AN INVESTIGATION OF CHANGES IN DIRECT LABOR REQUIREMENTS RESULTING FROM CHANGES IN AIRFRAME PRODUCTION RATE

by

LARRY LACROSS SMITH

A DISSERTATION

Presented to the Department of Marketing, Transportation
and Business Environment
and the Graduate School of the University of Oregon
in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

June 1976

APPROVED: Roy J. Sampson
Roy J. Sampson

VITA

NAME OF AUTHOR: Larry Lacross Smith

PLACE OF BIRTH: North Bend, Oregon

DATE OF BIRTH: November 23, 1936

UNDERGRADUATE AND GRADUATE SCHOOLS ATTENDED:

Oregon State University
University of Washington
University of Oregon

DEGREES AWARDED:

Bachelor of Science, 1958, Oregon State University
Master of Science, 1963, University of Washington
Master of Business Administration, 1974, University of Oregon

AREAS OF SPECIAL INTEREST:

Logistics Management
Procurement Management
Contract Pricing

PROFESSIONAL EXPERIENCE:

Course Director, School of Systems and Logistics, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio, 1969-1973

AWARDS AND HONORS:

Professional Designation in Contract Management, 1971
Outstanding Educator of America, 1972
Professional Designation in Logistics Management, 1973
Beta Gamma Sigma National Honorary Society, 1974

PUBLICATIONS:

Smith, Larry L., An Overview of Integrated Logistics Support in the Air Force. Dayton, Ohio: Air Force Institute of Technology, 1971.

Noe, Charles G.; Smith, Larry L.; and Jacobs, Grady L. Defense Cost and Price Analysis. 2 vols. Gunter Air Force Station, Alabama: Extension Course Institute, 1973-74.

PREFACE

This research topic is selected from a number of suggestions for business research by United States Air Force staff and operational activities. The topics are promulgated by the Air Force Business Research Management Center located at Wright-Patterson Air Force Base in Ohio. The Center serves as a central coordinating activity in assisting the investigator and the office needing the results. Travel funding, expert advice on the problem and assistance in gaining access to data are provided by the Comptroller Directorate of the Aeronautical Systems Division also located at Wright-Patterson AFB. Access to historical data is provided by the McDonnell Aircraft Company, the General Dynamics Corporation and the Boeing Company.

The author is indebted to all those individuals who aided in completing this dissertation. Special thanks are due Professor Roy J. Sampson for his service as program advisor to the author and research committee chairman, Professor Harold K. Strom for his thoughtful assistance and Professors Kenneth D. Ramsing and Robert E. Smith for their ideas and consideration.

CONTENTS

LIST OF TABLES

LIST OF TABLES-Continued

. . . .

CHAPTER I

INTRODUCTION AND OVERVIEW

Each year great sums are expended to produce military aircraft. For example, the United States budget reflects that an estimated $2.3 billion will be needed to pay for new U.S. Air Force combat aircraft purchased in fiscal year (FY) 1976.[1] The FY 1976 budget request for 108 F-15 air superiority aircraft is $1.4 billion.[2] Similarly, some $361 million is requested for production of 53 A-10 close air support aircraft in FY 1976.[3]

Programs of such magnitude require careful management. One facet of that task is cost estimating. Program managers must be able to predict costs in order to successfully advocate, budget, contract for and control the programs. This research is focused on developing a cost

---

[1]Executive Office of the President of the United States, Office of Management and Budget, Appendix, The Budget of the United States Government, Fiscal Year 1976, (Washington D.C.: U.S. Government Printing Office, 1975), p. 300.

[2]Prepared statement by Thomas C. Reed, Secretary of the Air Force, before the Committee on Armed Services, U.S. House of Representatives, Washington, D.C., January 29, 1976, quoted in Air Force Policy Letter for Commanders 3-1976, supplement (March 1976): 15.

[3]Ibid.

estimating technique to use when program requirements change after production of aircraft is begun.

At the outset of an aircraft production program, a tentative monthly production schedule for the life of the program is negotiated between the contracting parties. This schedule permits planning for such items as work force buildup, facility and tooling needs and the ordering of long lead time items. Although the planning delivery schedule covers the life of the program, formal contractual agreements between the Department of Defense and manufacturers usually cover only annual delivery requirements. Delivery requirements for subsequent years are funded through the exercise of options or separate contracts as funds are appropriated by the Congress.

These multiple year programs may result in a need to change the production rate. For example, when funding for a particular year is insufficient to cover the production scheduled under an existing production plan, it may be necessary to stretch out the production over a longer time span. A national emergency or changed mission requirement may dictate an accelerated rate of production. When such changes in delivery schedules are required, changes in cost estimates are also required to support contract negotiations and additional funding requests. It is suggested that the rate of production is an important independent variable that can be used to help project the change in costs due to either program accelerations or decelerations.

Industrial and governmental cost estimators have traditionally used learning curve techniques to predict the direct labor hours required to produce airframes.[4] These techniques assume that given certain production conditions, the direct labor hours needed to manufacture each airframe decrease in a regular pattern as the cumulative number of airframes produced increases. There are a few approaches to pricing changes in the rate of production, most of which are adaptations of learning curve techniques. But there is no generally accepted estimating technique that considers the rate of production on a systematic basis of prescribed data collection, analysis and prediction.

The purpose of this research is to develop a procedure to consider the effect of a production rate change on direct production labor requirements for additional airframe production. The procedure encompasses the needed

---

[4]The airframe can be viewed as an accounting entity that encompasses the manufacturer's production responsibility. For example, airframe costs would not include the direct labor hours required to produce engines and avionics but would include the hours required to install those components. In contrast, aircraft costs would include all the costs associated with producing the aircraft.

elements of data collection, variable formation, data analysis and forecasting.

In the following chapter, airframe cost elements are discussed and the topic is narrowed to estimating airframe direct production labor hours. A third chapter summarizes previous studies in the area. The fourth chapter outlines the approach for conducting the research and identifies sources of data. An analysis of the data is presented in the fifth chapter. The paper closes with a summary, some conclusions and suggestions for further research.

CHAPTER II

AIRFRAME COSTS

This chapter focuses on three main issues. First the discussion is narrowed to estimating airframe direct manufacturing labor hours. Then a brief overview of learning curve theory is presented. The chapter closes with a discussion of airframe production rate as a cost variable.

DIRECT MANUFACTURING LABOR HOURS

There are many ways to categorize airframe production costs into manageable elements for the purpose of analysis. One generally accepted procedure is to segregate costs into the elements specified on a Department of Defense Contract Pricing Proposal (DD Form 633).[5] These elements include; direct material, material overhead, direct engineering labor, engineering overhead, direct manufacturing labor, manufacturing overhead, general and administrative expenses and profit. One of these cost elements, direct manufacturing labor, is of particular significance since

---

[5]U.S. Department of Defense, Armed Services Procurement Regulation, 1973 edition (Washington D.C.; Government Printing Office, 1973), p. F131.

indirect costs such as manufacturing overhead and general and administrative expenses are often allocated as a proportion of direct cost. Accordingly, an error in estimating the direct manufacturing labor level would lead to a much larger error in the total cost estimate.[6]

Direct manufacturing labor cost can be further described as the product of the average labor rate and the labor hours. At a particular facility the average labor rate will vary with the skill mix, the change in labor force size and the passage of time.[7] In the recent past, the labor rate has steadily increased while the value of the dollar decreased and the level of output per worker increased. The problem of forecasting changes in the labor rate is a significant one but it is outside the scope of this research. This study focuses on predicting direct manufacturing labor hours required to produce an airframe.

The use of direct manufacturing labor hours as a cost element is widely practiced in the airframe industry. In general, this element includes those hours required to machine, process, fabricate and assemble all integral parts

---

[6]U.S. Department of Defense, Armed Services Procurement Regulation Manual for Contract Pricing (Washington D.C.: Government Printing Office, 14 February 1969), p. 6-1.

[7]Charles G. Noe, Larry L. Smith and Grady L. Jacobs, Defense Cost and Price Analysis, Vol. 2, (Gunter Air Force Station, Alabama: Extension Course Institute, 1974), pp. 6-11.

of the airframe structure. It also includes the installation (but not the production) of all aeronautical accessories and equipment. The hours required to produce raw materials and standard shelf items such as fasteners and gaskets are usually excluded.

Direct manufacturing labor hours are those hours of effort which can be readily identified to a given end item or lot of end items through a work order or comparable document. In the airframe industry, direct labor is sometimes further categorized as either fabrication or assembly effort. One reason for this breakout is because the labor required to fabricate parts is usually more machine paced than is the effort required to assemble those parts into major assemblies or an airframe. Thus, the rate of improvement or learning on the building of successive airframes is often viewed as slower for fabrication than for assembly.

Fabrication effort can also be viewed as the sum of set-up and run time. Set-up time per unit is inversely related to the size of a production lot release while run time per unit might be expected to follow a shallow learning curve. Assessing the impact of these subelements is difficult since actual set-up and run times are not often individually recorded.

It appears that different factors affect the behavior of fabrication hours than affect the behavior of assembly hours. Accordingly, it may be advantageous to separate these elements when analyzing historical data.

LEARNING CURVE THEORY

Direct manufacturing labor hours in the airframe industry are often predicted using learning curve techniques. These techniques are quantitative adaptations of the idea that individuals performing repetitive tasks exhibit a rate of improvement due to increased manual dexterity. Observations of complex airframe assembly operations reveal that management innovations such as work simplification, environment improvement and engineering changes also contribute to the rate of improvement. It is subsequently found that this rate of improvement occurs in a regular pattern and can be predicted by a simple model. One model, the unit learning curve model, can be expressed as; $y_x = C_0 \cdot x^{C_1}$ where:

> y_x represents the direct manufacturing labor hours required to make the xth unit,
>
> x is the cumulative unit number,
>
> C_0 is a coefficient that represents the theoretical number of direct manufacturing labor hours required to make the first unit and
>
> C_1 is a coefficient that reflects the rate of improvement that exists in a particular manufacturing environment.

The model is more intuitively appealing when one observes that as the total quantity of units produced doubles, the cost per unit decreases by some constant percentage.[8] The model is also frequently used in a form where the y value is expressed as the cumulative average hours rather than the unit hours.

Certain elements that characterize the airframe manufacturing environment appear to be important contributors to the cost behavior exhibited by the labor requirement for airframes.[9] The first is the building of a sizeable, complex end item which requires large numbers of direct labor hours. Another is production that is manually paced rather than machine paced. A third factor is continuity of production. A production break would permit loss of learning through worker dispersal or forgetfulness. Frequent engineering changes seem to be dynamically accommodated by the learning model as long as the changes are not major.

It should also be noted that while the learning curve is essentially a trend concept, it is not a time series trend form. Rather, the independent variable is taken

---

[8] Ibid., Vol. 1, p. 50.

[9] Herbert R. Kroeker and Robert Peterson, A Handbook of Learning Curve Techniques (Columbus, Ohio: The Ohio State University Research Foundation, 1961), p. 3.

to be the number of opportunities to learn while the dependent variable is cost input per unit of production.[10] Accordingly, 500 airframes produced at the rate of 50 per month would be predicted by the model to require the same number of direct manufacturing labor hours as 500 airframes produced at the rate of ten per month when the same rate of learning (C_1 coefficient) is assumed.

This lack of sensitivity of the learning curve model to the production rate is a problem if the rate is explicitly changed in midprogram. Logically, direct labor requirements per airframe should change as a result of a forced change in the production rate. The worker who senses the pressure of an increased production rate should be motivated to work faster than the worker who senses a production line slowdown. Higher rates of production should permit greater worker specialization than lower rates where one worker would be expected to accomplish multiple tasks. Tool and tooling set-up costs can be spread over a greater number of units at higher production rates. These factors suggest that the production rate should be considered as a variable in the models used to predict the direct labor hours required to manufacture an airframe.

---

[10] Ibid., p. 1.

THE PRODUCTION RATE

An airframe production rate can be defined in different ways for different purposes. Four definitions are discussed here. They are a peak rate, an equivalent rate, a lot average manufacturing rate and a delivery rate.

A peak production rate is the maximum output rate attained during a program. It is usually expressed as units per month. This variable relates to such investments as size of the facilities and the requirements for tools and tooling.

Peak production rate is often used as a variable in a parametric cost model. For example, using multiple regression analysis, Brents relates weight and peak rate to direct labor hours at the 300th unit for 13 jet fighter airframe programs.[11] He uses the resulting model, hours = $B_0 \cdot (\text{DCPR weight})^{B1} \cdot (\text{peak rate})^{B2}$, to estimate direct costs for a new jet fighter production program.[12]

---

[11] Interview with Thomas E. Brents, Jr., Estimating Division, General-Dynamics Corporation, Fort Worth Division, 12 September 1975.

[12] It is the practice of the Government to furnish many components to a contractor building a new aircraft. Engines, electronic systems, wheels and brakes, tires, batteries, certain instruments and auxiliary power units are illustrative of these kinds of components. To provide a common basis for comparing weight based cost estimating relationships, Government and Industry planners have agreed to a Defense Contractor Planning Report (DCPR)

An equivalent rate of production can be used to allow for the fact that work on a specific airframe usually takes place in more than one calendar month. Thus fractional parts of airframes are theoretically produced in different months. Assigning the fraction of each airframe or lot of airframes produced to the proper month and then summing the fractions develops the equivalent airframes produced per month. This method provides a formidable editing problem when manufacturers do not collect direct labor hour data for each airframe or lot of airframes broken out by month. Another problem is that it is difficult to forecast the equivalent rate beyond the historical data.

A lot average manufacturing rate is constructed by dividing the number of airframes in a lot by the time required to produce the lot. The lot release date for the first airframe in a lot and the airframe acceptance date for the last airframe in a lot define the extremes of the lot time span. This form of the production rate is easy to construct and appears to match well with lot average labor hour expenditures. However, the averaging process obscures any learning or rate effect within a lot. Thus when the lots are released infrequently, such as once a

---

weight that excludes the Government furnished equipment, fuels and lubricants. The DCPR weight was formerly called the Airframe Manufacturer Planning Report (AMPR) weight.

year, the number of observations is limited and the data are somewhat distorted.

A delivery rate can be developed by averaging monthly acceptances for each lot. If direct labor requirements are available for each airframe, the actual monthly airframe acceptance rate can also be used to develop cases for analysis. Historical acceptance data are readily available and delivery rates are easily forecasted from contract delivery schedules. Thus the delivery rate is also a possible candidate as a production rate proxy.

In this research both the lot average manufacturing rate and delivery rate are examined as proxies for the production rate. Additional discussions of their construction and characteristics are included in chapters four and five.

The rate of production can be viewed as either a dependent or independent variable. It is a dependent variable when holding the work force constant and allowing the improvement in unit process time due to learning to be realized in terms of an increased output rate.

The production rate is more frequently viewed as an independent variable. The contract delivery schedule states the needs of the buyer in terms of a monthly delivery rate. Then the manufacturer balances the scheduling of facilities, the work force and the production rate to best meet that contract delivery schedule. Typically, such production rates are low at the beginning of the program when the hours required per airframe are high. Then the program rate is accelerated to some peak production rate. Near the end of the program the rate is diminished. For the purposes of this study, the rate of production is viewed as an independent variable to be indirectly specified by the purchaser through the contract delivery schedule.

CHAPTER III

PREVIOUS FINDINGS

A number of people have investigated the effects of production rate on unit production costs. By no means have their findings been unanimous. Some have concluded that the effect of production rate on cost is insignificant or unpredictable. Others have concluded that production rate is an important independent variable that improves the accuracy of production cost models.

This chapter summarizes contributions to the literature on the topic. The summaries are presented in chronological sequence beginning with an older work by the French writer Guibert. Ideas from the writings of Dean, Asher, Alchian and Allen, Johnson, Orsini and Fazio and Russell are extracted. Some observations from an Air Force pricing and negotiation memorandum are included. A review of two recent works closes the chapter. One is by the team of Large, Hoffmayer and Kontrovich of the Rand Corporation and the other is by Joseph Noah writing under contract with the Navy.

P. GUIBERT

Guibert sets forth his ideas on planning for the production of airframes in a book published in 1945. In that study he introduces the production rate as a variable affecting the unit labor cost. Guibert considers the production rate to be an important cost determinant because of its proportional relationship to the number of tools required.[13]

Guibert observes that the number of airframes produced prior to achieving the peak production rate is approximately equal to the number of airframes in work when the peak rate is attained. He then uses this generalization to derive the cost of a hypothetical airframe which is a function of the production rate and the flow time at peak production.

Guibert's models relate unit production cost to cumulative quantity at four different rates of production; 10, 20, 50, and 100 units per month. The models are of the form $y_i = M - (V / (x_i - P))$ where:

y_i represents the man hours required to produce unit i and is expressed as the ratio of the actual

---

[13] P. Guibert, Mathematical Studies of Aircraft Production, translated into English by the U.S. Air Force, from Le Plan de Fabrication Aeronautique (Paris: Dunod) 1945, p. 64.

cost in hours to the cost of the hypothetical unit,

x_i represents the cumulative unit number and

M, V and P are constants derived through simultaneous solution of the model with observed values of x_i and y_i.

At any unit of production x_i, the models predict higher unit costs at higher rates of production when using Guibert's reported constants.

JOEL DEAN

Dean explores the relationship between cost and rate of output in his book on managerial economics. Based largely on previous evaluations of empirical data from a hosiery mill, a leather belt shop and a furniture factory that were generated in the 1930's, he concludes that the most important independent variable in determining short run cost is the rate of output.[14] He further observes that unit costs are generally constant over the range of different production rates included in the samples.[15]

When selecting the form of cost observations, Dean advises other investigators to analyze costs for an

---

[14] Joel Dean, Managerial Economics, (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1951), p. 292.

[15] Ibid., pp. 272-273.

accounting period rather than costs per unit of product.[16] This approach should facilitate editing the input data. However in the case of airframe production, where hours are usually aggregated and reported by production lot, some method of matching the input hours per time period with the output airframes per time period needs to be used. This is a problem because production of a particular airframe often spans more than one month or quarter.

HAROLD ASHER

Asher writes on the relationship between cost and quantity in the airframe industry. In the course of examining empirical data for many different airframe production programs, he subjectively evaluates the effect of production rate on direct labor hours. His conclusion is that production rate is not a very important predictive variable. Asher says that production rate " . . . is felt to be of minor importance, within a certain range of rates of production, and definitely subordinate to the effect of cumulative production."[17]

Asher does not attempt to separate statistically

---

[16] Ibid., p. 286.

