—LOWER HALF

NAME	A	B	(t)	F	D.W.	adjR^2	MSE
A-10 Engine	-.0073	1.80	(27.17)	737.99	1.93	.979	.00015
B-1 Model 4	.5301	2.057	(47.17)	2131.61	1.74	.996	.00029
B-1 Model 2*	11.439	3.22	(77.66)	6031.25	2.09	.999	.00071
XB-70	.8176	1.994	(16.84)	283.53	2.19	.976	.0077

* Not percentage

TABLE III

QUADRATIC CURVE FIT—UPPER HALF

NAME	A	B_1	(t)	B_2	(t)	F	D.W.	adjR^2	MSE
A-10 System	-.88	3.64	(29.97)	-1.76	(-22.79)	4312.53	2.10	.998	.000016
Engine	-1.76	5.86	(4.57)	-3.10	(- 3.95)	32.19	1.66	.795	.00028
Gun	-.66	3.60	(21.57)	-1.95	(-17.57)	1102.23	1.67	.994	.000066
Milestones	-.0881	1.94	(3.56)	- .87	(- 2.5)	78.55	1.92	.881	.00012
AGM 65 A	-.33	2.93	(5.16)	-1.64	(- 4.08)	54.12	2.88	.964	.00086
AWACS	-.5008	3.306	(14.46)	-1.8	(-11.98)	423.37	2.14	.994	.000082
B-1 Model 1 Cost Estimates	.3804	1.19	(18.19)	-.5794	(-14.13)	1335.68	2.21	.997	.0000032
B-1 Model 3 Cost Estimates	-163.377	31.46	(5.84)	-.1836	(- 4.8)	45.34	1.93	.936	1159.55
F-105 Milestones	.2225	1.1736	(11.58)	-.4080	(- 5.69)	1729.57	2.09	.998	.000084
F-100 Milestones	-.7607	3.664	(12.23)	-1.939	(- 9.07)	403.02	1.94	.987	.00060
C-141 Milestones	-1.43	4.825	(7.97)	-2.397	(- 6.29)	257.4	2.51	.987	.00028
C-5 Cost Initial Report Engineering	.4708	1.133	(8.43)	-.6069	(- 6.44)	128.76	2.04	.927	.000021
C-5 Non-Recurring Cost	.0587	1.612	(11.75)	-.6741	(- 7.49)	1011.28	1.77	.994	.000021
XB-70	-1.044	4.365	(11.21)	-2.325	(- 8.85)	450.06	1.63	.992	.00011
Tug R	-2.051	6.747	(18.39)	-3.711	(-14.86)	693.12	1.92	.996	.00034
Tug C	-.6194	3.146	(14.89)	-1.521	(-10.57)	908.54	2.36	.997	.00011
Tug E	-.3129	2.499	(22.59)	-1.185	(-16.69)	3104.38	2.42	.999	.000014
Tug G	-.4153	2.912	(12.04)	-1.501	(- 9.38)	451.56	2.17	.993	.000099
C-5 Quality Assurance Hours	-.3756	2.962	(2.86)	-1.595	(- 2.63)	7.92	1.80	.323	.000019
C-5 Production Hours	.9660	-.0948	(-.248)	.1208	(.492)	1.75	1.89	.057	.000024
B-1 Engineering Labor (Budget)	-.203	2.49	(184.4)	-1.29	(-136.09)	102,647	2.05	.999	.00000088
B-1 P & M Labor (Budget)	-.613	3.15	(19.93)	-1.54	(-15.94)	1384.46	2.00	.997	.0000081
B-1 Total Labor (Budget)	.149	1.76	(13.54)	-.91	(-10.63)	519.66	2.22	.991	.000027
B-1 Manufacturing Labor (Budget)	.523	1.06	(2.89)	-.588	(- 2.46)	13.73	1.99	.739	.00016