[17] Harold Asher, Cost-Quantity Relationships in the Airframe Industry, R-291 (Santa Monica, California: The Rand Corporation, July, 1956), p. 86.

the effect of production rate from the effect of cumulative production on unit cost. He notes however that there are at least two ways in which the rate of production can influence unit labor cost. First, it can affect the number of hours of machine set-up time charged to each unit. Second, it can affect the number of subassemblies employed in the manufacturing process. This in turn affects the number of hours for subassembly work charged to each unit. He concludes that except for these two effects, there is little reason to expect the unit hours for a 200 unit per month case to be significantly fewer than for a 30 unit per month case.[18]

ALCHIAN AND ALLEN

Alchian and Allen advance the idea that production cost is dependent on three production variables. The variables are the total volume of the item to be produced, the production rate and the time span from the decision to produce until the first item is output.[19]

They suggest that larger total volumes lead to smaller unit costs. This cost behavior is explained as

---

[18] Ibid., p. 87.

[19] Armen A. Alchian and William R. Allen, University Economics (Bilmont, California: Wadsworth Publishing Company, Inc., 1964), p. 308.

the result of increasing product standardization with increasing total volume.[20]

These writers further suggest that unit costs should increase with increasing production rates. This behavior is attributed to the fact that higher production rates require use of more overtime and reliance on less efficient workers.[21]

Alchian and Allen view the cost variable as increasing with decreasing initial production span. This behavior is explained as follows. When startup time is compressed, less efficient procedures and equipment are used than if time were allowed to prepare properly for production.[22] These inefficiencies must then be corrected as production progresses resulting in relatively higher unit costs.

The writers do not support these ideas with evaluations of data. Nevertheless, the notions that total volume, production rate and startup span explain production cost behavior may have some application to the airframe industry.

---

[20]Ibid., p. 313.

[21]Ibid., p. 315.

[22]Ibid., p. 322.

GORDON JOHNSON

Johnson describes an approach to estimating direct labor requirements when production rate changes occur on rocket motor production lines.[23] He develops a model that incorporates both the rate effect and the learning effect in estimating direct labor hours. Johnson's approach is to regress direct labor hours per month as a linear function of production rate in equivalent units per month and as a power function of cumulative units produced as of the end of each month. In equation form, the model is

$y = A + Bx_1 + Cx_2^Z$ where:

y = direct labor hours per month,

x_1 = production rate in equivalent units per month,

x_2 = cumulative units produced as of the end of each month and

A, B, C and Z are model parameters.

Johnson tests the model on four sets of rocket motor production data. He reports good results on two of the sets and fair results on one. Johnson explains that the fourth data set may have been distorted due to an inadequate accounting system. He subsequently uses the

---

[23] Gordon J. Johnson, "The Analysis of Direct Labor Costs for Production Program Stretchouts," National Contract Management Journal (Spring 1969): 25-41.

model to price the direct labor element in a rocket motor production program stretchout.

JOSEPH ORSINI

Orsini tests Johnson's rocket motor model on the C-141A airframe production data listed in Table 1 to determine if the model is adaptable to the airframe production industry.[24] In the C-141A program, there is no indication that the production rate is explicitly changed in midprogram although the data reflect that the production rate gradually varies throughout the program.

Orsini reports a method for editing lot airframe direct labor hour data into an equivalent quarterly airframe production rate. He first develops an average production rate for each lot by dividing the number of units in a lot by the lot total hours. Then, from a financial management report prepared by the manufacturer, he extracts the number of labor hours expended per quarter per lot. The product of the average lot production rate in airframes per hour and the hours per lot expended in each quarter produces equivalent units per lot per

---

[24]Joseph A. Orsini, "An Analysis of Theoretical and Empirical Advances in Learning Curve Concepts Since 1966," GSA/AM/72-12, M.S. Thesis, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio, 1970, pp. 53-80.

TABLE 1

C-141A LEARNING AND PRODUCTION RATE DATA IN QUARTERLY STRUCTURE

Quarter	Direct Hours per Quarter	Equivalent Production Rate	Cumulative Units
1	72,088.1	.17033	.17033
2	201,565.4	.47625	.64658
3	406,434.9	1.0192	1.6658
4	501,005.0	1.2479	2.9137
5	644,070.2	1.7871	4.7008
6	944,570.4	3.0159	7.7167
7	1,216,825.8	4.4853	12.202
8	1,254,365.4	5.9089	18.111
9	1,258,344.3	6.6978	24.809
10	1,517,091.4	9.0960	33.906
11	1,784,800.2	12.217	46.123
12	2,265,594.6	18.442	64.565
13	2,397,482.5	21.608	86.173
14	2,453,744.4	25.556	111.73
15	2,364,145.4	27.487	139.22
16	2,137,634.0	26.292	165.51
17	2,280,516.7	28.734	194.24
18	2,241,113.5	28.713	222.95
19	1,768,963.7	23.105	246.05
20	1,444,044.0	18.274	264.32
21	1,095,218.6	13.210	277.53
22	412,769.7	4.9630	282.49
23	71,639.9	.92512	283.42
24	24,199.1	.31937	283.74

SOURCE: Joseph A. Orsini, "An Analysis of Theoretical and Empirical Advances in Learning Curve Concepts Since 1966," GSA/SM/70-12, M.S. Thesis, Air Force Institute of Technology, Wright Patterson Air Force Base, Ohio, 1970, p. 64.

quarter. Summing the equivalent units for all lots in each quarter produces equivalent airframes per quarter. This editing is necessary to fit the data to Johnson's model.

Orsini reports that the three variable model developed by Johnson fits the edited C-141A data better than the Johnson model with the production rate term omitted. He concludes from this comparison that rate is a significant factor in determining manufacturing labor hours.

For further comparison, Orsini also tests the C-141A data in a linear form of the multiplicative model

$y = e^{A} x_1^{B} x_2^{C}$ where: y represents direct labor hours per month,

x_1 represents the production rate in equivalent units per quarter,

x_2 represents cumulative units produced as of the end of each month and

A, B and C are model parameters.

This multiplicative model fits the data better than does the Johnson model. Orsini suggests that " . . . the multiplicative models may be more suitable for analysis than Johnson's model because they eliminate the requirement for estimating the additional parameter Z."[25]

---

[25] Ibid., p. 71.

FAZIO AND RUSSELL

Fazio and Russell overview the entire airframe production cost estimating problem in terms of the sensitivity of each decision variable to the production rate.[26] Their objective is to find the optimum rate of production given the decision variables of production scheduling. Their approach is analytical as contrasted to the statistical investigations of historical data reported by Johnson and Orsini.

Fazio and Russell observe that the number of parallel production lines or stations is an important determinant of direct labor hours per unit. They conclude that the installation of " . . . duplicate load centers, which may be required for higher rates of production, will in fact reduce the overall rate of learning and thereby increase total manufacturing hours for a fixed buy."[27]

They further note that the efficiency of the labor force varies by shift. Accordingly, the direct labor hours required to produce each unit could also be expected to vary with the number of shifts employed.

---

[26] Peter F. Fazio and Stephen H. Russell, "An Analytical Approach To Optimizing Airframe Production Costs as a Function of Production Rate," SLSR 30-74A, M.S. Thesis, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio, 1974, p. 1.

[27] Ibid., p. 66.

THE FISCAL YEAR 74 F-4 AIRFRAME CASE

The rate of production of F-4 airframes at the manufacturer's plant was increased at the request of the Air Force. The Government pricing and negotiation memorandum reflects that the Contracting Officer had some problems in pricing the assembly hour increase associated with the changed production rate. "This element of the proposal was one of the more controversial elements of cost in the FY 74 procurement."[28]

In this instance, the manufacturer was requested to increase the production rate from 15 airframes per month in 1973 to 18 per month in 1974. The rate had already been increased from six airframes per month in 1972 to 15 per month in 1973. The manufacturer indicated that a substantial number of recalls and new hires would be required to meet the increase in delivery rate. Since these personnel would not have worked on the F-4 airframe for some time or not at all, their efficiency would not be as high as the personnel currently building the airframe.[29] Thus, the learning curve model alone would not suffice to

---

[28] An extract from the Pricing and Negotiation Memorandum for the Fiscal Year 1974 F-4 airframe procurement, unpublished document, Aeronautical Systems Division, Wright-Patterson Air Force Base, Ohio, p. 17.

[29] Ibid.

project the cost of the additional airframes.

The manufacturer's approach in supporting his proposal that a higher rate of production requires a higher rate of labor expenditure per aircraft is summarized here. In 1967, there was a production rate increase on the F-4 program from an average of 43 airframes per month to a peak of 67 per month. The manufacturer estimated the direct labor hours that would have been required to produce the additional airframes with the standard learning curve model. This model makes no allowance for a changed rate. These estimated hours were related to the hours actually experienced as the production rate increased. The relationship was expressed in terms of a per cent increase in unit labor requirements due to the rate increase for each lot of airframes.

This per cent increase in delivery rate was then adjusted downward to an "equivalent aircraft delivery rate" to compensate " . . . for the learning that takes place each month during the period of increased deliveries."[30] To achieve this, an equivalent airframe flow time reduction curve based upon a 75 per cent learning factor was developed for the anticipated delivery schedule. This adjustment reflects that additional

---

[30] Ibid., p. 18.

learning takes place each month that a given rate of delivery remains above the base rate.

The equivalent per cent increase in rate was then plotted against per cent increase in assembly hours. From this graph, the manufacturer estimated the per cent increase in assembly hours from the per cent increase in equivalent delivery rate. This approach was used to support the assembly labor hour portion of the cost proposal.

In a note to the Air Force analysis of the proposal, the analyst wrote that the procedure summarized above predicted a higher impact due to changed production rate than was actually experienced in the 1973 procurement at a rate of 15 per month. With this new information, the predicted increase in labor requirements due to the change in production rate was adjusted downward by some 75 per cent.

In searching for reasons why the cost estimating relationship developed by the manufacturer predicted incorrectly, it is noted that it was based on only one set of previous observations. In addition, the nominal values of production rate used in constructing the relationship were in the range of 40 to 70 airframes per month while the application of the relationship was to the range of six to 18 airframes per month. Even though percentage change does not consider order of magnitude, it is possible the prediction was produced outside the relevant range of the relationship.

LARGE, HOFFMAYER AND KONTROVICH

Three investigators from the Rand Corporation report on a search for parametric cost models that would enable program planners to project the magnitude of costs that could be expected at different production rates.[31] They use statistical techniques to examine the effects of airframe production rate and other selected design parameters on major cost elements. The investigation is an aggregative approach where a few descriptors from each of many programs are evaluated simultaneously through multiple regression analysis. The purpose is to develop a general cost model suitable for predicting costs of other programs.

Observations from 29 different programs are examined. The model selected for direct manufacturing labor hours is $y_i = A \cdot w^B \cdot s^C \cdot r^D$ where:

---

[31]Joseph P. Large, Karl Hoffmayer and Frank Kontrovich, Production Rate and Production Cost, R-1609-PA&E, (Santa Monica, California: The Rand Corporation, December 1974), p. 1.

y_i represents the cumulative direct manufacturing labor hours through unit number i,

w represents the program average DCPR weight expressed in pounds,

s represents the maximum design airspeed in knots,

r represents the production rate expressed as the acceptance span in months for the first i airframes. For the investigation, i is arbitrarily chosen at 100 or 200 airframes.

A, B, C and D are model parameters.

The data are evaluated in the model at y_{100}, y_{200} and $y_{200} - y_{100}$. Then to determine the importance of the acceptance span as a proxy for the production rate, the data are revaluated in the model without the acceptance span term. This comparison shows that the acceptance span proxy for the production rate is of little value in explaining the variation in cumulative labor hours among the different programs. The investigators conclude that the influence of production rate cannot be predicted with confidence on the basis of the analysis performed.[32]

The investigators observe that the programs included in the sample are typical of modern production runs in that the number of airframes produced is low (less than 300). The production rates are also low at less than 13 airframes per month. Thus their findings are limited to the effects of production rate for programs where total

---

[32]Large, Hoffmayer and Kontrovich, pp. 50-51.

output amounts to a few hundred airframes or less.

The use of an acceptance span in months as a proxy for the production rate is explained to be necessary because within each program the production rate changes a number of times. Thus in order to have a single observation to represent each program production rate, an averaging approach is needed. This use of an average for the production rate will almost surely mask the effects of a rate change after a particular program is in progress.

Large, Hoffmayer and Kontrovich also evaluate the cost model proposed by Johnson for rocket motor direct cost estimating and later tested by Orsini with airframe production data. Using data from seven different airframe production programs, they found that "In none of the programs did the inclusion of production rate improve the coefficient of determination, R^2, by as much as one per cent over what was obtained using cumulative quantity alone."[33] In one of the four cases where the contribution of the rate was found to be statistically significant, costs were found to increase as rate increased. Thus, the Johnson model does not appear to be appropriate for the task at hand.

---

[33]Ibid., p. 49.

JOSEPH NOAH

Noah reports on a statistical analysis of cost data to discover the effect of production rate on airframe cost.[34] His research includes the major cost elements for two fighter airframe production programs, the A-7 and the F-4. The programs are characterized by the large numbers of airframes produced. The A-7 data cover 1252 airframes which were produced over a period of nine years while the F-4 data span 4645 airframes and eight years.

Noah analyzes all the major elements of airframe cost but his findings on direct labor hours are of particular interest here. For the two programs studied, he reports that the following relationship models the data well:

$y = e^{A} \cdot x_1^{B} \cdot x_2^{C} \cdot x_3^{D}$ where:

> y represents the average direct labor hours expended per pound of airframe produced for each airframe lot,
>
> e is the base of the natural system of logarithms,
>
> x_1 represents the cumulative volume expressed as pounds of airframe produced through the midpoint of each successive airframe lot.

---

[34] J.W. Noah, "Resource Input vs. Output Rate and Volume in the Airframe Industry," Draft Technical Report TR-204-USN, Contract N00014-73-0319, (Alexandria, Virginia: J. Watson Noah Associates, Inc., December 1974), p. 1.

x_2 represents the production rate expressed as the average pounds of airframe delivered per month for the period spanning the first and last delivery of the lot,

x_3 represents the annual volume of aircraft ordered, expressed in airframe pounds and

A, B, C and D are model parameters.

The data adjustment from airframes to pounds is accomplished with the observed data element, DCPR weight per airframe, averaged over each production lot. Using this model, Noah reports multiple coefficients of determination for the F-4 and A-7 data of 0.99 and 0.80 respectively.

Noah attempts to generalize the cost model by averaging the estimated regression coefficients derived from the two sets of cost data. He suggests that this generalized model can be used to predict the effects on cost of changing the production rate of the F-14 airframe production program. Noah concludes that "Our findings suggest that delivery rate as a proxy for production rate has a significant effect, and an important one, on the production cost of airframes."[35]

The conversion of airframe numbers to DCPR pounds is a common practice when estimating airframe costs. This practice adjusts the data for variations resulting from

---

[35]Ibid., p. 41.

minor design changes in the airframe as production proceeds. A change that adds weight to an airframe is assumed to add labor hours to the cost of manufacturing that airframe. One problem with this conversion is that the direct labor hours required to install Government furnished equipment on the airframe are usually included in the recorded hours required to assemble the airframe. As previously noted, the weight of those Government furnished components is excluded from the DCPR weight. This presents a possible data mismatch.

The first independent variable, cumulative volume, is analogous to the cumulative airframe variable used in learning curve analysis. The choice of the lot midpoint of cumulative volume in pounds relates logically to the choice of average hours per pound per lot for the dependent variable.

The lot average airframe delivery rate is a practical choice to represent the production rate. The data are easily obtained and manipulated to form the variable. However, it appears that the average delivery rate variable will lag the average expenditure of the hours required to produce the delivered airframes.

There is no clear indication of the adequacy of the Noah approach. The two programs examined are very large. Noah's approach of averaging coefficients derived from the two programs to construct a generalized cost model for a third program defies logic. Nevertheless, the best test of any predictive technique is how well it predicts. Until this model is tested on additional programs no firm conclusions can be drawn.

SUMMARY

All of the literature reviewed reflects a common interest in the relationship between production rate and direct labor requirements. But the authors do not agree on the form or the importance of that relationship. Some write of increasing unit cost with increasing rate while others express the opposite viewpoint. Some write of no significant effect of rate on unit cost while others suggest that rate is an important independent variable. This research will test the idea that the production rate changes can explain changes in direct labor requirements. Furthermore, it is postulated that an increase in the rate will cause a decrease in the unit labor requirement within the relevant capacity of a given plant. The next chapter sets forth some theory and an approach for testing these ideas.

CHAPTER IV

RESEARCH OUTLINE

This chapter outlines the procedures used to conduct the research. For ease of presentation, the chapter is divided into five sections. First the general purpose and approach are described. Then the variables and some theory on their relationships are discussed. A cost model is suggested to explain the variation in the direct labor hours as production proceeds and the rate changes. A fourth section sets forth hypotheses for testing the adequacy of the model. Included under each hypothesis are tests and conditions for acceptance. The data used in the research are discussed in the final section.

PURPOSE AND APPROACH

The effect of a change in production rate on the direct labor hours required to manufacture airframes is the focus of this research. Should the findings indicate that production rate provides an explaination of variation in unit labor requirements, the purpose is to develop procedures for collecting, editing and analyzing historical airframe production data. These procedures are designed to assist analysts in predicting the effects of production rate changes on the direct labor requirement for airframe manufacture.

The approach is to evaluate the effects of a production rate on direct labor requirements within an individual program. This contrasts with developing a general cost model to apply to all airframe production programs. The individual program approach necessitates gathering and evaluating many observations for each program investigated. The desired product from each set of data is a tailored cost model suitable for predicting the effects of a production rate change on the direct labor hours required to manufacture additional airframes in that program. Each data set should be considered as the basis for a separate investigation. The results of the investigations are not to be aggregated into a generalized cost model.

THE VARIABLES

The dependent variable to be investigated is the average number of direct labor hours required to manufacture each pound of airframe. When expressed as a total, this variable includes all the hours required by the manufacturer and major subcontractors to fabricate parts, assemble those parts into components and the airframe and finish the aircraft. The hours required to produce raw materials, bench stock such as rivets and standard fasteners and aeronautical accessories and equipment such as avionics and engines are not included.

It has been suggested that direct labor hours for fabrication and assembly operations behave differently with respect to cumulative production.[36] Accordingly, when the data permit, the dependent variable is also examined in parts as well as in total. For example, in two programs, the dependent variable is analyzed as fabrication, assembly and total hours required to manufacture an airframe.

Results may also be improved if the data are stratified according to different airframe models that are produced simultaneously. When the data permit, this proposition is also tested.

The procedures used by manufacturers to collect the hours required to produce an airframe vary in accuracy for different parts of the process. For example, the direct labor hours required to fabricate bits and pieces in the plant are usually prorated to each airframe according to

---

[36]Interview with Charles Meranda, Cost Analyst, Comptroller Directorate, Aeronautical Systems Division, Wright-Patterson AFB, Ohio, 9 September 1975.

the size of the batch released to the machine shop. On the other hand, the hours required to assemble those parts into an airframe can often be accurately traced to a specific airframe. The direct labor hours required to manufacture purchased parts such as standard fasteners are often ignored. The hours required for subcontractors to manufacture major parts and assemblies to the manufacturer's specifications are included but frequently they must be estimated from the results of subcontract negotiations. Accordingly, when working with the different levels of data aggregation, one must remember that different levels of accuracy accompany the data from the different parts of the process.

Many airframe manufacturers collect data by lot or block of airframes. This generalization is particularly valid for the fabrication of parts in a batch operation. When the data are in lot form it is convenient to express the dependent variable as the lot average direct labor hours required to fabricate, assemble or output each airframe.

When the data are available by individual airframe, it is convenient to aggregate the data by sublots where each sublot includes the hours required to produce the airframes output in a specific time frame such as a month. This action reduces the number of observations to a manageable number and permits a logical matching of the hours expended to a delivery schedule.

The design and weight of an airframe often change over the life of a program. In general, the direct labor hours required to manufacture an airframe will increase with an increase in weight. Thus, variation in direct labor hours required to produce an airframe can theoretically be reduced by expressing those hours on the basis of pounds produced. The average DCPR weight of each lot of airframes is usually available from airframe manufacturers. In this report, the dependent variable is expressed as the unit average direct labor hours per DCPR pound of airframe.

The independent variable of primary interest is the production rate. One can define different forms of a production rate that might logically be used to explain variation in the labor hours required to produce an airframe. For this research it is important to restrict these definitions to the confines of the available data. Since models developed from the research procedure are intended to be used for prediction, one must also be able to forecast the production rate variable in terms of future output. Finally, the production rate variable must logically relate to the dependent variable. With these restrictions in mind, two definitions of production rate are selected for testing. They are the lot average manufacturing rate and the lot average delivery rate.

The lot average manufacturing rate is the number of airframes in a production lot divided by the production time span. This span bridges the lot release date of the first airframe in a lot and the completion date of the last airframe in a lot. The lot release date is defined as the date work orders are issued to fabricate the first batch of parts in a lot. The completion date is defined as the date the customer signs for acceptance of the completed aircraft. Construction of the lot average manufacturing rate variable requires the lot size, the lot release date and the acceptance date as raw data.

The lot average delivery rate is the number of airframes in a lot divided by the time span over which those airframes are delivered. The time span is bounded by the dates the first and last aircraft in the lot are accepted by the customer. Construction of the delivery rate variable requires only the lot size and the acceptance dates as raw data. When direct labor hour data are aggregated for each airframe accepted in a month, the actual airframe acceptance rate may be used for this variable.

It is clear that other independent variables may need to be examined simultaneously with the production rate in order to statistically isolate the effect of rate on cost. The cumulative number of airframes produced is one such variable. One method used to explain much variation in the direct labor hours required to produce an airframe is through application of a unit cumulative learning model. One model is $y_i = B_0 \cdot x_{1i}^{B_1} \cdot 10^{e_i}$ where:

> y_i represents the unit average direct labor hours required to output each pound of airframe in lot i,
>
> x_{1i} represents the cumulative learning achieved on all airframes of the same type through lot i,
>
> e_i is an error term that accounts for the variation in each lot i that is not explained by the independent variable and
>
> B_0 and B_1 are model parameters.

When the dependent variable is expressed as lot average direct labor hours per pound of airframe, cumulative production can logically be expressed as one-half the corresponding lot size plus the cumulative total number of airframes produced in prior lots. This lot midpoint is selected to represent the entire lot since it matches conceptually with the use of an average to represent the direct labor hours.

THE CUMULATIVE PRODUCTION AND PRODUCTION RATE COST MODEL

This study is based on the assumption that the production rate affects the quantity of direct labor hours required to manufacture an airframe. Since cumulative production has been shown to be a strong explainer of changes in the direct hours required to produce an airframe, addition of a production rate variable to the learning curve model should explain additional variation in the direct labor requirement. Accordingly, a three variable model is suggested as more suitable for explaining and predicting variation in direct labor requirements than is the cumulative learning model. The model suggested is $y_i = B_0 \cdot x_{1i}^{B_1} \cdot x_{2i}^{B_2} \cdot 10^{e_i}$ where:

y_i represents the unit average direct labor hours required to output each pound of airframe in lot i,

x_{1i} represents the cumulative learning achieved on all airframes of the same type through lot i,

x_{2i} represents the lot i production rate for all airframes of the same type,

e_i represents the variation in each dependent variable that is not explained by the two independent variables and

B_0, B_1 and B_2 are parameters in the model.

The production rate is chosen for inclusion in the model in the multiplicative form for a number of reasons. Other writers have suggested that it might be a good predictor in this application. Multiple regression analysis is facilitated by this choice. Finally, investigation of some test data indicates that it works well.

An increase in the production rate within the planned plant capacity should cause a decrease in the hours required to manufacture each airframe. A rate decrease should produce the opposite reaction. This theory is supported by the ideas of specializing labor, prorating set up time and motivating workers to produce.

The management concept of specializing labor to increase efficiency supports the idea that the rate and the direct production labor hours should be inversely related. At higher rates of production, more workers are added. This permits the foreman to assign each worker fewer tasks which are performed more frequently. Thus, in addition to the increased proficiency which accrues due to the learning process, some of the time that would be required to change from one task to another is saved. At decreased rates of production, the opposite cost behavior is expected as workers are removed from the production line and the remaining workers are required to perform additional tasks.

At higher rates of production, the hours required to set up machines for fabricating airframe parts can be spread over more airframes because of the accompanying larger batch sizes. This should also contribute to an inverse relationship between production rate and labor requirements.

Finally, it appears logical that the worker who senses the pressure of a high production rate will be motivated to work faster than the worker who is not pressed to finish a task because the level of effort is diminishing. This assumption is supported by the "toe-up" phenomenon familiar to students of the airframe production process. The "toe-up" phenomenon is an increase in the direct hours required to manufacture airframes that is frequently experienced as a production program is concluded and the production rate is decreased toward zero. The C-141A data in Table 1 display this phenomenon.

As previously implied, least squares multiple regression analysis is used to examine historical airframe production data in the cumulative production and production rate cost model. To facilitate the regression analysis, the model is transformed by extracting the common logarithm of each term. The transformed model is $\log y_i = \log B_0 + B_1 \log x_{1i} + B_2 \log x_{2i} + e_i$. The model is now linear in each term.

Some assumptions about the error terms in the model are required to develop tests for the regression results. For this analysis, the error terms are assumed to be normally distributed, not related to the independent variables or each other, with a mean of zero and a constant variance equal to that of the dependent variable.

Collinearity of the independent variables, cumulative production and production rate, is likely to exist. It is obvious that as the production rate increases, the cumulative number of airframes produced also increases. But as production rate decreases, cumulative production continues to increase as long as some airframes are being completed. When the independent variables in a multiple regression are highly correlated with each other, the standard error of individual regression coefficients may become unreliably large.[37] However, it may not alter the predictive power of the total regression model. Since the purpose of developing the model is for prediction, some collinearity is not considered a problem as long as estimates of individual coefficients do not fail statistical tests yet to be described.

However, collinearity may produce some peculiar effects in regression analysis besides indicating unreliability of individual regression coefficients.[38] For example, the two variables cumulative production and production rate, when expressed as logarithms, may be positively correlated with each other and negatively correlated

---

[37] William A. Spurr and Charles P. Bonini, Statistical Analysis for Business Decisions (Homewood, Illinois: Richard D. Irwin, Inc., 1967), p. 610.

[38] Ibid., p. 611.

with the logarithm of hours per pound of airframe. But the net effect of the production rate, when taking cumulative production into account, may be positive. Based on the earlier discussion of learning curve and production rate theory, the expected signs of the estimated regression coefficients associated with the cumulative production and production rate variables are negative. Yet the above discussion points out that if sufficient collinearity exists, the sign of the coefficient associated with the weaker independent variable may change to positive while the model still supports the theory.

HYPOTHESES

There are three hypotheses to be tested in this research effort. The first is that the production rate, when expressed in a suitable form in a proper model, can explain an important part of the variation in the direct labor requirements to build an airframe. The second is that the cumulative production and production rate cost model can be employed to explain the variation in the hours required to fabricate and assemble the airframe. The third is that the cost model is suitable to predict the hours required to produce additional airframes.

HYPOTHESIS ONE

The first hypothesis is that the production rate is an important explainer of variation in total direct labor requirements when included in an appropriate model. It can be tested by examining the results of regression analysis of historical data in the selected model. The model to be used for the tests is the cumulative production and production rate cost model.

For this first hypothesis, the dependent variable is the logarithm of total hours per pound rather than fabrication or assembly hours. The hours are stratified by homogeneous models of airframes when practical. The independent variables are, in turn, the logarithm of the cumulative production lot midpoint and the logarithm of the production rate. Both definitions of the production rate are examined when data permit.

Three subhypotheses are used to examine the model in parts. The first is that cumulative production and production rate are related to hours per pound as indicated in the model. This hypothesis is stated more formally in the null and alternate form as :

(1A) H_0: B_1 and $B_2 = 0$

H_a: not both B_1 and $B_2 = 0$

The null hypothesis is rejected and the alternate hypothesis accepted if it is shown that the statistic F^* is larger than the theoretical value of F at the 0.05 level of significance. In this instance, $F^* = (SSE(0) - SSE(x_1,x_2))/2 \,/\, SSE(x_1,x_2)/(n-3)$. SSE represents the sum of the square of each residual term. The residual term is the difference between the observed value of the hours per pound and the fitted regression line at the corresponding values of cumulative production and production rate. Then SSE(0) is SSE about the mean value of the dependent variable. $SSE(x_1,x_2)$ is SSE with both the independent variables in the model. The number of observations is represented by n.[39] Throughout this discussion, the theoretical percentiles of the F distribution are obtained from a table in the Neter and Wasserman text.[40]

It is of primary interest to test the hypothesis that the production rate can explain additional variation in the hours per pound of airframe in the presence of the cumulative production variable. This second subhypothesis is the equivalent of stating that the B_2 coefficient in the cumulative production and production rate cost model

---

[39] John Neter and William Wasserman, Applied Linear Statistical Models (Homewood, Illinois: Richard D. Irwin, Inc., 1974), p. 242.

[40] Ibid., pp. 807-813.

has a non zero value at some predetermined confidence level. In the null and alternate form, the hypothesis is:

(1B) H_0: $B_2 = 0$

H_a: $B_2 \neq 0$

As before, the alternate hypothesis is accepted if the test statistic F^* is larger than the theoretical value of F at the 0.05 level of significance.

In this case the F^* statistic is calculated by measuring the reduction in the sum of the square of each residual that is attributable to adding the production rate variable to a reduced model containing only the cumulative production independent variable. This statistic is calculated from the relationship:

$F^* = (SSE(x_1) - SSE(x_1,x_2)) \ / \ SSE(x_1,x_2)/(n-3).$

$SSE(x_1)$ represents the SSE for the reduced model, without the production rate variable.[41]

The aptness of the model is evaluated through an analysis of the residuals or the observed errors. If the model is appropriate for the data, the residuals should reflect the properties assumed for the theoretical error values. Accordingly, the residuals are examined for departures from the assumptions of constant variance, independence and normality. A third subhypothesis

---

[41] Ibid., pp. 262-264.

is advanced to test for aptness of the model:

(1C) H_0: The model is not appropriate.

H_a: The model cannot be rejected.

The assumption of a constant error variance is evaluated by plotting the residual values against the fitted values of the dependent variable. Lack of a discernable pattern and a scattered distribution of points is accepted as evidence that the assumption of constant variance is not violated.

The assumption of error term independence is checked by plotting the residual values against each of the independent variables. If the residuals fluctuate in a more or less random pattern around a base line of zero, independence is assumed to exist.

The residual values may also be plotted in the sequence in which they were gathered against a time variable. This check for autocorrelation may be somewhat distorted in regression analysis of unadjusted labor requirements for airframe manufacture. Peaks in the unit data caused by major engineering or model changes show up as runs of positive and negative residuals. Since these are symptoms similar to that of autocorrelation, one may be mislead into rejecting a model when it does not exist. Therefore, in testing for autocorrelation, care is taken to account for alternative sources of runs in the residuals.

The assumption that the error terms are normally distributed is studied by examining the cumulative frequency distribution of the residuals when plotted on normal probability paper. A characteristic of this type of paper is that a normal distribution plots in a straight line. Therefore, the assumption of normality is rejected if a plot of the cumulative frequency distribution of the residuals on normal probability paper produces substantial departures from a straight line.

The null hypothesis is rejected and the alternate hypothesis accepted if the subjective tests described above do not indicate substantial departures from the assumptions of constant variance, independence and normality for the calculated residuals.

Another statistic of interest is the mean squared error (MSE). Calculated by dividing SSE by the associated degrees of freedom, MSE is convenient to use when comparing the reduced and full models. Specifically one would expect to observe a reduction in MSE when a production rate variable is added to the reduced model.

The total sum of squares (SSTO) for the regression model is calculated by finding the difference between the logarithm of each observed value and the mean of those values for the dependent variable. The multiple coefficient of determination (R^2) is calculated by subtracting from one the sum of the square of each residual (SSE) from SSTO and dividing that result by the total sum of squares. The resulting ratio represents the variation in the dependent variable that is explained by the regression model. But since each variable is transformed to logarithms to linearize the terms, R^2 actually represents the variation in the logarithm of the hours per pound explained by the model

This transformation to logarithms somewhat obscures the interpretation of R^2 with respect to the true variable of interest, hours per pound. Therefore, another statistic is introduced that reduces the effect of the transformation. Termed R^2 actual, a calculation analogous to that for R^2 is performed on the variables expressed in their original form. Specifically, a SSTO actual is calculated by summing the square of the difference between each observed value of hours per pound and their mean. Each observation is predicted by the model in logarithms and then transformed to the original form. Actual residuals are then calculated, squared and summed producing an SSE actual. The R^2 actual statistic is calculated by dividing SSE actual by SSTO actual and subtracting the quotient from one. R^2 actual represents the variation in the actual hours per pound explained by the cost model.

If all three subhypotheses are not rejected, the MSE is relatively small, both R^2 and R^2 actual statistics are high and estimates of the model parameters follow the theory, the model is accepted as suitable for fitting the data. From this conclusion, it follows that the production rate is an important variable to be used in explaining variation in labor requirements to produce airframes.

HYPOTHESIS TWO

The direct labor hours required to fabricate and assemble an airframe are collected at a number of different process levels. For example, the total direct labor hours required to produce an airframe can be separately identified as in-plant or major subcontractor hours. Within a particular plant, a major portion of the direct labor hours can be identified as fabrication or assembly hours. It is of interest to know if the cumulative production and production rate cost model can be used to explain the behavior of direct labor hours at sublevels of aggregation.

Data at lower levels of aggregation are analyzed using the same cost model and regression techniques that are described under the first hypothesis. The model may not be as suitable for this investigation if there is variation in subcontracting as a percentage of the total workload as the program progresses. In order to exclude this source of variation it may be necessary to limit observations to those where major subcontracting is a relatively constant percentage.

The only difference between the tests and subhypotheses described under hypothesis one and those described under hypothesis two is the definition of the dependent variable. Under hypothesis two, the dependent variables are hours per pound required to fabricate or assemble the airframe. The definitions of the independent variables remain the same.

For each set of data examined in the model, the subhypotheses are:

(2A) H_0: B_1 and $B_2 = 0$

H_a: not both B_1 and $B_2 = 0$

(2B) H_0: $B_2 = 0$

H_a: $B_2 \neq 0$

(2C) H_0: The model is not appropriate.

H_a: The model cannot be rejected.

The same tests and rejection criteria are used to evaluate these subhypotheses as are used to evaluate the first three subhypotheses.

HYPOTHESIS THREE

One purpose of this research is to develop a model form and define variables so that model parameters can be tailored to a continuing airframe production program. These tailored models would then be used to predict the direct labor component of the cost of additional airframes. After successfully developing the models suggested under the first two hypotheses, they should be tested to see how well they predict.

In a real application of the model, the prediction would be beyond the range of the historical data. The only way to test the accuracy of the prediction would be to wait and see how many hours it takes to build the next airframe lot. To simulate this situation, the regression coefficients in the model are estimated with the last few observed data points omitted. Then using this new model, omitted values (which are known but not used in estimating the model coefficients) are predicted. Comparisons are then drawn between the actual and predicted hours as a subjective measure of predictive ability.

Hypothesis three is that the predictive capability of each model is good for one year into the future. Subjective evaluation of the accuracy of the forecasts is the basis for accepting or rejecting the hypothesis.

DATA SELECTION

The selection of programs to be examined is guided by convenience and accessability of data. There is no random sampling from all possible airframe production programs. Data from the F-4, F-102, and KC-135A production programs are examined.

The F-4 data are the most complete and comprehensive of the three programs. These data are used to test all three hypotheses. The F-102 data appear to be very accurate but are limited to total hours. Accordingly, hypothesis two is not evaluated with F-102 data. The KC-135A data are available in both total hours and lower process level hours.

The research procedure is to examine data sets from individual programs through application of the statistics and logic outlined here. Coefficients are estimated from a data set that tailor the proposed cost model to that set. Then the predictive ability of that model is tested.

There is no intent to develop a generalized cost model, only a generalized approach to building tailored cost models. In this sense, each data set from each program represents a unique population. Therefore, the departure from randomness in selecting the programs should not in itself bias the conclusions.

SUMMARY

The approach is to use multiple regression analysis of historical airframe production data to estimate coefficients in the proposed cumulative production and production rate cost model for a number of different data sets. Both statistical and subjective tests are used to verify the model and the assumption that the production rate is an important explainer of the variation in labor requirements to produce airframes. Finally, the ability of the model to predict the cost of new lots of airframes is tested. The next chapter discusses the results of the research.

CHAPTER V

ANALYSIS AND EVALUATION

This chapter is arranged in four sections. The first three are findings about individual programs. Each includes a discussion of the raw data, construction of the variables, results of the hypotheses testing and comments on the findings. The first section reports on the analysis of nine sets of F-4 data. Two sets of F-102A data are examined in the second section while five sets of KC-135A data are analyzed in the third. The chapter closes with a summary of the findings.

THE F-4 PROGRAM

The F-4 data are the most comprehensive of the three programs evaluated in this research. They include direct labor hours collected at many process levels as well as in total. The data cover 4665 airframes produced in 12 models. These airframes are produced in 57 lots from 1958 to 1975. The DCPR weight of the airframe increases from 15,300 pounds for early models to over 21,300 pounds for a more recently produced F-4E. After an initial peak of 71 airframes per month in 1967, the total delivery rate declines to a low of five per month in 1972 and then increases to 20 per month in 1975. The data are complex with many possible sources of variation for the direct labor hours. But the change in delivery rate over time makes the data unusually interesting for an investigation of the effect of a changing production rate on labor requirements.