TABLE IV

LOG FIT--UPPER HALF

NAME	A	B	(t)	F	D.W.	adjR^2	MSE
AGM 65A	.02213	.342	(6.45)	41.64	2.00	.931	.00062
B-1 Model 4	-.00086	.9432	(27.97)	782.11	2.11	.988	.000058

TABLE V

MILESTONES vs. EXPENDITURES RELATIONSHIP STATISTICS

NAME	A	B	(t)	F	D.W.	adjR^2	MSE
F-102	-1.68	2.69	(11.64)	135.53	1.98	.971	.0011
A-10	.241	.760	(13.85)	191.76	2.13	.864	.00022
T-38	-.795	1.86	(12.0)	144.09	1.75	.973	.0028
F-104	-.0166	.865	(4.65)	21.64	2.39	.838	.0208
KC-135	-.799	1.85	(3.71)	13.79	2.00	.681	.0234

TABLE VI

INFLECTION POINT LOCATIONS

	% Expenditures or % Milestones	% Time
B-1	.5573	.3594
F-105 Milestones	.5814	.3400
C-141 Milestones	.5385	.5050
C-5 Non-Recurring	.5155	.3400
XB-70	.4659	.4431
Tug R	.5316	.5500
Tug C	.5026	.4600
Tug E	.6394	.5050
Tug G	.6626	.5050
B-1 Engineering Labor	.5169	.3590
B-1 P & M Labor	.4681	.3590
B-1 Total Labor	.4218	.3010
B-1 Manufacturing Labor	.5866	.3590
C-5 Quality Assurance Hours	.4363	.4616
C-5 Production Hours	.5246	.4769
A-10 System	.4934	.4741
A-10 Engine	.6899	.6245
A-10 Gun	.5582	.4500
A-10 Milestones	.5388	.4205
AGM-65 A	.6221	.4343
AWACS	.7011	.5050
Mean Inflection Point Locations: [The four B-1 labor categories are not included in these calculations to avoid duplication of data.]	.562	.462

Standard Error:

$$S_{time} = \sqrt{\frac{\Sigma(X-\bar{X})^2}{n-1}} \qquad S_{expenditures} = \sqrt{\frac{\Sigma(Y-\bar{Y})^2}{n-1}}$$

$$S_t = \sqrt{\frac{.0852043}{16}} = \sqrt{.0053252} \qquad S_e = \sqrt{\frac{.0466886}{16}} = \sqrt{.002918}$$

$$S_t = .0729744 = 1\sigma_t \qquad S_e = .0540189 = 1\sigma_e$$

TABLE VII

PROGRAM B DATA

PERCENTAGE OF LEVEL 6 PROPOSED BUDGET FIGURES

Q/YEAR	TOTAL TIME	TOTAL BUDGET	ENG.LABOR	MFG.LABOR	P.M.LABOR	
2/1	.0100	.001206	.001905	.000015	.001322	
4/1	.0680	.025086	.045683	.001122	.058681	
2/2	.1260	.070831	.111638	.019189	.148909	
4/2	.1850	.158547	.199698	.101142	.229059	
2/3	.2430	.284803	.313899	.251838	.315084	
4/3	.3010	.421767	.414406	.407334	.390737	
2/4	.3590	.557690	.516945	.586607	.468142	Inflection Points
4/4	.4180	.668184	.611508	.722934	.533521	
2/5	.4760	.759586	.690111	.83054	.597940	
4/5	.5340	.833228	.759773	.925338	.661209	
2/6	.5920	.874918	.818965	.955615	.722603	
4/6	.6510	.908006	.872539	.965701	.786322	
2/7	.7090	.937340	.914294	.974886	.849762	
4/7	.7670	.960294	.947851	.983690	.902436	
2/8	.8250	.978752	.974348	.992814	.940941	
4/8	.8840	.991546	.990384	.998511	.970686	
2/9	.9420	.993368	.998021	.999866	.991682	
4/9	1.000	.999999	.999999	.999866	.999999	

KEY: Percentages in blocks are inflection points.