Because the F-4 data are so comprehensive, evaluations of many different data sets are possible. In this study, evaluations are classified along two major lines, total direct production labor requirements and direct production labor requirements at lower process levels. There are five data sets used to evaluate the research procedure for total hours. Then the procedure is tested using four data sets from lower process levels.

COST MODEL 1 - TOTAL HOURS: F-4A-F

The first evaluation of hypotheses one and three is accomplished on input data that include the total direct labor hours required to produce F-4A, B, C, D, E and F model airframes in all 57 lots. These airframe models represent a direct evolution of the F-4 and provide an opportunity to examine the entire program for effects of the production rate on labor requirements. Labor for reconnaissance or other fighter versions are excluded from the lot average calculations in an effort to generate a more homogeneous data set.

Cost Model 1 has the form $y = B_0 \cdot x_1^{B_1} \cdot x_2^{B_2} \cdot 10^e$ as do all the models in this investigation. But in this case the dependent variable is the weighted average direct labor hours required to produce a pound of airframe in each lot. This variable is calculated by first developing the weighted average direct labor hours per airframe for each lot. Here the average labor requirement for each of the six airframe models is weighted by the corresponding number of airframes in the lot. Similarly, a weighted average DCPR weight is calculated. The average hours per airframe are then divided by the average pounds per airframe producing the dependent variable, hours per pound. The effects of design changes are assumed to be eliminated by including the DCPR weight in the dependent variable.

Cumulative production plot point is the first independent variable. It is obtained by dividing the total number of airframes in each lot by two and adding the quotient to the cumulative number of airframes already produced. In accordance with established learning curve technique, the plot point for the first lot is adjusted to allow for the steep drop in hours per pound for the first few airframes. This first lot plot point is extracted from learning curve tables using an arbitrarily selected 70 per cent learning factor.[42]

The second independent variable is the program lot average delivery rate. It is developed from the actual acceptance schedule as follows. First the months are identified during which the airframes from a particular lot are accepted by the customer. Then the total number accepted during those months is divided by the number of months. This produces a lot average delivery rate that closely approximates the rate that the airframes in the lot move through the plant. The three variables described above are listed in columns 2, 4 and 5 of Table 2.

The data are evaluated through multiple regression analysis in the following regression model: $\log y = \log B_0 + B_1 \log x_1 + B_2 \log x_2 + e$. The variables are transformed to logarithms to produce linearity in the model. Some important products of the regression are listed in Table 3.[43]

The three part test of hypothesis one indicates that the model fits the data set well. The fact that the F^*

---

[42]H. E. Boren, Jr. and H. G. Campbell, Learning Curve Tables: Volume II, 70-85 Per Cent Slopes (Santa Monica, California: The Rand Corporation, April 1970), p. 2.

[43]Throughout this study the primary regression results are obtained through application of a FORTRAN (WATFIV) program written by the author and tailored to the needs of the research. The program is listed in Appendix A.

TABLE 2

VARIABLES FOR ANALYSIS OF F-4A, B, C, D, E and F
TOTAL DIRECT PRODUCTION HOURS PER POUND

Lot Number (1)	Cumulative Production Plot Point (2)	Lot Average Manufacturing Rate (3)	Lot Average Monthly Delivery Rate (4)	Total Production Hours per Pound (5)
1	3.0		0.50	masked[a]
2	12.5		0.50	
3	20.5	0.19	0.86	
4	28.0	0.38	1.70	
5	40.0	0.58	7.67	
6	59.0	1.09	10.67	
7	83.0	1.26	10.00	
8	107.0	1.26	6.75	
9	131.0	1.41	11.00	
10	155.0	1.60	11.33	
11	179.0	1.71	13.50	
12	206.0	2.14	10.75	
13	243.0	2.93	12.20	
14	287.0	2.75	10.75	
15	333.0	2.82	11.50	
16	382.0	2.78	13.50	
17	432.0	2.63	24.00	
18	502.0	5.00	33.33	
19	602.0	5.79	35.00	
20	717.0	6.32	33.40	
21	837.0	6.67	41.25	
22	957.0	7.06	41.50	
23	1082.0	7.65	43.00	
24	1214.5	7.94	37.75	
25	1349.5	7.50	44.25	
26	1484.5	7.50	51.33	
27	1619.5	7.50	55.00	
28	1763.5	9.00	62.33	
29	1930.0	10.59	63.75	
30	2125.0	12.35	67.75	
31	2340.0	12.22	66.75	
32	2555.0	11.05	52.33	
33	2762.5	9.76	46.20	
34	2947.5	7.86	55.67	

TABLE 2-Continued

Lot Number (1)	Cumulative Production Plot Point (2)	Lot Average Manufacturing Rate (3)	Lot Average Monthly Delivery Rate (4)	Total Production Hours per Pound (5)
35	3110.0	7.62	60.00	masked[a]
36	3252.5	6.25	44.67	
37	3375.0	6.00	39.00	
38	3475.0	4.21	40.33	
39	3555.0	4.44	39.67	
40	3635.0	4.44	31.00	
41	3715.0	4.44	25.50	
42	3800.0	4.74	21.20	
43	3894.0	4.90	25.25	
44	3939.0	4.60	28.00	
45	4062.5	2.89	23.67	
46	4118.0	2.95	15.75	
47	4175.0	2.90	13.33	
48	4220.5	1.43	6.29	
49	4252.5	1.29	5.67	
50	4283.0	1.30	6.80	
51	4313.5	1.48	10.67	
52	4348.5	1.95	12.25	
53	4387.5	2.05	14.50	
54	4436.5	3.11	14.80	
55	4495.0	3.05	16.40	
56	4558.5	3.45	17.20	
57	4701.0	3.13	12.57	

SOURCES: Data for calculating the plot point are obtained from the "F-4 Procurement Summary", a single page periodic report developed by Department 926 of the McDonnell Douglas Corporation and dated 3 February 1974.

The lot average manufacturing rate is developed from four sources. The lot release date for lots 16 through 57 is read from a series of graphs entitled "F4 Manufacturing Schedule", developed by Department 149A of the McDonnell Douglas Corporation and dated from 24 January 1963 to 11 August 1975. The release dates for lots 2 through 15 are read from Section 3.3.2 of MCAIR Report 7290 entitled "F 4 Cost Data" and updated in August of 1975. The lot sizes are read from the "F-4 Procurement Summary" listed above. The lot completion dates are extracted from

TABLE 2-Continued

McDonnell Douglas Report Number 3.0, file number 3231, entitled "F-4 Scheduled Deliveries vs Actuals", prepared by Department 013 and revised 7 August 1975.

The program delivery rate is also developed from McDonnell Douglas Report Number 3.0 described above.

The total hours per pound variable is developed from three sources. The first 50 lots of data are extracted from a summary report entitled "McDonnell-Douglas F-4 Airplane Production Man-hours per Pound, F-4B, C, D/E Airplanes" and dated 10 January 1974. The labor hour data for lots 51 through 57 are extracted from Section 2.2.4 of MCAIR report 7290 listed above. DCPR weights for lots 51 through 57 are extracted from a McDonnell Douglas summary report entitled "DCPR Weight-Pounds per Airplane" and dated 17 November 1975. For airframes 34 through 57, the DCPR weight includes useful load items at 1039 pounds. The hours required for major subcontracting are assumed to be constant for lots 51 through 57 at the same level as that experienced for lot 50.

[a]The Total Production Hours per Pound are considered proprietary by the manufacturer. Accordingly, these data are masked in the published version of this dissertation. Access to these data can be provided by the author upon approval of the McDonnell Aircraft Co. representatives in St. Louis. The masked data have been reviewed in the draft dissertation by the Doctoral Dissertation Committee.

TABLE 3

REGRESSION RESULTS - MODEL 1
(F-4A-F Total Hours Versus Plot Point and Delivery Rate)

Estimate B_0	= masked[a]	R^2 actual	= 0.982
Estimate B_1	= -0.261	t ratio B_1	= 26.10
Estimate B_2	= -0.169	t ratio B_2	= 11.27
F^* (full)	= 1206.57	$F(2,54), 0.05$	= 3.17
F^* (incremental)	= 122.85	$F(1,54), 0.05$	= 4.02
MSE (full)	= 0.0016	MSE (reduced)	= 0.0053
R^2 (full)	= 0.978	R^2 (reduced)	= 0.928

Case	Observed Value	Predicted Value	Residual	Per Cent Deviation
1	masked[a]	masked[a]	masked[a]	masked[a]
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				

TABLE 3-Continued

Case	Observed Value	Predicted Value	Residual	Per Cent Deviation
27	masked[a]	masked[a]	masked[a]	masked[a]
28				
29				
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				

[a]Total Production Hours per Pound are considered proprietary by the manufacturer. Accordingly, these data are masked in the published version of this dissertation. Access to these data can be provided by the author upon approval of McDonnell Aircraft Co. representatives in St. Louis. The masked data have been reviewed in the draft dissertation by the Doctoral Dissertation Committee.

statistic of 1206.56 is larger than the theoretical F value of 3.17 indicates that the set of variables, logarithm plot point and logarithm rate, are related to the dependent variable, logarithm hours per pound. The F^* incremental statistic of 122.85 is larger than the theoretical F value of 4.02. This indicates that the logarithm of the delivery rate contributes importantly to the model. All F tests are conducted at the 0.05 significance level.

An examination of the residuals indicates no serious departures from the assumptions of the model. A plot of the residuals versus the corresponding fitted values of the dependent variables reveals little pattern. A plot of the residuals against each of the independent variables gives no evidence of a relationship. Plotting the cumulative frequency of the residuals on normal probability paper indicates no significant departure from a straight line.

A plot of the residuals versus the time sequence in which they were observed presents a suggestion of correlation. But a check of the data reveals that much of the pattern is due to introduction of major design changes. For example, a C model airframe is introduced at lot 15 resulting in an increase in average direct labor hour expenditures. This increase causes the immediately succeeding residuals to become positive until the change is absorbed into the process. This cause of patterns in the residuals does not in itself invalidate the assumption of independence. It does indicate that dividing the hours per pound by the DCPR weight does not completely eliminate the effects of major design changes on labor requirements.

The theory advanced in the previous chapter for the behavior of the variables is supported by the estimated model coefficients. The estimate of the B_1 coefficient is negative indicating that an increase in cumulative production leads to a decrease in unit labor requirements. The B_2 coefficient estimate is also negative. This supports the idea that a delivery rate increase (up to plant capacity) can be expected to decrease unit labor requirements.

Comparisons of statistics produced from a reduced model (with the delivery rate variable omitted) and a full model support the conclusion that the delivery rate is an important explanatory variable. The MSE calculated from data in the reduced model is 0.0053 as compared to a MSE for the full model of 0.0016. This indicates a 70 per cent reduction in the MSE due to addition of the rate variable. Similarly, the multiple coefficient of determination, R^2, increases 4.98 per cent as the result of adding the rate term to a reduced model.

The R^2 actual statistic, which is calculated from predicted and actual values in their original form, is 0.982. This ratio indicates that 98.2 per cent of the variation in hours per pound is explained by the cost model.

The t ratios for the B_1 and B_2 estimates are listed with the regression results for readers that prefer this statistic. In this analysis, the incremental test based on the F statistic is equivalent to the familiar t test. Therefore no further comment is offered on the t statistic.

These tests sustain the first hypothesis that the production rate, when expressed in the cumulative production and production rate cost model, explains an important part of the variation in the total direct production labor required to produce F-4 airframes. But this conclusion is of little value if the model does not predict well.

One way to test the predictive ability of the full model is to successively remove a few of the most recent cases from the data set. One may then perform the regression analysis with the remaining data to develop new regression coefficients. Using the model developed from the truncated data set, one can forecast the points that are removed. A comparison between the forecast and the observed values gives a measure of the predictive ability of the model tailored from the truncated data set.

As successively larger increments of data points are removed from the original data set, a family of estimated regression coefficient sets is developed for the cost model. Each of these tailored models can be used to predict some particular future data point. These predictions, compared to an observed value, are the basis for subjective evaluation of the ability of the model to forecast.

Table 4 lists the results of estimating the full cumulative production and production rate cost model from successively fewer observations. For example, in the row for 50 cases are estimated regression coefficients and statistics produced from regression analysis of the first 50 of the 57 observations in the data set. Also included

TABLE 4

PREDICTIVE ABILITY - FULL MODEL 1
(Total Hours: F-4A-F)

Cases	Estimate B_0	Estimate B_1	Estimate B_2	Forecast Lot 57[a]	Per Cent Deviation
55	32.654	-0.261	-0.169	2.34	-2.63
50	33.027	-0.266	-0.162	2.31	-1.32
45	32.556	-0.259	-0.172	2.36	-3.50

[a]The observed value of the dependent variable for lot 57 is 2.28 hours per pound.

in the table are forecasts for lot 57 using the different sets of estimated coefficients. During this program, a new lot is released approximately every fourth month. Therefore removing seven observations and developing the model on 50 cases allows about a 28 month forecast to lot 57. The column captioned "Forecast Lot 57" lists the forecast direct hours per pound after transformation from common logarithms to their original form.

The per cent deviation column reflects the difference between the observed and the predicted value of the hours per pound divided by the observed value. These ratios are expressed as a percentage. The choice of lot 57 for the comparison is arbitrary. It is clear that the observed value of hours per pound is a random variable subject to many sources of error. It could well be that choice of another data point for the comparison would give better (or worse) relative results. Nevertheless, the approach provides a stationary mark for the purpose of comparison.

As one might anticipate, the model accuracy (as measured at lot 57) decreases as fewer data points are used to estimate the coefficients and the prediction span becomes longer. With 50 cases in the model, where per cent deviation happens to be the smallest, the forecast value for lot 57 is 2.34 as compared to the observed value of of 2.28 hours per pound. When the coefficients are estimated from 45 cases, the forecast for lot 57 increases to 2.37 hours per pound which is still within 3.5 per cent of the target.

A reduced model (without the delivery rate variable) is employed to estimate regression coefficients using the same observation sets used to develop the full model. This reduced model is identical to the unit learning curve model. The results of these evaluations are listed in Table 5.

TABLE 5

PREDICTIVE ABILITY - REDUCED MODEL 1

Cases	Estimate B_0	Estimate B_1	Forecast Lot 57	Per Cent Deviation
55	33.339	-0.336	1.95	14.5
50	35.138	-0.347	1.86	18.4
45	37.906	-0.363	1.76	22.8

Once more using the actual direct hours per pound for lot 57 as the prediction target, per cent deviation from the observed value is calculated. The reduced model predicts low in every case. The magnitude of the differences is much larger than those produced from the full model. For example, for a reduced model developed from 50 observations, the deviation for a prediction of lot 57 is 18.4 per cent. This result is much less accurate than the prediction deviation of 1.32 per cent from the full model.

One can conclude from these observations that for this data set, the full cumulative production and production rate cost model is superior in predicting power to the reduced model. Furthermore, the accuracy of prediction of the full model appears to warrant its use as a forecasting tool. Accordingly, hypothesis three is sustained for Model 1.

COST MODELS 2 AND 3 - TOTAL HOURS: F-4B-F

The production rate variable can be expressed in a number of different ways. In Cost Model 1, it is a lot average delivery rate. In Cost Model 2, the production rate is represented by the lot average manufacturing rate. It is constructed by dividing the total number of airframes in a lot by the lot manufacturing time span. This span bridges the date the first work orders are released to fabricate parts for a lot until the date the last airframe in the lot is accepted by the customer. Because this span is so much larger than the span of months during which airframes from a particular lot are delivered, the lot average manufacturing rate is a smaller number than the lot average delivery rate.

Cost Model 2 is identical to the first cost model in all other respects. The data used to test the model are listed in columns 2, 3 and 5 of table 2. Observations for lots one and two are not constructed because the lot release dates are not included in the available data. Accordingly, the analysis is accomplished with 55 lots of F-4B, C, D, E and F data. Regression analysis results are listed in Table 6.

TABLE 6

REGRESSION RESULTS - MODEL 2

Estimate B_0	=	masked	R^2 actual	=	0.971
Estimate B_1	=	-0.246	t ratio B_1	=	26.74
Estimate B_2	=	-0.183	t ratio B_2	=	11.37
F^* (full)	=	920.19	F(2,52),0.05	=	3.18
F^* (incremental)	=	129.30	F(1,52),0.05	=	4.02
MSE (full)	=	0.0013	MSE (reduced)	=	0.0045
R^2 (full)	=	0.973	R^2 (reduced)	=	0.904

The statistics indicate that the lot average manufacturing rate logarithm is also an important variable for explaining variation in the logarithm of hours per pound. The full regression model F^* statistic of 920.19 is larger than the theoretical F value of 3.18. This indicates that both of the regression coefficients are not zero with a 95 per cent level of confidence. Similarly, the F^* statistic of 129.30 produced by incrementally adding the manufacturing rate to the reduced model is larger than 4.02. Again, at the 95 per cent level of confidence, one can conclude that the B_2 coefficient is not zero and that the manufacturing rate is an important variable. The increase in the R^2 statistic from 0.904 to 0.973, which results from adding manufacturing rate to the reduced model, provides further evidence that the manufacturing rate variable contributes importantly to the model.

Other tests also validate the model. As in Model 1, the MSE statistic for full regression Model 2 is less than one-third the value produced from the reduced regression model. The high value of R^2 actual indicates that the full model explains some 97 per cent of the total variation in the hours per pound for the 55 observations. An analysis of the residuals by the methods outlined previously reveals no serious departures from the assumptions of the model. Accordingly, Model 2 is also considered to be valid and hypothesis one is accepted.

The predictive ability of Model 2 is examined by again successively removing observations from the end of the data set, estimating model coefficients from these truncated sets and predicting the hours per pound value for lot 57. The results of this examination for the full model are listed in Table 7. The model appears to be quite

TABLE 7

PREDICTIVE ABILITY - FULL MODEL 2
(Total Hours: F-4B-F)

Cases	Estimate B_0	Estimate B_1	Estimate B_2	Forecast Lot 57[a]	Per Cent Deviation
50	22.442	-0.249	-0.180	2.23	2.19
45	22.652	-0.252	-0.176	2.20	3.51
40	20.042	-0.224	-0.216	2.36	-3.51

[a]The observed dependent variable for lot 57 (case 55) is 2.28 hours per pound.

stable for this test. It predicts within plus or minus 3.5 per cent for some 45 months at the arbitrarily selected test points. This stability is further evidenced by the relatively small change in the regression coefficients as up to 15 observations are removed from the original data set.

When compared to the predictive ability of a reduced Model 2, the full model is clearly superior. The reduced model test results are listed in Table 8. The family of reduced model coefficients is estimated from data points that are common to those used to develop the coefficients listed in Table 7. Thus the results are directly comparable. For example, for coefficients estimated from the

TABLE 8

PREDICTIVE ABILITY - REDUCED MODELS 2 AND 3
(F-4B-F Total Hours Versus Plot)

Cases	Estimate B_0	Estimate B_1	Forecast Lot 57[a]	Per Cent Deviation
50	28.676	-0.317	1.97	13.60
45	31.944	-0.337	1.86	18.42
40	33.534	-0.347	1.78	21.93

[a]The observed dependent variable for lot 57 (case 55) is 2.28 hours per pound.

first 50 cases, the full and reduced models forecast the hours per pound for case 55 (lot 57) at 2.23 and 1.97 respectively. When compared with the observed value of 2.28 hours per pound for lot 57, the improvement in predictive ability resulting from adding a manufacturing rate variable to the reduced model is apparent. This same improvement in predictive ability exists at all of the test points.

It is concluded that the predictive ability of full Model 2 is superior to that of the reduced Model 2. Furthermore, the ability of the full model to predict beyond the range of the data appears to be good. This is evidenced by the small per cent deviation for predictions of up to 45 months. Accordingly, hypothesis three is also accepted for Model 2.

It is desirable to compare the production rate when expressed as a delivery rate (Model 1) or as a manufacturing rate (Model 2). To facilitate the comparison, coefficients for a third regression model are estimated. This model is based on the same observations as is Model 1 except that the data for lots 1 and 2 are omitted. The data are listed in columns 2, 4 and 5 of Table 2. With this data set, the comparison between the effectiveness of the manufacturing rate and the delivery rate in the cumulative production and production rate cost model is simply a comparison between the tests of Models 2 and 3.

Before this comparison is made, the tests to validate the model and the delivery rate in the model are performed. Because the data used in estimating the coefficients for Model 3 are so nearly the same as for Model 1, one would expect similar results from the tests. This is so, and hypothesis one is sustained for Model 3.

The primary purpose for constructing Model 3 is to compare the effectiveness between the production rate expressed as a lot average manufacturing rate or delivery rate. Some statistics and characteristics of the models are listed in Table 9 to assist in the comparison.

TABLE 9

COMPARISON OF MANUFACTURING AND DELIVERY RATES IN MODELS 2 AND 3

Cost Model	2	3
Production Rate	Manufacturing Rate	Delivery Rate
Estimate of B_0	masked	masked
Estimate of B_1	-0.246	-0.257
Estimate of B_2	-0.183	-0.161
F^* (full)	920.19	742.13
F^* (incremental)	129.30	95.20
MSE (full)	0.0013	0.0016
MSE (reduced)	0.0045	0.0045
R^2 (full)	0.973	0.966
R^2 (reduced)	0.904	0.904
R^2 actual	0.971	0.966

For these data sets of 55 observations, the statistics suggest that the manufacturing rate is the better representative of the production rate. The MSE (expressed in logarithms) for Model 2 is 0.0003 smaller than that for Model 3. The R^2 actual statistic for Model 2 is 0.005 larger than for Model 3. The other measures similarly indicate that manufacturing rate is better. However, the differences are so small that they could easily be due to chance variation in the data. Thus, little confidence is given to the statement that one is better than the other.

On a time axis, the manufacturing rate aligns conceptually with the average hours required to produce a pound of airframe in a particular lot. Conversely, airframe delivery as related by the delivery rate must lag the expenditure of the average hours per pound by some months. This observation may provide some clue as to why lot average production rate appears to be the stronger representative of production rate.

COST MODELS 4 AND 5

The manufacturer's data collection system for F-4 direct labor hours changed between lots 15 and 16. Although grand total production hours are available, data for the fabrication and assembly portions of the manufacturing process do not match before and after that change point. To provide a basis for comparison of models developed from fabrication hours, assembly hours and total hours, coefficients for a model are developed from total F-4B, C, D, E and F hours per pound for lots 15 through 57.

Cost Model 4 is the cumulative production and production rate cost model with the manufacturing rate form of the x_2 variable. The data for lots 16 through 57, listed in columns 2, 3 and 5 of Table 2, are used to estimate the model coefficients. Table 10 presents the regression analysis results.

TABLE 10

REGRESSION RESULTS - MODEL 4
(F-4B-F Total Hours)

Estimate B_0	= masked	R^2 actual	= 0.887
Estimate B_1	= -0.230	t ratio B_1	= 14.62
Estimate B_2	= -0.157	t ratio B_2	= 8.41
F^* (full)	= 112.96	F(2,39),0.05	= 3.24
F^* (incremental)	= 70.85	F(1,39),0.05	= 4.09
MSE (full)	= 0.0009	MSE (reduced)	= 0.0025
R^2 (full)	= 0.853	R^2 (reduced)	= 0.585

The statistical tests of the regression results indicate that the model is appropriate for the set of 42 observations. The F tests indicate that the coefficient estimates are not 0 with 95 per cent confidence. Examination of the residuals reveals no important departures from assumptions of the model.

While the R^2 actual statistic of 0.887 leaves some 11 per cent of the variation in hours unexplained, it still indicates that this simple model captures 89 per cent of the variation in a most complex process. Nevertheless, it is also apparent that the fit is not as good as that of Model 2 which is estimated with more observations. Perhaps it is important to include the early observations in models of F-4 data.

When comparing the full model with the reduced model, two statistics indicate that delivery rate is an important explainer of direct labor requirements. First, the R^2 statistic is improved more than 23 per cent by adding the rate variable to the reduced model. Second, the MSE for the full model is less than one-half the MSE for the reduced model. These and other tests of the three sub-hypotheses provide no basis for rejecting the model. Therefore, hypothesis one is accepted for Model 4.

Table 11 lists the results of tests for the pre-

TABLE 11

PREDICTIVE ABILITY - FULL MODEL 4
(Total Hours: F-4B-F)

Cases	Estimate B_0	Estimate B_1	Estimate B_2	Forecast Lot 57[a]	Per Cent Deviation
40	19.012	-0.232	-0.156	2.24	1.75
38	19.166	-0.233	-0.156	2.23	2.19
36	19.256	-0.235	-0.151	2.22	2.63
34	19.234	-0.236	-0.149	2.21	3.07
32	19.023	-0.240	-0.126	2.16	5.27
30	18.865	-0.236	-0.136	2.20	3.51

[a]The observed value at lot 57 (case 42) is 2.28 hours per pound.

dictive ability of Model 4. The tests show that the model form is well suited to prediction beyond the range of the data. For example, with coefficients estimated from the first 40 of the 42 cases, the model predicts 1.75 per cent high for lot 57, a span of about eight months. With coefficients estimated from 36 of the 42 cases, the prediction accuracy for the 57th lot decreases to 2.63 per cent. The time span for the six lots is some 23 months. The other calculations reflect similar results.

For this data set, the procedure appears to be quite stable. As additional cases are removed from each regression analysis, one would anticipate changes in the estimated coefficients and the resulting predictions. Yet as Table 11 reflects, the coefficients do not change very much as observations are removed. The predictions for lot 57 resulting from these different sets of coefficients are also relatively constant. The range for the seven different estimates is 0.08 hours per pound over a time span of about 45 months.

The results of the predictive ability test for a reduced Model 4 are listed in Table 12. The new coefficients are estimated with the same cases used in the tests reported in Table 11. This permits a comparison of a model that includes the manufacturing rate with one that lacks it.

The stabilizing effect of the manufacturing rate on the full model is even more apparent in this comparison.

TABLE 12

PREDICTIVE ABILITY - REDUCED MODELS 4 AND 5
(F-4B-F Total Hours versus Plot)

Cases	Estimate B_0	Estimate B_1	Forecast Lot 57[a]	Per Cent Deviation
40	10.514	-0.187	2.16	5.26
38	10.815	-0.191	2.15	5.70
36	11.682	-0.202	2.12	7.02
34	12.950	-0.217	2.07	9.21
32	15.118	-0.239	2.00	12.28
30	15.789	-0.246	1.97	13.60

[a]The observed value of the dependent variable for lot 57 (case 42) is 2.28 hours per pound.

In the reduced model, the per cent deviation of the forecast value from the observed value increases as fewer observations are included in estimating the coefficients of the reduced model. As previously noted, per cent deviation for predictions from the full model remain relatively constant.

With all 42 observations in the data set, the prediction accuracy of the full model is much better than that of the reduced model. This difference is maintained as different numbers of cases are used to evaluate the model. In addition, the predictions from the full model form are quite accurate when compared to the target observed value. These indications of accuracy permit one to conclude that the full Model 4 is a good predictor of total hours per pound and that hypothesis three may be accepted.

The coefficients of Model 5 are estimated from the same 42 observations used in Model 4 except that the x_2 variable is the delivery rate. The variables are listed in columns 2, 4 and 5 of Table 2 for lots 16 through 57. Results of the regression analysis are listed in Table 13.

TABLE 13

REGRESSION RESULTS - MODEL 5
(F-4B-F Total Hours)

Estimate B_0	= masked	R^2 actual	= 0.854
Estimate B_1	= -0.229	t ratio B_1	= 13.10
Estimate B_2	= -0.136	t ratio B_2	= 7.14
F^* (full)	= 88.96	F(2,39),0.05	= 3.24
F^* (incremental)	= 50.96	F(1,39),0.05	= 4.09
MSE (full)	= 0.0011	MSE (reduced)	= 0.0025
R^2 (full)	= 0.820	R^2 (reduced)	= 0.585

Tests of the statistics and the residuals reflect that the model is appropriate for the data set. The production rate proxy, delivery rate, is again an important explanatory variable in the model.

As with the comparison between Cost Models 2 and 3, Model 4 with the manufacturing rate variable produces slightly better statistics than Model 5 with the delivery rate variable. In both comparisons, the R^2 statistics are 3.3 per cent higher in the models with the manufacturing rate. Similarly, the MSE statistic is smaller for the models with the manufacturing rate. This comparison also favors the manufacturing rate variable as the better production rate representative although the evidence is not conclusive.

The results of tests for the predictive ability of full Model 5 are listed in Table 14. The model form appears to forecast well. This is indicated by pre-

TABLE 14

PREDICTIVE ABILITY - FULL MODEL 5
(Total Hours: F-4B-F)

Cases	Estimate B_0	Estimate B_1	Estimate B_2	Forecast Lot 57[a]	Per Cent Deviation
40	23.050	-0.229	-0.137	2.35	3.07
38	23.064	-0.229	-0.136	2.36	3.51
36	23.176	-0.233	-0.131	2.32	1.75
34	22.869	-0.235	-0.123	2.30	0.09
32	21.529	-0.239	-0.099	2.22	-2.70
30	22.052	-0.234	-0.115	2.28	0.00

[a]The observed value for lot 57 (case 42) is 2.28 hours per pound.

dictions of within five per cent of the observed value of 2.28 hours per pound for spans of up to 48 months. When compared with the predictive ability of reduced Model 5 (listed in Table 13), the full model is more stable and predicts better. It is concluded that Model 5 is also a good predictive model and hypothesis three is accepted.

COST MODELS 6 AND 7
F-4 FABRICATION HOURS PER POUND

The second hypothesis concerns the direct labor hours required to produce the airframe at lower process levels. Two additional levels of data aggregation are examined. They are the direct labor hours required to fabricate the airframe parts and the hours required to assemble those parts into an airframe. These data are analyzed in the cumulative production and production rate cost model with the manufacturing rate and the delivery rate serving as alternate representatives of the production rate. Four models, numbered 6 through 9 are tested to determine if hypothesis two is viable. Variables used to test the procedure with the four models are listed in Table 15.

Models 6 and 7 are the models with the dependent variable expressed as fabrication hours per pound. The cumulative production variable for the two models is the program total cumulative production plot point. The production rate representative for Model 6 is the manufactur-

TABLE 15

VARIABLES FOR ANALYSIS OF F-4 FABRICATION AND ASSEMBLY HOURS PER POUND

Lot Number (1)	Cumulative Production Plot Point (2)	Lot Average Monthly Manufacturing Rate (3)	Lot Average Monthly Delivery Rate (4)	Fabrication Hours Per Pound (5)	Assembly Hours Per Pound (6)
16	382.0	2.78	13.50	1.574	1.740
17	432.0	2.63	24.00	1.459	1.783
18	502.0	5.00	33.33	1.311	1.514
19	602.0	5.79	35.00	1.174	1.200
20	717.0	6.32	33.40	1.151	1.054
21	837.0	6.67	41.25	1.103	1.002
22	957.0	7.06	41.50	0.954	0.900
23	1082.0	7.65	43.00	0.927	0.836
24	1214.5	7.94	37.75	0.911	0.808
25	1349.5	7.50	44.25	0.892	0.837
26	1484.5	7.50	51.33	0.919	0.917
27	1619.5	7.50	55.00	0.891	0.895
28	1763.5	9.00	62.33	0.849	0.873
29	1930.0	10.59	63.75	0.795	0.862
30	2125.0	12.35	67.75	0.808	0.861
31	2340.0	12.22	66.75	0.837	0.883
32	2555.0	11.05	52.33	0.873	0.961
33	2762.5	9.76	46.20	0.896	0.966
34	2947.5	7.86	55.67	0.860	0.928
35	3110.0	7.62	60.00	0.859	0.854
36	3252.5	6.25	44.67	0.803	0.816

TABLE 15-Continued

Lot Number (1)	Cumulative Production Plot Point (2)	Lot Average Monthly Manufacturing Rate (3)	Lot Average Monthly Delivery Rate (4)	Fabrication Hours Per Pound (5)	Assembly Hours Per Pound (6)
37	3375.0	6.00	39.00	0.816	0.773
38	3475.0	4.21	40.33	0.842	0.765
39	3555.0	4.44	39.67	0.830	0.723
40	3635.0	4.44	31.00	0.830	0.773
41	3715.0	4.44	25.50	0.813	0.733
42	3800.0	4.74	21.20	0.767	0.721
43	3894.0	4.90	25.25	0.782	0.688
44	3989.0	4.60	28.00	0.815	0.715
45	4062.5	2.89	23.67	0.942	0.689
46	4118.0	2.95	15.75	0.891	0.676
47	4175.0	2.90	13.33	0.928	0.662
48	4220.5	1.43	6.29	0.969	0.959
49	4252.5	1.29	5.67	1.010	0.900
50	4283.0	1.30	6.80	0.970	0.924
51	4313.5	1.48	10.67	0.927	0.786
52	4348.5	1.95	12.25	0.907	0.861
53	4387.5	2.05	14.50	0.882	0.847
54	4436.5	3.11	14.80	0.796	0.786
55	4495.0	3.05	16.40	0.763	0.738
56	4558.5	3.45	17.20	0.776	0.764
57	4701.0	3.13	12.57	0.765	0.745

TABLE 15-Continued

SOURCES: Fabrication and assembly hours per lot are extracted from Section 2.2.4 of MCAIR report 7290 entitled "F 4 Cost Data", produced by the McDonnell Douglas Corporation and updated in August of 1975.

DCPR weights are read from a McDonnell Douglas summary report entitled "DCPR Weight-Pounds per Airplane" and dated 17 November 1975.

Plot points, manufacturing rates and delivery rates are identical to those listed in Table 2.

ing rate and for Model 7 it is the delivery rate. Observations are taken at lots 16 through 57. Thus Models 6 and 7 are identical except for the different production rate variables. The results of the regression analysis of these models are presented in Tables 16 and 17 for evaluation and comparison.

The statistical tests indicate rejection of the corresponding null subhypotheses for both of the models. The theoretical F value of 3.24 is smaller than the F^* statistic for each model. The theoretical F incremental value of 4.09 is also smaller than the F^* incremental statistic in each case.

At the outset of this investigation, the labor required to fabricate parts is assumed to be production rate sensitive. Prorating the hours required to set up machine tools and tooling over larger quantities at higher production rates can explain this behavior. The assumption holds for the F-4B-F fabrication hours for lots 16 through 57. The sign of the estimated B_2 coefficient in Model 6 is negative as anticipated. Furthermore, the explanatory ability of the full model becomes 27 per cent greater than the reduced model. Finally, addition of the manufacturing rate variable to the reduced model causes the MSE to drop some 70 per cent in magnitude. The same type of improvement over reduced model statistics is noted

TABLE 16

REGRESSION RESULTS - MODEL 6
(F-4 Fabrication Hours)

Estimate B_0	= 6.328	R^2 actual	= 0.919
Estimate B_1	= -0.221	t ratio B_1	= 17.22
Estimate B_2	= -0.148	t ratio B_2	= 9.75
F^* (full)	= 155.93	F(2,39),0.05	= 3.24
F^* (incremental)	= 95.10	F(1,39),0.05	= 4.09
MSE (full)	= 0.0006	MSE (reduced)	= 0.0020
R^2 (full)	= 0.889	R^2 (reduced)	= 0.618

TABLE 17

REGRESSION RESULTS - MODEL 7
(F-4 Fabrication Hours)

Estimate B_0	= 7.601	R^2 actual	= 0.889
Estimate B_1	= -0.219	t ratio B_1	= 14.71
Estimate B_2	= -0.127	t ratio B_2	= 7.82
F^* (full)	= 111.60	F(2,39),0.05	= 3.24
F^* (incremental)	= 61.21	F(1,39),0.05	= 4.09
MSE (full)	= 0.0008	MSE (reduced)	= 0.0020
R^2 (full)	= 0.851	R^2 (reduced)	= 0.618

for the data in Model 7.

The explanatory ability of Models 6 and 7 is also good. As indicated by the R^2 actual statistic, Model 6 explains 92 per cent of the variance in the hours required to fabricate parts. The comparable statistic for Model 7 is 89 per cent.

These other tests of the analysis products also show the production rate as an important variable in explaining variation in fabrication hours per pound. Therefore, hypothesis two is accepted for Models 6 and 7.

As with the other model pairs examined so far, the statistics associated with the model that includes the manufacturing rate are slightly better than those from the model that includes the delivery rate. For example, the R^2 actual statistic for Model 6 is three per cent higher than the value for Model 7.

The predictive ability of the models is evaluated by estimating the model coefficients with successively smaller data sets. Omitted observations are predicted from each new model as if they were unknown. The observed value for lot 57 is again chosen as the prediction target. Conclusions about the model's stability and predictive ability are drawn from a comparison between this target and its prediction. Table 18 presents the results of estimating new coefficients for the full Model 6.

TABLE 18

PREDICTIVE ABILITY - FULL MODEL 6
(Fabrication Hours: F-4)

Cases	Estimate B_0	Estimate B_1	Estimate B_2	Forecast Lot 57[a]	Per Cent Deviation
42	6.328	-0.221	-0.148	0.825	-7.84
40	6.166	-0.217	-0.150	0.830	-8.50
38	5.990	-0.212	-0.152	0.839	-9.67
36	5.971	-0.211	-0.155	0.840	-9.80
34	5.986	-0.209	-0.162	0.850	-11.11
32	5.991	-0.209	-0.164	0.849	-10.98
30	6.042	-0.212	-0.154	0.844	-10.33

[a]The observed value of fabrication hours per pound at lot 57 (case 42) is 0.765.