TABLE VIII

PROGRAM B

PERCENTAGE OF ACTUAL FIGURES

ENG.	MFG.	P.M.	BUDGET
0	0	0	0
.019	.0006	.036	.024
.074	.0066	.101	.080
.174	.038	.196	.145
.282	.105	.300	.280

```
*************************************** DETROIT REGRESSION ANALYSIS ***************************************

REGRESSION SERIES TITLE:  GENERAL R AND D EXPENDITURE PATTERN CURVE BOTTOM

THE DURBIN-WATSON RHO-TRANSFORMATION FOR AUTOCORRELATION HAS BEEN APPLIED TO THE VARIABLES INCLUDED IN THE PREVIOUS REGRESSION
    ORDER OF AUTO CORRELATION:          1
RESPECTIVE ESTIMATE(S) OF RHO:  0.298249
                    ***** FOLLOWING REGRESSION STATISTICS DERIVED FROM 300 OBS. INDEXED FROM  2 TO 303 *****

   DEPENDENT VARIABLE:  INDEX   LAG   TITLE
                          2      0   PERCENT EXPENDITURE

INDEPENDENT VARIABLES:  INDEX   LAG   TITLE
                          1      0   PERCENT TIME
                          3      0   PERCENT TIME SQUARED

                                INTERCEPT TERM:  ALPHA =   -.126055E-01          (RHOALPHA =   -.923188E-02)
            COEFFICIENT OF MULTIPLE DETERMINATION:  R-SQUARED = 0.768687
ADJUSTED COEFFICIENT OF MULTIPLE DETERMINATION:  ADJ R-SQUARED = 0.767140     DEGREES OF FREEDOM:   < (NUMERATOR), 299 (DENOMINATOR)
                                F-TEST STATISTIC:  F =    496.8104  DEGREES OF FREEDOM:   < (NUMERATOR), 299 (DENOMINATOR)
                           DURBIN-WATSON STATISTIC:  D =    2.0429
                   ESTIMATED RHO (FOR AUTOCORRELATION):  RHO =   -0.034180
                       RESIDUAL MEAN SQUARE:  SIGMA SQUARED =     .993818E-02
```

VARIABLE INDEX	ESTIMATED COEFFICIENT VALUE	STANDARD ERROR	STUDENT T RATIO	COEFFICIENT OF INCREMENTAL CONTRIBUTION	AVERAGE VALUE OF VARIABLE
1	0.537633	0.172393	3.11864	0.177496	0.228383
3	1.39401	0.279053	5.00409	0.277987	.931458E-01
2				DEPENDENT VARIABLE AVERAGE VALUE:	0.283608
				STANDARD ERROR:	0.206546

```
ANALYSIS OF VARIANCE STATISTICS:
F-STATISTIC IS BASED ON THE ADDITIONAL REGRESSION SUM OF SQUARES                          DEGREES OF FREEDOM:
                    ATTRIBUTABLE TO THE VARIABLES INDEXED:          F-STATISTIC:      (NUMERATOR)       (DENOMINATOR)
                                                        3              25.04             1                 299

SIMPLE CORRELATION MATRIX:
          1          3
VAR ********** **********
 1    1.0000     0.9867
 3    0.9867     1.0000

VARIANCE-COVARIANCE MATRIX:
          1          3
VAR ********** **********
 1    .29719E-01  -.4449E-01
 3   -.44492E-01   .7781E-01
```