The full model appears to be quite stable when predicting fabrication hours per pound. This conclusion is indicated by the small range of the predictions produced by the seven sets of estimated coefficients. They vary from 0.825 to 0.850 hours per pound over spans from zero to 45 months.

The accuracy of the seven predictions for lot 57 is not very good. The forecasts range from 7.84 to 11.11 per cent higher than the observed value for this particular point. These data suggest that a prediction of fabrication hours for lots beyond 57 would also be high. For such a forecast, logic would indicate a downward adjustment of the prediction to compensate for the error.

The indicated predictive ability of the model is not good. According to the test plan, deviations larger than five per cent are unacceptable. Therefore, hypothesis three is not accepted for Model 6.

The predictive ability of the full Model 7 is also tested. As compared to the tests run on Model 6, the results are not as good. The per cent deviation of each prediction is larger. Furthermore, the stability of the model is not as good with a range of 0.038 hours per pound for the seven estimates. Therefore, hypothesis three is also rejected for Model 7.

COST MODELS 8 AND 9
F-4 ASSEMBLY HOURS PER POUND

The hours required to assemble parts into an airframe are also assumed to be inversely affected by the production rate. An increase in rate would logically be implemented with higher loading per station and more stations. These management actions should lead to greater task specialization per worker. This should be accompanied by a conservation of some time that would be expended when changing from one task to another at the lower rate. Furthermore, it is assumed that the sense of urgency that accompanies a higher rate of production motivates the worker to produce faster. The opposite results are expected for a production rate decrease.

Cost Models 8 and 9 examine assembly hours per pound as a function of cumulative production and production rate. The proxy for production rate in Model 8 is the program manufacturing rate while the program delivery rate represents the production rate in Model 9. Data for the models are also listed in Table 15. Regression analysis results are reported in Table 19 for Model 8 and Table 20 for Model 9.

The statistics produced from Models 8 and 9 support the assumption that assembly hours are also inversely and importantly related to the production rate. For Model 8, with the manufacturing rate independent variable, the estimate of the B_2 coefficient is negative. The F^* incremental statistic is larger than the theoretical value of 4.09. Similarly, Model 9, with the delivery rate variable, also has a negative B_2 estimate and an acceptable F^* incremental statistic of 10.96. Another measure of the contribution of the production rate to the model is the improvement in the coefficient of determination from the reduced model to the full model. For Model 8, the improvement is 8.6 per cent while for Model 9 it is 7.5 per cent.

Neither model with assembly hours fits the data as

TABLE 19

REGRESSION RESULTS - MODEL 8

Estimate B_0	= 9.016	R^2 actual	= 0.797
Estimate B_1	= -0.279	t ratio B_1	= 10.64
Estimate B_2	= -0.112	t ratio B_2	= 3.62
F^* (full)	= 56.59	F(2,39),0.05	= 3.24
F^* (incremental)	= 13.10	F(1,39),0.05	= 4.09
MSE (full)	= 0.0025	MSE (reduced)	= 0.0033
R^2 (full)	= 0.744	R^2 (reduced)	= 0.658

TABLE 20

REGRESSION RESULTS - MODEL 9

Estimate B_0	= 10.400	R^2 actual	= 0.775
Estimate B_1	= -0.278	t ratio B_1	= 10.33
Estimate B_2	= -0.097	t ratio B_2	= 3.31
F^* (full)	= 53.46	F(2,39),0.05	= 3.24
F^* (incremental)	= 10.96	F(1,39),0.05	= 4.09
MSE (full)	= 0.0026	MSE (reduced)	= 0.0033
R^2 (full)	= 0.733	R^2 (reduced)	= 0.658

well as Models 4 and 5 with total hours or Models 6 and 7 with fabrication hours. For example, the R^2 actual statistic for Model 8 is 0.797 and for Model 9 it is 0.775. The corresponding statistics for the total hours models are 0.887 and 0.854. Similarly, the corresponding statistics for the fabrication hours models are 0.919 and 0.889. Nevertheless, the statistical and subjective tests do not indicate model rejection and Models 8 and 9 are indicated as appropriate for the data set. Hypothesis two is accepted for both models.

The results of the predictive ability test for Model 8 (Table 21) are good. Little change is observed when removing the first eight observations in steps of two. The forecast of the assembly hours required to produce lot 57 with a model developed from 34 cases spans about 34 months. Yet, the deviation of the prediction from the target is only 0.53 per cent. The prediction with 32 cases is the only one that exceeds the arbitrarily selected five per cent limit.

The predictive ability of the model developed with different numbers of cases appears to be stable. The range of the seven forecasts is only 0.052 hours per pound.

The predictive ability results for Model 9 (Table 22) are not quite as good. While the predictions center nicely

TABLE 21

PREDICTIVE ABILITY - FULL MODEL 8
(Assembly Hours: F-4)

Cases	Estimate B_0	Estimate B_1	Estimate B_2	Forecast Lot 57[a]	Per Cent Deviation
42	9.016	-0.279	-0.112	0.750	-0.67
40	9.042	-0.279	-0.112	0.752	-0.94
38	9.041	-0.279	-0.112	0.752	-0.01
36	9.111	-0.282	-0.105	0.745	0.00
34	9.060	-0.283	-0.097	0.741	0.53
32	8.812	-0.294	-0.041	0.700	6.04
30	8.601	-0.283	-0.069	0.726	2.55

[a]The observed value at lot 57 (case 42) is 0.745 assembly hours per pound.

TABLE 22

PREDICTIVE ABILITY - FULL MODEL 9
(Assembly Hours: F-4)

Cases	Estimate B_0	Estimate B_1	Estimate B_2	Forecast Lot 57[a]	Per Cent Deviation
42	10.406	-0.278	-0.097	0.776	-4.16
40	10.402	-0.277	-0.098	0.780	-4.70
38	10.365	-0.277	-0.099	0.775	-4.03
36	10.420	-0.281	-0.092	0.767	-2.95
34	10.142	-0.283	-0.081	0.755	-1.34
32	8.587	-0.294	-0.013	0.692	7.11
30	9.066	-0.283	-0.049	0.732	1.74

[a]The observed value for lot 57 (case 42) is 0.745.

on the target of 0.745 assembly hours per pound, they range over 0.098 hours per pound. This is almost twice the range of the predictions from Model 8 and indicates that Model 9 is less stable than Model 8.

Tests for the predictive ability of reduced Models 8 and 9 are reported in Table 23. All of the predictions are low and the magnitude of the deviations is larger than those for full Models 8 or 9. This reduced model is fairly stable with a prediction span of 0.056 hours per pound.

TABLE 23

PREDICTIVE ABILITY - REDUCED MODELS 8 AND 9
(Assembly Hours: F-4)

Cases	Estimate B_0	Estimate B_1	Forecast Lot 57[a]	Per Cent Deviation
42	5.852	-0.245	0.737	1.07
40	5.912	-0.247	0.732	1.74
38	5.984	-0.249	0.729	2.15
36	6.430	-0.259	0.711	4.56
34	6.996	-0.271	0.707	5.10
32	8.185	-0.294	0.681	8.59
30	7.931	-0.288	0.688	7.65

[a] The observed value for lot 57 (case 42) is 0.745 assembly hours per pound.

The production rate variable helps in the accuracy of prediction in both models and hypothesis three is accepted. But in comparing Models 8 and 9, Model 8 with the manufacturing rate representing the production rate is the better of the two. This correlates with the findings on all of the model pairs examined to this point.

Another interesting fact is that the fit of the model to the data set improves with an increasing percentage of fabrication hours in the dependent variable. For example, Models 8, 4 and 6 are identical except for the dependent variable. This variable is assembly, total and fabrication hours per pound respectively. The corresponding R^2 actual statistics are 0.797, 0.887 and 0.919. One can estimate the fabrication hours for Models 8, 4 and 6 as a percentage of total hours. This percentage is 0, 33 and 100 respectively. This increasing percentage of fabrication hours with increasing R^2 actual may indicate that prorating fabrication set up time is the most important of the reasons why the production rate is an explainer of variation in hours per pound.

THE F-102 PROGRAM

The F-102 airframe production data are extracted from the "F-102 Program Cost History" which is prepared by the manufacturer to assist in pricing, estimating and cost evaluation.[44] This history is a comprehensive publication containing information on many facets of the F-102 program. One is left with the impression that great care has been taken to assign accurately the cost elements to the proper category or end item.

For the purpose of this study, the data are limited in that lot release dates and hours for lower process levels by lot are not available. But total direct production hours by individual airframe are available. These data are matched with airframe delivery data to produce the variables needed to test the procedure in the cumulative production and production rate cost model.

Deliveries of 1000 aircraft are made during the years 1953 through 1958. Of these 1000, 889 are F-102A and 111 are TF-102A aircraft. The monthly delivery rate of the 1000 is somewhat irregular as reflected in column 3 of Table 24. For the first 26 months, the maximum monthly delivery rate is three. Over these 26 months,

---

[44] "F-102 Program Cost History," CRA-1-7, a report prepared by the Cost Research Department, Fort Worth Division, General Dynamics Corporation, June 1965.

TABLE 24

VARIABLES FOR ANALYSIS OF F-102A
TOTAL HOURS PER POUND

Month (1)	Cumulative Production Plot Point (2)	Monthly Delivery Rate (3)	Total Hours Per Pound (4)
1	1.	1	33.33
3	1.5	1	31.13
7	3.0	2	21.72
8	4.5	1	20.72
9	5.5	1	20.38
10	7.0	2	18.81
11	8.5	1	17.44
12	9.5	1	17.28
15	10.5	1	26.92
16	11.5	1	23.57
17	12.5	1	22.32
18	13.5	1	21.74
21	14.5	1	20.35
22	16.0	2	18.27
23	17.5	3	15.22
24	21.0	2	13.44
27	26.0	6	13.30
28	32.0	6	10.57
29	38.0	6	9.66
30	45.5	9	8.57
31	52.0	4	7.62
32	55.5	3	7.74
33	60.0	6	7.31
34	68.0	10	6.73
35	80.0	14	6.60
36	92.5	11	6.18
37	112.5	29	5.56
38	148.5	43	5.19
39	185.5	31	4.69
40	221.0	40	4.79
41	257.5	33	4.14
42	293.5	39	4.01
43	337.0	48	3.84
44	384.5	47	3.72
45	435.0	54	3.38
46	482.0	40	3.17
47	526.5	49	3.26

TABLE 24-Continued

Month (1)	Cumulative Production Plot Point (2)	Monthly Delivery Rate (3)	Total Hours Per Pound (4)
48	567.5	33	3.24
49	604.0	40	3.11
50	649.5	51	3.02
51	696.0	42	2.97
52	740.5	47	3.00
53	787.5	47	2.90
54	835.0	48	2.85
55	879.5	41	2.87
56	917.5	35	2.86
57	950.0	30	2.85
58	975.0	20	2.92
59	990.0	10	3.02
60	997.5	5	3.42

SOURCE: The data for constructing the variables in this table are extracted from the "F-102 Program Cost History," numbered CRA-1-7 and dated June 1965. The summary is developed from the joint accounting records of the General Dynamics Corporation Convair and Fort Worth Divisions and published by the Fort Worth Division. The section for construction of the dependent variable is entitled "F-102A and TF-102A Final Tabulation of Charted Hours and Manhours per Pound". It includes direct labor hour and DCPR data for each airframe produced. The delivery information used for constructing the independent variables is taken from a section entitled "F-102 Deliveries".

there are ten during which no airframes are delivered. Then the rate increases gradually over the next 24 months to peak at 51. The last ten months are characterized by a generally decreasing rate.

The cumulative production and production rate cost model, $y = B_0 \cdot x_1^{B_1} \cdot x_2^{B_2} \cdot 10^e$, is once more used to examine the data. In this case, construction of the variables is easier because the hourly data are collected by individual airframe.

The dependent variable is average direct production hours per pound for each month. It is constructed by calculating the arithmetic average of the hours required to build the F-102A airframes delivered in each month. This number is divided by the arithmetic average DCPR weight of those same airframes. Column 4 of Table 24 lists the dependent variable for the 50 months during which deliveries are made.

The first independent variable is the cumulative production plot point. It is calculated on a monthly basis for the program total of 1000 airframes even though the dependent variable is constructed for the 889 F-102A airframes. This is done because it is assumed that learning on all similar airframes contributes to the unit cost behavior of a particular airframe model. The cumulative production plot point is simply the middle value for monthly production plus the cumulative total production for the previous months. It is listed in column 2 of Table 24.

The second independent variable is the program monthly delivery rate as determined by the date of customer acceptance of the airframe. It is listed in column 3 of Table 24 for those months in which some delivery is made. The ten months during which there are no deliveries are omitted as cases for the regression analysis. As previously indicated, the raw data needed to construct a manufacturing rate variable are not available.

The F-102A data of Table 24 reflect the familiar toe-up phenomenon at the end of the program. For the last three months, the average hours per pound increase sharply due to the effects of program completion actions. This phenomenon is lacking in the F-4 data just examined because production of F-4 airframes continues through the most recent observations.

For the F-102A part of the investigation, two versions of the cumulative production and production rate cost model are examined using the test procedures. They are Model 10 with all 889 F-102A airframes aligned in 50 monthly cases and Model 11 with the last 879 F-102A airframes delivered over 42 months.

COST MODEL 10
F-102A TOTAL HOURS - 50 CASES

Table 25 lists the results of regression analysis of the F-102A data in the cost model. The statistics reflect that the delivery rate is an important explanatory variable in the cost model. The F^* incremental statistic of 38.39 is larger than the theoretical F value of 4.05. The R^2 statistic for the full model improves by 1.8 per cent over the same statistic for the reduced model. Similarly, the MSE of the full model is about one-half that of the reduced model.

The model also appears to fit the data well. The R^2 statistic of 0.979 indicates that 97.9 per cent of the variation in the logarithm of hours per pound is explained by the logarithm of the two independent variables. The R^2 actual statistic of 0.933 indicates that some 93 per cent of the observed variation in hours per pound is explained by the cost model. The F^* statistic of 1071.11 is much greater than the theoretical value of 3.20. This supports the theory of a relationship between the dependent variable and the set of independent variables.

An examination of the residuals shows a sharp discontinuity between cases eight and nine. At this point the hours per pound dependent variable increases abruptly and begins what appears to be a new trend. Perhaps there

TABLE 25

REGRESSION RESULTS - MODEL 10
(F-102A Total Hours - 50 Cases)

Estimate B_0 =	38.371	R^2 actual =	0.933
Estimate B_1 =	-0.299	t ratio B_1 =	14.95
Estimate B_2 =	-0.158	t ratio B_2 =	6.08
F^* (full) =	1070.11	F(2,47),0.05 =	3.20
F^* (incremental) =	38.39	F(1,47),0.05 =	4.05
MSE (full) =	0.0029	MSE (reduced) =	0.0051
R^2 (full) =	0.979	R^2 (reduced) =	0.961

Case	Observed Value	Predicted Value	Residual	Per Cent Deviation
1	33.33	38.37	-5.04	-15.12
2	31.13	33.99	-2.86	-9.19
3	21.72	24.76	-3.04	-14.00
4	20.72	24.48	-3.75	-18.15
5	20.38	23.05	-2.68	-13.10
6	18.81	19.22	-0.41	-2.18
7	17.44	20.24	-2.80	-16.06
8	17.28	19.58	-2.30	-13.31
9	26.92	19.00	7.92	29.42
10	23.57	18.49	5.08	21.55
11	22.32	18.04	4.28	19.18
12	21.74	17.63	4.11	18.91
13	20.35	17.25	3.10	15.23
14	18.27	15.01	3.25	17.84
15	15.22	13.71	1.51	9.92
16	13.44	13.84	-0.40	-2.98
17	13.30	10.91	2.39	17.97
18	10.57	10.25	0.31	2.93
19	9.66	9.74	-0.08	-0.83
20	8.57	8.66	-0.09	-1.05
21	7.62	9.46	-1.84	-24.15
22	7.74	9.71	-1.96	-25.32
23	7.31	8.50	-1.19	- 6.28
24	6.73	7.55	-0.82	-12.18
25	6.60	6.82	-0.22	-3.33
26	6.18	6.78	-0.60	-9.71

TABLE 25-Continued

Case	Observed Value	Predicted Value	Residual	Per Cent Deviation
27	5.56	5.49	0.07	1.26
28	5.19	4.75	0.44	8.48
29	4.69	4.68	0.01	0.21
30	4.79	4.26	0.53	11.06
31	4.14	4.20	-0.06	-1.45
32	4.01	3.93	0.08	2.00
33	3.84	3.65	0.19	4.95
34	3.72	3.52	0.20	5.38
35	3.38	3.32	0.06	1.78
36	3.17	3.38	-0.21	-6.62
37	3.26	3.18	0.08	2.45
38	3.24	3.31	-0.07	-2.16
39	3.11	3.16	-0.05	-1.61
40	3.02	2.97	0.05	1.66
41	2.97	3.00	-0.03	-1.01
42	3.00	2.89	0.11	3.67
43	2.90	2.84	0.06	2.07
44	2.85	2.78	0.07	2.46
45	2.87	2.81	0.06	2.09
46	2.86	2.84	0.02	0.70
47	2.85	2.88	-0.03	-1.05
48	2.92	3.05	-0.13	-4.45
49	3.02	3.39	-0.37	-12.25
50	3.42	3.78	-0.35	-10.23

is a process change or major design change introduced at this point. Except for the toe-up at the end, the rest of the residuals do not reflect abnormal distribution. Hypothesis one is therefore accepted for Model 10.

The predictive ability of the full Model 10 is examined in the data of Table 26. Since the purpose of the model is to predict follow-on production, month 57 (case 47) prior to the toe-up effect is chosen as the prediction target.

TABLE 26

PREDICTIVE ABILITY - FULL MODEL 10

Cases	Estimate B_0	Estimate B_1	Estimate B_2	Forecast Month 57[a]	Per Cent Deviation
50	38.371	-0.299	-0.158	2.88	-1.05
48	36.852	-0.273	-0.189	2.98	-4.56
46	36.171	-0.263	-0.200	3.02	-5.96
44	35.975	-0.260	-0.202	3.04	-6.67
42	35.933	-0.259	-0.202	3.06	-7.37
40	35.826	-0.258	-0.203	3.06	-7.37
38	35.655	-0.255	-0.205	3.09	-18.42
36	35.370	-0.252	-0.207	3.11	-9.21
34	35.034	-0.247	-0.209	3.16	-10.88

[a]The observed value for month 57 (case 47) is 2.85 total hours per pound.

The predictive ability of the procedure for this set of data deteriorates more rapidly than it did with similar F-4 data. As cases are removed from the original 50, the resulting models appear to systematically develop higher predictions. When predicting only 13 months into the future, the 34 case model predicts almost 11 per cent high. This exceeds the test limit and hypothesis three is rejected.

The results of the predictive ability tests for a reduced Model 10 are presented in Table 27. Since the reduced model has a high R^2 statistic of 0.961, one might expect the predictive ability of these reduced models to be good. They are not. For a prediction of only one month from the model developed with 46 cases, the prediction is 8.77 per cent low. Longer predictions with coefficients developed from fewer cases are increasingly poor.

When comparing the predictive results of the full model with those of the reduced model, one can conclude that the full model is better. For example, with 38 cases in the regression analysis, the full model predicts 8.4 per cent high while the reduced model predicts 11.6 per cent low. While neither prediction is particularly good for only a nine month span, the comparison shows that the delivery rate is an important additional variable in the model.

TABLE 27

PREDICTIVE ABILITY - REDUCED MODEL 10

Cases	Estimate B_0	Estimate B_1	Forecast Month 57	Per Cent Deviation
50	44.162	-0.408	2.69	5.61
48	44.883	-0.414	2.63	7.72
46	45.282	-0.417	2.60	8.77
44	45.623	-0.420	2.56	10.18
42	45.859	-0.521	2.55	10.53
40	46.083	-0.423	2.53	11.23
38	46.178	-0.424	2.52	11.58
36	46.321	-0.425	2.51	11.93
34	45.966	-0.422	2.55	10.53

The discontinuity in the dependent variable data between cases eight and nine suggests that a better model can be constructed from the last 42 cases of the 50 case F-102A data set. Model 11 is developed from these 42 cases. The results of the analysis are listed in Table 28.

Once again the statistics indicate that the delivery rate is an important explainer of variation in hours per pound. In Model 11, the R^2 statistic is increased by two per cent when adding the delivery rate to the reduced model. The F^* incremental statistic of 36.37 is larger than the theoretical F value of 4.09 at the 0.05 significance level.

TABLE 28

REGRESSION RESULTS - MODEL 11
(F-102A Total Hours - 42 Cases)

Estimate B_0	= 47.290	R^2 actual	= 0.955
Estimate B_1	= -0.344	t ratio B_1	= 16.44
Estimate B_2	= -0.144	t ratio B_2	= 6.03
F^* (full)	= 909.46	F(2,39),0.05	= 3.24
F^* (incremental)	= 36.37	F(2,39),0.05	= 4.09
MSE (full)	= 0.0022	MSE (reduced)	= 0.0041
R^2 (full)	= 0.979	R^2 (reduced)	= 0.959

The model appears to be appropriate for the data set. The R^2 actual statistic is 0.955 indicating that over 95 per cent of the variation in hours per pound is explained. The residuals indicate no important problems with the model. The F^* statistic is much larger than the critical value. Therefore, hypothesis one is sustained for Model 11.

Tests of the predictive ability of models estimated from the data set of Model 11 are also made. These predictions of month 57 are low in every test with generally larger per cent deviations than those from the data set of Model 10. Hypothesis three is also rejected for Model 11.

The stability of Model 11 is also not as good as that of Model 10. The range of nine forecasts using the data of Model 11 is 0.63 hours per pound. These forecasts span 16 months in steps of two. The corresponding forecasts for Model 10 range over only 0.27 hours per pound. As with the F-4 program data, it appears that the early observations are important in both locating and stabilizing the regression plane.

TABLE 29

PREDICTIVE ABILITY - FULL MODEL 11

Cases	Estimate B_0	Estimate B_1	Estimate B_2	Forecast Month 57[a]	Per Cent Deviation
42	47.290	-0.344	-0.144	2.75	3.51
40	45.994	-0.330	-0.158	2.80	1.75
38	46.017	-0.330	-0.158	2.80	1.75
36	47.263	-0.341	-0.149	2.74	3.51
34	51.254	-0.374	-0.125	2.58	9.47
32	48.961	-0.356	-0.138	2.69	5.61
30	53.964	-0.394	-0.110	2.49	12.63
28	59.428	-0.433	-0.083	2.30	19.30
26	64.186	-0.462	-0.065	2.17	23.86

[a]The observed value for month 57 (case 39) is 2.85 hours per pound.

THE KC-135A PROGRAM

The last program examined in this project is production of the KC-135A four engine tanker airframe. Deliveries from this program span the years 1957 through 1965. There are 820 aircraft in the C-135 family of which 732 are KC-135As. The balance of the aircraft are a mix of six other models produced during the last half of the program. The KC-135A model is selected for the analysis because it is produced from the beginning to the end of the program and accounts for the great majority of the direct labor hours spent in production.

Data for this analysis are derived from the historical records of the manufacturer. Two different data sets are constructed to facilitate the tests described in chapter four. One is based on 96 months of deliveries and the other is based on seven production lots. From these data, five models are developed and examined.

COST MODEL 12
TOTAL HOURS: KC-135A-96 MONTHS

Table 30 lists the variables needed to develop cost Model 12. The variables are constructed in the same manner as those for the F-102A. Deliveries per month are a count of all C-135 family airframes accepted in each month. The cumulative production plot point is the middle value of

TABLE 30

VARIABLES FOR ANALYSIS OF KC-135A TOTAL HOURS

Month	Cumulative Production Plot Point	Monthly Delivery Rate	Total Hours Per Pound
1	1.	1	9.82
2	1.5	1	7.59
3	2.5	1	7.31
4	4.0	2	7.36
5	6.5	3	7.62
6	10.0	4	6.05
7	14.0	4	5.69
8	16.5	1	5.52
9	19.0	4	5.29
10	23.5	5	5.22
11	28.0	4	5.15
12	33.5	7	4.94
13	41.0	8	4.22
14	50.0	10	3.68
15	60.5	11	3.16
16	72.0	12	2.75
17	84.5	13	2.46
18	98.5	14	2.34
19	112.5	15	2.15
20	127.5	15	2.04
21	142.5	15	2.01
22	157.5	15	1.99
23	172.5	15	1.96
24	187.5	15	1.91
25	202.5	15	1.86
26	217.5	15	1.82
27	232.5	15	1.79
28	247.5	15	1.75
29	262.5	15	1.60
30	277.5	15	1.53
31	292.5	15	1.47
32	306.0	12	1.41
33	322.0	10	1.36
34	326.5	9	1.31
35	335.0	8	1.31
36	343.0	8	1.29
37	351.0	8	1.24
38	359.0	8	1.22

TABLE 30-Continued

Month	Cumulative Production Plot Point	Monthly Delivery Rate	Total Hours Per Pound
39	366.5	7	1.18
40	373.5	7	1.17
41	380.5	7	1.16
42	388.0	8	1.13
43	395.5	7	1.12
44	402.5	7	1.11
45	409.5	7	1.09
46	417.0	8	1.06
47	424.5	7	1.06
48	432.0	6	1.06
49	438.0	6	1.05
50	444.0	6	1.05
51	450.0	6	1.05
52	456.0	6	1.04
53	462.0	6	1.03
54	468.0	6	1.03
55	472.5	5	1.02
56	478.0	6	1.01
57	484.0	6	1.00
58	490.0	6	1.00
59	497.0	8	1.01
60	504.5	7	1.01
61	511.5	7	1.01
62	518.5	7	1.00
63	527.0	10	1.06
64	537.0	10	1.09
65	548.0	12	1.11
66	560.0	12	1.10
67	569.5	9	1.09
68	579.0	8	1.06
69	587.0	8	1.04
70	596.5	11	1.03
71	604.5	5	1.04
72	611.0	8	1.02
73	619.0	8	1.01
74	627.5	9	1.01
75	635.5	7	1.04
76	543.0	8	1.07
77	651.0	8	1.00
78	659.0	8	1.00
79	667.0	8	1.00
80	675.0	8	0.98

TABLE 30-Continued

Month	Cumulative Production Plot Point	Monthly Delivery Rate	Total Hours Per Pound
81	683.0	8	0.98
82	691.0	8	0.99
83	699.0	8	1.00
84	707.5	9	1.04
85	715.5	7	1.02
86	722.5	7	1.02
87	729.5	7	1.03
88	736.5	7	1.02
89	744.0	8	1.01
90	752.5	9	0.98
91	762.5	11	0.97
92	773.0	10	0.95
93	782.0	8	0.95
94	790.5	9	0.92
95	799.0	8	0.90
96	803.5	7	0.89

SOURCES: A summary of important dates for each airframe is contained in a series of charts entitled "Military Program Actuals" prepared by the Program Planning Unit, Airplane Division, The Boeing Company. Each report in the series bears the chart number 1-135-25. The charts are dated 1 April 1960 (8 pages), 30 June 1962 (2 pages), 31 March 1964 (2 pages) and 31 October 1966 (2 pages). These charts are the source of the monthly delivery rate and the cumulative production plot point.

The total hours per pound per month are read and in some cases estimated from a summary chart entitled "Manhours Per Pound KC-135" and dated 30 September 1965. This chart is numbered 4-135-7, consists of two pages and is prepared by the Manufacturing Department, Airplane Division, The Boeing Company. Although data for the first 20 and the last 13 airframes are reported individually, some of the intermediate airframes are reported in steps of five or ten. Therefore it is sometimes necessary to estimate the manhours per pound for the average airframe delivered in a particular month.

C-135 family acceptances in each month plus the cumulative total of prior acceptances. Total hours per pound are for the KC-135A airframe only. They represent the average total direct production labor hours divided by the average DCPR weight for the airframes accepted in each month.

The data reflect an early peak production rate of 15 airframes per month, a decline to five per month followed by a second peak of 12 per month. After another decline to seven per month, there is a final peak of 11 per month.

The data needed to develop fabrication and assembly hour variables to match this total hour data set are not available. Also, a manufacturing rate variable cannot be produced from the available data. Therefore, only one model is examined with this data set.

The results of the regression analysis are listed in Table 31. All of the statistics indicate that the model fits the data set well. The F^* full statistic is much larger than the theoretical F value of 3.10 indicating that the set of independent variables is related to the dependent variable. The R^2 statistic (in logarithms) reflects that 97 per cent of the variance in the dependent variable is captured by the model. The R^2 actual statistic, showing explained variance after transformation back to the original variable form, indicates that 89 per cent of the variance in hours per pound is explained by the model.

TABLE 31

REGRESSION RESULTS: MODEL 12
(KC-135A Total Hours-96 Cases)

Estimate B_0	=	13.133	R^2 actual	=	0.890
Estimate B_1	=	-0.453	t ratio B_1	=	47.60
Estimate B_2	=	0.164	t ratio B_2	=	6.56
F^* (full)	=	1558.37	F(2,93),0.05	=	3.10
F^* (incremental)	=	42.92	F(1,93),0.05	=	3.95
MSE (full)	=	0.0022	MSE (reduced)	=	0.0032
R^2 (full)	=	0.971	R^2 (reduced)	=	0.958

The delivery rate is also shown to be an important explanatory variable. The F^* incremental statistic of 42.92 is larger than the critical F value of 3.95 at the 0.05 significance level. The R^2 statistic shows an improvement of 1.3 per cent after adding the rate variable to the reduced model. Similarly, the MSE statistic shows a one-third reduction due to inclusion of the rate variable in the model.

The fact that the R^2 statistic improves by only 1.3 per cent after adding the rate term to the reduced model raises doubt about the practicality of including the rate term in the KC-135A model. Further analysis of the variance eases these doubts. The $SSE(x_1)$ for the 96 cases is 0.301. It represents the unexplained variance (in logarithms) when only cumulative production is in the model. $SSE(x_1, x_2)$ is 0.206 and represents the unexplained variance with both cumulative production and delivery rate in the model. The reduction in unexplained variance due to adding the rate variable when divided by $SSE(x_1)$ is the partial coefficient of determination. It is (0.301 - 0.206)/0.301 = 0.316.[45] From this analysis, one can conclude that about one-third of the unexplained variance from the reduced model is accounted for by adding the rate variable.

One unusual result of the KC-135A analysis is that the estimate of the B_2 coefficient in the full model is positive. If one assumes that an increase in production rate causes a decrease in unit labor requirements, the expected sign of the B_2 coefficient is negative. The unexpected result can be explained as a twist in the regression plane brought about by the difference in relative strength of the independent variables and their collinearity.

The reduced form of the full model provides a clue to this twisting effect cause. The estimate of the B_1 coefficient with the x_2 variable removed is -0.415. The estimate of the B_2 coefficient with the x_1 variable removed

---

[45] Neter and Wasserman, Statistical Models, p. 411.

is -0.573. Yet when both the x_1 and x_2 variables are in the full model, the B_2 sign becomes positive. In their text on statistical analysis, Spurr and Bonini attribute this type of behavior to collinearity among the independent variables.[46]

The x_1 and x_2 variables are somewhat correlated in this KC-135A case. The coefficient of correlation is 0.619. This statistic must be evaluated with some caution because as production rate declines, cumulative production continues to increase until the rate reaches zero.

The x_1 variable appears to be much stronger than the x_2 variable when explaining the variance of the dependent variable. For example, the R^2 statistic of the reduced model with x_2 removed is 0.958. The R^2 statistic of a model with x_1 removed is only 0.265.

This combination of collinearity and a large difference in relative strength of the independent variables causes the regression plane to tilt and the B_2 coefficient to become positive. There does not appear to be any theoretical significance to this change in sign and the model should be useful for predictive purposes.

Subjective analysis of the model behavior and the

---

[46] William A. Spurr and Charles P. Bonini, Statistical Analysis for Business Decisions (Homewood, Illinois: Richard D. Irwin, Inc., 1967), p. 611.

residuals indicates no serious departures from the model assumptions. Accordingly, hypothesis one is accepted for the 96 case KC-135A data set.

The predictive ability of the model is also examined. Results of the tests are listed in Tables 32 and 33. Parallel forecasts for the full and reduced models are made in steps of 12 months.

Both model forms forecast fairly well. The full model form forecasts low in every case with a range of only 0.04 hours per pound over the five forecasts. The reduced model form predicts high in every case with a range of 0.05 hours per pound. One cannot say that adding the production rate improves the predictive ability of the full model over the reduced model very much. One can accept hypothesis three for data set of Model 12.

The data for Model 12 are not adaptable to evaluating the different levels of the KC-135A manufacturing process. Therefore a different data set is constructed to evaluate hypothesis two for the KC-135A program.

The 820 airframes in the C-135 family include seven different airframe models. They are 732 KC-135A, 17 EC-135C, 4 RC-135A, 10 RC-135B, 15 C-135A, 30 C-135B and 12 C-135F aircraft. Labor requirement data are collected by the manufacturer at different process levels for these airframes by lot. These data are identified

TABLE 32

PREDICTIVE ABILITY - FULL MODEL 12

Cases	Estimate B_0	Estimate B_1	Estimate B_2	Forecast Month 96[a]	Per Cent Deviation
96	13.133	-0.453	0.164	0.87	2.2
84	13.244	-0.461	0.176	0.85	4.5
72	13.373	-0.468	0.186	0.84	5.6
60	13.421	-0.473	0.195	0.83	6.7
48	13.384	-0.468	0.186	0.84	5.6

[a]The observed value for month 96 is 0.89 hours per pound.

TABLE 33

PREDICTIVE ABILITY - REDUCED MODEL 12

Cases	Estimate B_0	Estimate B_1	Forecast Month 96[a]	Per Cent Deviation
96	14.782	-0.415	0.92	-3.3
84	14.891	-0.417	0.92	-3.3
72	14.945	-0.418	0.91	-2.2
60	14.917	-0.417	0.92	-3.3
48	14.225	-0.401	0.97	-9.0

[a]The observed value for month 96 is 0.89 hours per pound.

to a particular model through lot seven but are aggregated for all airframes for lots eight through ten. Therefore only seven lots of data are available for the KC-135A analysis. Table 34 lists the data needed to evaluate Models 13 through 16.

The cumulative production plot point is constructed by adding one-half the lot size to the cumulative total number of airframes produced in prior lots. In accordance with learning curve technique, the plot value for the first lot is adjusted to compensate for the rapid change in the slope of the cumulative production curve for the first few units. A plot point value of 10.14 is read from the tables for a first lot size of 29 airframes with a slope of 78 per cent.[47]

The lot average manufacturing rate is constructed by dividing the number of airframes in a lot by the time span in months required to manufacture those airframes. This span bridges the time the first batch of parts in the lot is released to fabrication until the last airframe in the lot is accepted. For this program, the first month that any fabrication hours are recorded for each lot is used as a proxy for the lot release date.

The lot average delivery rate is constructed by

---

[47]Boren and Campbell, Learning Curve Tables, p. 162.

TABLE 34

VARIABLES FOR ANALYSIS OF KC-135A FABRICATION AND ASSEMBLY HOURS

Lot	Cumulative Production Plot Point	Lot Average Monthly Manu-facturing Rate	Lot Average Monthly Delivery Rate	Fabri-cation Hours Per Pound	Assembly Hours Per Pound
1	10.14	0.81	3.63	0.509	2.933
2	63.00	2.52	9.71	0.235	1.655
3	156.00	4.21	14.75	0.186	0.978
4	280.00	4.33	13.00	0.191	0.722
5	385.50	3.00	7.36	0.179	0.511
6	459.00	2.44	5.50	0.188	0.439
7	534.50	3.27	6.54	0.161	0.471

SOURCES: Lot composition data is obtained from a report entitled "Major Production Airplanes-Military" prepared by Industrial Engineering, Airplane Division, The Boeing Company and dated 3 January 1966.

The months during which the first fabrication hours are expended are extracted from a machine listing of the KC-135 Program data. The data bank is maintained by Production Planning, Industrial Engineering, 707/727/737 Division of the Boeing Commercial Airplane Company.

Delivery data needed to construct the two forms of the production rate are extracted from the "Military Program Actuals" series of reports more completely described in Table 30.

Fabrication and assembly hours by lot are read from a three page report numbered 132.3115 and entitled "Production Hours, Model KC-135, By Work Order By Control Code." It is prepared by Program Planning, Industrial Engineering, Airplane Division, The Boeing Company and dated March 1968.

DCPR weights for constructing the dependent variables are estimated from a Boeing Company summary report entitled "Manufacturing On and Off-Site Direct Manhours by Lot." This one page report is dated 1 March 1965.

dividing the number of airframes in a lot by the time span in months from the first to last acceptance. This span is calculated directly from the actual acceptance dates which are available in the data.

The fabrication and assembly hours per pound variables are constructed by dividing the total hours required to fabricate or assemble the parts for each lot by the number of airframes in the lot and the lot DCPR weight. The assembly hours are the sum of two reporting codes used by the manufacturer, minor and major assembly.

The cumulative production and production rate cost model form is used to examine four data sets listed in Table 34. The fabrication and assembly hours per pound are used as alternate dependent variables while the independent variables are the cumulative production plot point and one of the two forms of the production rate.