BEST AVAILABLE COPY

EXHIBIT I

GENERAL CURVE--BOTTOM HALF

```
********************************************** DETROIT REGRESSION ANALYSIS ***********************************************

REGRESSION SERIES TITLE:  GENERAL  B AND D EXPENDITURE PATTERN CURVE TOP

THE DURBIN-WATSON RHO-TRANSFORMATION FOR AUTOCORRELATION HAS BEEN APPLIED TO THE VARIABLES INCLUDED IN THE PREVIOUS REGRESSION
    ORDER OF AUTO CORRELATION:          1
RESPECTIVE ESTIMATE(S) OF RHO:  0.289003

                    ***** FOLLOWING REGRESSION STATISTICS DERIVED FROM 376 OBS. INDEXED FROM  .2 TO 377 *****

   DEPENDENT VARIABLE:  INDEX   LAG   TITLE
                          2      0    PERCENT EXPENDITURE

INDEPENDENT VARIABLES:  INDEX   LAG   TITLE
                          1      0    PERCENT TIME
                          3      0    PERCENT TIME SQUARED

                                    INTERCEPT TERM:  ALPHA =  -0.534448          (RHO*ALPHA =  -0.379511    )
            COEFFICIENT OF MULTIPLE DETERMINATION:  R-SQUARED = 0.815732
ADJUSTED COEFFICIENT OF MULTIPLE DETERMINATION:  ADJ R-SQUARED = 0.814744    DEGREES OF FREEDOM:  2 (NUMERATOR), 373 (DENOMINATOR)
                                     F-TEST STATISTIC:  F =    825.6126  DEGREES OF FREEDOM:  2 (NUMERATOR), 373 (DENOMINATOR)
                                DURBIN-WATSON STATISTIC:  D =    2.0950
                     ESTIMATED RHO (FOR AUTOCORRELATION):  RHO =   -0.055059
                         RESIDUAL MEAN SQUARE:  SIGMA SQUARED =     .746826E-02

VARIABLE INDEX   ESTIMATED COEFFICIENT VALUE   STANDARD ERROR   STUDENT T RATIO   COEFFICIENT OF INCREMENTAL CONTRIBUTION   AVERAGE VALUE OF VARIABLE
      1              .11500                    0.183030          16.9357                0.655313                          0.444224
      3             -1.58347                   0.142985         -11.0772               -0.497525                          0.313124
      2                                                                   DEPENDENT VARIABLE AVERAGE VALUE:                0.508297
                                                                                          STANDARD ERROR:                 0.208781

ANALYSIS OF VARIANCE STATISTICS:
F-STATISTIC IS BASED ON THE ADDITIONAL REGRESSION SUM OF SQUARES                          DEGREES OF FREEDOM:
                    ATTRIBUTABLE TO THE VARIABLES INDEXED:          F-STATISTIC:     (NUMERATOR)      (DENOMINATOR)
                                                        3             122.7             1                373

SIMPLE CORRELATION MATRIX:

VAR ************ ************
            1            3
 1    1.0000       0.9882
 3    0.9882       1.0000

VARIANCE-COVARIANCE MATRIX:

VAR ************ ************
            1            3
 1    .33520E-01   -.29089E-01
 3   -.29089E-01    .24445E-01
```

BEST AVAILABLE COPY

EXHIBIT II

GENERAL CURVE--TOP HALF

```
********************************************** OFFUTT REGRESSION ANALYSIS **********************************************

REGRESSION SERIES TITLE: BUDGETED TOTAL VS TIME FOR THE FIRST HALF

THE DURBIN-WATSON RHO TRANSFORMATION FOR AUTOCORRELATION HAS BEEN APPLIED TO THE VARIABLES INCLUDED IN THE PREVIOUS REGRESSION
      ORDER OF AUTO CORRELATION:          1         2
RHO-ESTIMATE(S) OF RHO:  0.105956  0.305817

------------ FOLLOWING REGRESSION STATISTICS DERIVED FROM  6 OBS. INDEXED FROM  3 TO  8 ------

   DEPENDENT VARIABLE:  INDEX   LAG   TITLE
                          2      0   PERCENT EXPENDITURE

INDEPENDENT VARIABLES:  INDEX   LAG   TITLE
                          1      0   PERCENT TIME
                          3      0   PERCENT TIME SQUARED

                                  INTERCEPT TERM:  ALPHA =  -0.103572          (RHOALPHA =  -0.172076  1
              COEFFICIENT OF MULTIPLE DETERMINATION:  R-SQUARED = 0.996950
ADJUSTED COEFFICIENT OF MULTIPLE DETERMINATION:  ADJ R-SQUARED = 0.994916    DEGREES OF FREEDOM:  2 (NUMERATOR)/  3 (DENOMINATOR)
                                   F-TEST STATISTIC:  F =    490.2611  DEGREES OF FREEDOM:  2 (NUMERATOR)/  3 (DENOMINATOR)
                              DURBIN-WATSON STATISTIC:  D =    2.1344
                     ESTIMATED RHO (FOR AUTOCORRELATION):  RHO =  -0.124536
                          RESIDUAL MEAN SQUARE:  SIGMA SQUARED =    .220907E-03

VARIABLE   ESTIMATED                                                  COEFFICIENT OF                    AVERAGE VALUE
INDEX      COEFFICIENT VALUE        STANDARD ERROR        STUDENT T RATIO    INCREMENTAL CONTRIBUTION      OF VARIABLE
   1          1.88402                 3.452321                 4.18523            0.322347              0.398236
   3          0.470383                3.850475                 0.587255           0.321093              .999109E-01
   2                                                                        DEPENDENT VARIABLE AVERAGE VALUE:  0.332122
                                                                                       STANDARD ERROR:  3.208453

ANALYSIS OF VARIANCE STATISTICS:
F-STATISTIC IS BASED ON THE ADDITIONAL REGRESSION SUM OF SQUARES                    DEGREES OF FREEDOM:
                ATTRIBUTABLE TO THE VARIABLES INDEXED:          F-STATISTIC:    (NUMERATOR)     (DENOMINATOR)
                                                   3              0.3449            1               3

SIMPLE CORRELATION MATRIX:

VAR ************ ************
    1          3
 1  1.0000     0.98944
 3  0.98944    1.0000

VARIANCE-COVARIANCE MATRIX:

VAR ************ ************
    1          3
 1  0.22459    -0.35759
 3  -0.35759   0.44076
```

BEST AVAILABLE COPY

EXHIBIT III

BUDGET CURVE--BOTTOM HALF

```
************************************************ DETROIT REGRESSION ANALYSIS ************************************************

REGRESSION SERIES TITLE: BUDGETED TOTAL VS TIME FOR THE SECOND HALF

THE DURBIN-WATSON RHO TRANSFORMATION FOR AUTOCORRELATION HAS BEEN APPLIED TO THE VARIABLES INCLUDED IN THE PREVIOUS REGRESSION
ORDER OF AUTO CORRELATION:  1
DESCRIPTIVE ESTIMATE(S) OF RHO:  0.219260

*** THE FOLLOWING REGRESSION STATISTICS DERIVED FROM  10 OBS. INDEXED FROM   2 TO  11 ***

DEPENDENT VARIABLE:  INDEX  LAG  TITLE
                       2     0   PERCENT EXPENDITURE

INDEPENDENT VARIABLES:  INDEX  LAG  TITLE
                          1     0   PERCENT TIME
                          3     0   PERCENT TIME SQUARED

                                  INTERCEPT TERM:  ALPHA =   0.149350          (RHOALPHA =   0.116602  )
                COEFFICIENT OF MULTIPLE DETERMINATION:  R-SQUARED = 0.993310
ADJUSTED COEFFICIENT OF MULTIPLE DETERMINATION:  ADJ R-SQUARED = 0.991398    DEGREES OF FREEDOM:  2 (NUMERATOR),   7 (DENOMINATOR)
                                  F-TEST STATISTIC:  F =  519.6626  DEGREES OF FREEDOM:  2 (NUMERATOR),   7 (DENOMINATOR)
                          DURBIN-WATSON STATISTIC:  D =   2.2248
                ESTIMATED RHO (FOR AUTOCORRELATION):  RHO =  -0.247600
                     RESIDUAL MEAN SQUARE:  SIGMA SQUARED =   .271942E-04

VARIABLE INDEX   ESTIMATED COEFFICIENT VALUE   STANDARD ERROR   STUDENT T RATIO   COEFFICIENT OF INCREMENTAL CONTRIBUTION   AVERAGE VALUE OF VARIABLE
      1          -1.75836                      .132873          -13.2390          -0.981438                                 0.508991
      3          -0.610848                     .858805E-01      -10.6312          -0.973401                                 0.465174
      2                                                                            DEPENDENT VARIABLE AVERAGE VALUE:          0.728990
                                                                                   STANDARD ERROR:                            .582277E-01

ANALYSIS OF VARIANCE STATISTICS:
F-STATISTIC IS BASED ON THE ADDITIONAL REGRESSION SUM OF SQUARES
                                                                                   DEGREES OF FREEDOM:
                ATTRIBUTABLE TO THE VARIABLES INDEXED:   F-STATISTIC:   (NUMERATOR)   (DENOMINATOR)
                                                   3       113.0          1              7

SIMPLE CORRELATION MATRIX:

VAR ************ ************
          1          3
 1     1.0000     0.99527
 3     0.99527    1.00000

VARIANCE-COVARIANCE MATRIX:

VAR ************ ************
          1          3
 1    .188957E-01  -.110758E-01
 3   -.110758E-01   .734012E-02
```

BEST AVAILABLE COPY

EXHIBIT III (Cont.)

BUDGET CURVE--TOP HALF

********** DETROIT REGRESSION ANALYSIS **********

REGRESSION SERIES TITLE: ACTUAL TOTAL VS TIME

----- FOLLOWING REGRESSION STATISTICS DERIVED FROM, 5 OBS. INDEXED FROM 1 TO 5 -----

	INDEX	LAG	TITLE
DEPENDENT VARIABLE:	2	0	PERCENT EXPENDITURE
INDEPENDENT VARIABLES:	1	0	PERCENT TIME
	3	0	PERCENT TIME SQUARED

INTERCEPT TERM: ALPHA = .275322E-02

COEFFICIENT OF MULTIPLE DETERMINATION: R-SQUARED = 0.994733

ADJUSTED COEFFICIENT OF MULTIPLE DETERMINATION: ADJ R-SQUARED = 0.964466 DEGREES OF FREEDOM: 2 (NUMERATOR), 2 (DENOMINATOR)

F-TEST STATISTIC: F = 148.8645 DEGREES OF FREEDOM: 2 (NUMERATOR), 2 (DENOMINATOR)

DURBIN-WATSON STATISTIC: D = 3.0117

ESTIMATED RHO (FOR AUTOCORRELATION): RHO = -0.285277

RESIDUAL MEAN SQUARE: SIGMA SQUARED = .132815E-03

VARIABLE INDEX	ESTIMATED COEFFICIENT VALUE	STANDARD ERROR	STUDENT T RATIO	COEFFICIENT OF INCREMENTAL CONTRIBUTION	AVERAGE VALUE OF VARIABLE
1	-.57272E-01	0.233377	-0.240655	-0.157488	0.129400
3	4.34246	0.808870	5.32799	0.966532	.227486E-01
2				DEPENDENT VARIABLE AVERAGE VALUE:	0.105800
				STANDARD ERROR:	0.112286

ANALYSIS OF VARIANCE STATISTICS:

F-STATISTIC IS BASED ON THE ADDITIONAL REGRESSION SUM OF SQUARES

ATTRIBUTABLE TO THE VARIABLES INDEXED:	F-STATISTIC:	DEGREES OF FREEDOM: (NUMERATOR)	(DENOMINATOR)
3	28.39	1	2

SIMPLE CORRELATION MATRIX:

VAR	1	3
1	1.0000	0.96501
3	0.96501	1.0000

VARIANCE-COVARIANCE MATRIX:

VAR	1	3
1	.55824E-01	-0.20907
3	-0.20907	0.82604

BEST AVAILABLE COPY

EXHIBIT IV

ACTUALS--BOTTOM HALF