It is unfortunate that the lots are so large and are released so infrequently (about once a year). More observations would permit a more precise evaluation of the rate effect on labor requirements. Nevertheless, after working with many different data sources for the KC-135A, one is left with the impression that the raw data for these seven cases are carefully collected and allocated.

COST MODELS 13 AND 14
KC-135A FABRICATION HOURS PER POUND

Cost Model 13 examines fabrication hours per pound as a function of cumulative production and manufacturing rate. The regression analysis results are presented in Table 35.

TABLE 35

REGRESSION RESULTS-MODEL 13

Estimate B_0	= 0.674	R^2 actual	= 0.986
Estimate B_1	= -0.165	t ratio B_1	= 4.47
Estimate B_2	= -0.305	t ratio B_2	= 3.30
F^* (full)	= 74.51	F(2,4),0.05	= 6.94
F^* (incremental)	= 10.88	F(1,4),0.05	= 7.71
MSE (full)	= 0.0011	MSE (reduced)	= 0.0034
R^2 (full)	= 0.974	R^2 (reduced)	= 0.903

Case	Observed Value	Predicted Value	Residual	Per Cent Deviation
1	0.509	0.491	0.018	3.55
2	0.235	0.257	-0.022	-9.40
3	0.186	0.189	-0.003	-1.80
4	0.191	0.170	0.021	10.74
5	0.179	0.181	-0.002	-1.04
6	0.188	0.187	0.001	0.45
7	0.161	0.167	-0.006	-3.69

Although only seven observations are used in the analysis, the statistics support the procedure and the theory. The F^* statistic is larger than its critical value indicating that the set of independent variables is related to the dependent variable. The F^* incremental statistic is also larger than its critical value. This indicates that the manufacturing rate variable explains additional variation in the fabrication hours per pound variable in the presence of the cumulative production variable. Both of these conclusions are supportable at the 0.05 level of significance.

The following observations indicate that the model is appropriate for the data. The residuals do not exhibit behavior that indicates violation of model assumptions. The R^2 actual statistic of 0.986 indicates that almost 99 per cent of the variance in the fabrication hours per pound is explained by the model. Both the MSE and R^2 regression statistics improve substantially upon addition of the manufacturing rate variable to the reduced model.

It is interesting to note that the sign of the B_2 coefficient estimate does not become positive as it does in Model 12. This fact supports the observation made in the F-4 analysis that fabrication hours are more sensitive to rate than assembly or total hours. Model 13 with the x_1 variable excluded has a negative B_2 coefficient estimate and an R^2 statistic of 0.843. With the x_2 variable excluded, the B_1 coefficient is negative and the R^2 statistic is 0.903. The cumulative production variable is not enough stronger than the rate variable to cause a sign change as in Model 12.

The analysis indicates that hypothesis two is acceptable. Cost Model 13 is appropriate to explain variation in KC-135A fabrication hours per pound as a function of cumulative production and production rate for the first seven lots.

The test of hypothesis three, predictive ability, is impractical in this and the next three models because there are only seven observations. Removal of observations for further analysis leaves insufficient degrees of freedom to attach much importance to the results of a predictive ability test.

Cost Model 14 examines KC-135A fabrication hours per pound as a function of cumulative production and delivery rate. The only difference between Models 13 and 14 is that 13 uses the manufacturing rate and 14 uses the delivery rate form of the production rate. The regression results of Model 14 are listed in Table 36.

As one would anticipate, the regression results are similar to those of Model 13. The objective and subjective tests of the results do not indicate rejection and hypothesis two is also accepted for Model 14.

TABLE 36

REGRESSION RESULTS-MODEL 14

Estimate B_0	= 1.123	R^2 actual	= 0.985
Estimate B_1	= -0.233	t ratio B_1	= 9.18
Estimate B_2	= -0.222	t ratio B_2	= 3.03
F^* (full)	= 65.90	F(2,4),0.05	= 6.94
F^* (incremental)	= 9.20	F(1,4),0.05	= 7.71
MSE (full)	= 0.0013	MSE (reduced)	= 0.0034
R^2 (full)	= 0.971	R^2 (reduced)	= 0.903

Case	Observed Value	Predicted Value	Residual	Per Cent Deviation
1	0.509	0.491	0.018	3.57
2	0.235	0.257	-0.022	-9.57
3	0.186	0.190	-0.004	-2.09
4	0.191	0.170	0.021	10.80
5	0.179	0.179	-0.000	-0.25
6	0.188	0.184	0.004	2.22
7	0.161	0.171	-0.010	-6.03

As is the case with the F-4 data, comparison of the statistics of Models 13 and 14 leaves an indication that the manufacturing rate form of the production rate is better than the delivery rate. For example, the F^* incremental statistic is 10.88 for Model 13 as compared to 9.20 for Model 14. Similarly, the MSE and R^2 statistics are slightly better for Model 13. Again, there is insufficient evidence to conclude that one is clearly better than the other.

COST MODELS 15 AND 16
KC-135A ASSEMBLY HOURS PER POUND

Model 15 examines assembly hours per pound as a function of cumulative production and manufacturing rate. The results of regression analysis of the data in the cumulative production and production rate cost model are presented in Table 37.

TABLE 37

REGRESSION RESULTS-MODEL 15

Estimate B_0	= 13.338	R^2 actual	= 0.993
Estimate B_1	= -0.608	t ratio B_1	= 18.77
Estimate B_2	= 0.361	t ratio B_2	= 4.44
F^* (full)	= 327.85	F(2,4),0.05	= 6.94
F^* (incremental)	= 19.74	F(1,4),0.05	= 7.71
MSE (full)	= 0.0009	MSE (reduced)	= 0.0042
R^2 (full)	= 0.994	R^2 (reduced)	= 0.964

Case	Observed Value	Predicted Value	Residual	Per Cent Deviation
1	2.933	3.020	-0.087	-2.97
2	1.655	1.498	0.157	9.51
3	0.978	1.038	-0.060	-6.16
4	0.722	0.735	-0.013	-1.77
5	0.511	0.530	-0.019	-3.68
6	0.439	0.442	-0.003	-0.73
7	0.471	0.448	0.023	4.88

The statistics show that the null assumption for the three subhypotheses may be rejected. The F^* full statistic of 327.85 is larger than the theoretical F value of 6.94. The F^* incremental statistic of 19.74 is larger than the theoretical F value of 7.71. A plot of the residuals against the test variables does not indicate that the model should be rejected.

The subjective tests of the statistics also reflect a good fit and the importance of the rate variable. The R^2 actual statistic of 0.993 shows that most of the variance is explained. The improvement in the R^2 statistic from 0.964 to 0.994 due to adding the rate variable to the reduced model emphasizes the importance of the manufacturing rate. Similarly, adding the manufacturing rate variable to the reduced model decreases the MSE statistic by more than 75 per cent.

The behavior of the B_2 coefficient estimate is similar to that demonstrated in the analysis of Model 12. The sign of the coefficient changes from negative to positive when cumulative production is added to a reduced model. The reduced model with x_1 excluded has a B_2 estimate of -0.855 and an R^2 statistic of 0.460. A reduced model with x_2 excluded has a B_1 estimate of -0.494 and an R^2 statistic of 0.964. In the full model, the B_2 estimate becomes positive. This again indicates the effect of the collinearity of x_1 and x_2 and the large difference in the relative explanatory strength of these two variables. There does not appear to be any theoretical significance to the change. The full model should be useful for prediction.

None of the tests indicate that Model 15 is unsuitable for the data set. Therefore hypothesis two is accepted.

Model 16 is the last model examined in this study and is the companion to Model 15. It is used to evaluate assembly hours per pound as a function of cumulative production and delivery rate. The results of the regression analysis are listed in Table 38.

TABLE 38

REGRESSION RESULTS-MODEL 16

Estimate B_0	=	7.303	R^2 actual	=	0.992
Estimate B_1	=	-0.527	t ratio B_1	=	22.28
Estimate B_2	=	0.263	t ratio B_2	=	3.84
F^* (full)	=	259.14	F(2,4),0.05	=	6.94
F^* (incremental)	=	14.80	F(1,4),0.05	=	7.71
MSE (full)	=	0.0011	MSE (reduced)	=	0.0042
R^2 (full)	=	0.992	R^2 (reduced)	=	0.964

These results are similar to those of Model 15. All the statistical tests indicate a good fit of the model to the data. Therefore hypothesis two is accepted for Model 16.

A comparison of the statistics produced by Models 15 and 16 indicates 15 to be better. Since the only difference in the models is the production rate proxy, once again one can conclude that manufacturing rate is the better of the two in this particular application.

SUMMARY

In this chapter sixteen different sets of data are examined in the cumulative production and production rate cost model. There are nine from the F-4 program, two from the F-102 program and five from the KC-135 program.

With the F-4 program data, all three hypotheses are tested and accepted. The production rate is found to be an important explainer of variation in labor requirements at the three levels of production when expressed in the cost model. Furthermore, the rate variable in the presence of the cumulative production variable generally improves the ability of the model to predict as compared to the reduced model. It also appears to improve the stability of the model as predictions are made over longer spans of time.

In the F-102 program, only the first and third hypotheses are evaluated as the labor requirement data for lower process levels are not available. As in the F-4 data, the production rate is found to be an important explainer of total labor requirements. Again, the tests indicate that using the full model generally improves the capability to predict future requirements over that of the reduced model although the accuracy of the predictions are outside the test limits.

The analyses of the KC-135A data also validate the model and the theory. The production rate variable generally improves the ability to explain variance in the labor requirement at the three levels of production. Tests of improvement in the predictive ability are inconclusive for the model with total hours per pound. The tests for predictive ability are not made at lower process levels.

Results of the F-4 and KC-135 tests indicate that the fabrication hours per pound variable is more sensitive to the production rate than are the total or assembly hours. However, the production rate variable improves the explanatory power of all three levels of the process that are investigated.

Although no statistical proof exists, the manufacturing rate appears to be a better proxy for the production rate than the delivery rate. This observation is based on the relative strength of statistics produced from models with only the rate variable changed.

CHAPTER VI

SUMMARY AND CONCLUSIONS

There are many instances where the rate of production in an airframe manufacturing program is changed due to forces external to the production process. These changes are accompanied by a need for managers to know the effects on cost. The few methods that exist to estimate these effects are based on a limited number of cases and produce conflicting results. The purpose of this research is to develop and test a procedure to consider the effect of a production rate change on the direct production labor requirements to produce additional airframes.

The literature on the subject is inconclusive. Some writers believe that the change in production rate is an important predictor of variation in unit cost while others conclude that it is insignificant. Some writers indicate that an increase in production rate leads to an increase in unit cost and others believe the opposite effect occurs. This study shows through an empirical evaluation that the production rate can be an important predictor of variation in unit direct labor requirements. Furthermore, an increase in rate up to plant capacity can lead to a decrease in unit labor requirements.

One primary element in a cost estimate is the quantity of direct production labor hours. Therefore, a logical place to start the estimates of the effects of a changed rate of production on airframe production cost is with these hours.

Many estimators of direct labor hours use the learning curve cost model. But this model is driven only by the cumulative number of airframes produced, not the rate of production. Thus, even if the production rate is changed due to some requirement external to the process, the number of labor hours predicted by an unadjusted learning curve model to produce each unit is not changed.

This result is contrary to logic. Up to some limit that is associated with plant capacity, an increase in production rate should result in a decrease in unit labor requirements. Greater spreading of fabrication set up charges, increased efficiencies resulting from labor specialization and a more highly motivated worker all contribute to this expected result. For the same reasons, a production rate decrease should result in a unit labor requirement increase.

A cost model is proposed to capture this rate effect in the presence of the learning effect. It is $y = B_0 \cdot x_1^{B1} \cdot x_2^{B2} \cdot 10^e$ where:

y represents the direct production labor hours per DCPR pound of airframe in each lot,

x_1 represents the cumulative production plot point as in the learning curve model and

x_2 represents the production rate.

The dependent variable is expressed as the lot average direct labor hours per DCPR pound of airframe. When the data permit, the effects are examined at three levels of aggregation, total hours, assembly hours and fabrication hours.

The first independent variable is identical to that used in learning curve analysis. The cumulative production plot point is developed for all airframes of the same type produced in the plant. This procedure holds whether or not the dependent variable is developed from data for a single model or aggregated from a number of models.

Two variables are evaluated as representatives of the second independent variable. They are a manufacturing rate and a delivery rate. The manufacturing rate is the total number of airframes in a lot divided by the lot manufacturing time span. This span is bounded by the date the first airframe in a lot is released to fabrication until the last airframe in a lot is accepted. The delivery rate is the total number of airframes in a lot divided by the time span bounded by the first and last lot airframe acceptances.

Sixteen data sets are constructed from data for three airframe production programs. They include nine from the F-4 program, two from the F-102 program and five from the KC-135 program. The data sets are examined in the cumulative production and production rate cost model through regression analysis.

The estimating procedure proposed here consists of three major steps. First, the data are analyzed to make sure the model is appropriate to the program. Then if the results of that analysis are favorable, the predictive ability and stability of the model are tested. Again assuming a favorable outcome, in a practical application a third step would be to forecast labor requirements for additional production using the cost model developed from the analysis. Each step is discussed in turn along with associated conclusions from the research.

To linearize the model and facilitate the regression technique, each variable is transformed to logarithms so that: $\log y = \log B_0 + B_1 \log x_1 + B_2 \log x_2 + e$. Then each data set is analyzed in the regression model producing estimated coefficients and statistics. These become the basis for answering questions about the process. Findings are summarized in paragraphs that follow.

The cumulative production and production rate cost model fits the individual data sets well. In each of the sixteen sets of data the model is found to be appropriate.

Statistical tests to support this conclusion are conducted at the 0.05 level of significance. More specifically, one indicator of goodness of fit is the R^2 statistic for the full model. As reflected in Table 39 which summarizes some regression results for all 16 models, the R^2 full statistic is respectable in each case.

The production rate variable contributes importantly to the explanatory ability of the model. A statistical test of this conclusion is performed and accepted at the 0.05 level of significance for each model. A more intuitively appealing test is the improvement in the R^2 statistic when the production rate is added to a reduced model. Table 39 lists the R^2 full and R^2 reduced statistics side by side for each model. In each case the R^2 statistic is improved by an amount that indicates the production rate is a valuable contributor to the explanatory ability of the model.

Within some upper boundary related to plant capacity, production rate and unit labor requirements move in opposite directions with a cause and effect relationship. In every model examined, the rate variable is negatively correlated with unit direct labor requirements. This negative correlation is also reflected in the negative sign for the B_2 coefficient estimate in 13 of the 16 models (Table 39). In three KC-135A models, the combined effects

TABLE 39

REGRESSION MODEL SUMMARY

Cost Model	Airframe Model	Cases	Level	Rate	R^2 full	R^2 reduced	B_0	B_1	B_2
1	F-4A-F	57	Total	Del	0.978	0.928	masked	-0.261	-0.169
2	F-4B-F	55	Total	Manu	0.973	0.904	"	-0.246	-0.183
3	F-4B-F	55	Total	Del	0.966	0.904	"	-0.257	-0.161
4	F-4B-F	42	Total	Manu	0.853	0.585	"	-0.230	-0.157
5	F-4B-F	42	Total	Del	0.820	0.585	"	-0.229	-0.136
6	F-4B-F	42	Fabri	Manu	0.889	0.618	6.328	-0.221	-0.148
7	F-4B-F	42	Fabri	Del	0.851	0.618	7.601	-0.219	-0.127
8	F-4B-F	42	Assem	Manu	0.744	0.658	9.016	-0.279	-0.112
9	F-4B-F	42	Assem	Del	0.733	0.658	10.400	-0.278	-0.097
10	F-102A	50	Total	Del	0.979	0.961	38.371	-0.299	-0.158
11	F-102A	42	Total	Del	0.979	0.959	47.290	-0.344	-0.144
12	KC-135A	96	Total	Del	0.958	0.971	13.133	-0.453	0.164
13	KC-135A	7	Fabri	Manu	0.974	0.903	0.674	-0.165	-0.305
14	KC-135A	7	Fabri	Del	0.971	0.903	1.123	-0.233	-0.222
15	KC-135A	7	Assem	Manu	0.994	0.964	13.338	-0.608	0.361
16	KC-135A	7	Assem	Del	0.992	0.964	7.303	-0.527	0.263

of collinearity and great relative strength of the cumulative production variable causes the B_2 estimate sign to change to positive in the full model.

When comparing the effectiveness of a manufacturing or delivery rate representative of the production rate, the manufacturing rate gives better results in all six of the direct comparisons. But the difference is not great and either proxy is an important contributor to the explanatory power of the model.

The model fits fabrication hours per pound data better than assembly or total data for the F-4 and KC-135 programs. This may indicate that the fabrication labor requirements are more sensitive to rate than the other two levels of the process examined. Fabrication and assembly hours are not evaluated for the F-102 program.

This procedure does not appear to be suited to constructing general cost models with coefficients that are applicable to other programs. The wide variation in the coefficients listed in Table 39 for models with common production levels and rates suggests that any averaging of coefficients would lead to unreliable results. The procedure appears to be suited only for predicting additional production for continuing programs.

When practical, the ability of a model to predict accurately beyond the data is tested. This is accomplished by estimating the model coefficients with successively smaller data sets formed by truncating the most recent observations. Omitted observations are predicted with the new model as if they are unknown. These forecasts are then compared with the known values and those forecast from a learning curve model to measure the predictive ability of the full cost model.

The rate variable stabilizes and improves the predictive ability of the cost model for the F-4 and F-102 program data. This improvement is particularly marked for the nine F-4 models analyzed. The predictive accuracy of the models developed from the F-4 data sets make them attractive alternatives for estimating labor requirements for additional production. Although the model fit for the F-102A data is excellent, it does not forecast the prediction target very well. Tests for predictive ability improvement are either inconclusive or impractical for the KC-135 program data.

The procedure is well suited to forecasting the direct labor requirements for additional production lots. The steps for such a forecast are reviewed here. After the coefficients are estimated and the cost model is tailored to a particular program, one first must decide on the acceptability of the model using an approach similar to that described in Chapter 5. If the model fits, its predictive ability should be tested to gain some measure of confidence in a forecast. Assuming that the results of this test are acceptable, one must calculate the cumulative production plot point and the delivery or manufacturing rate for the unknown production lot or lots. Then using the tailored cost model, the unknown direct labor requirement for the new lot can be predicted.

The conclusions described above are necessarily limited to the data sets examined. But there is great temptation to generalize the results to other airframe production. In particular, the notion that a production rate variable correlates negatively and importantly with direct labor requirements has much application in the airframe industry if universally true.

An obvious extension of this research effort is to duplicate the procedure on additional programs. It would be of particular value to examine the procedure on new airframe production to discover if changes in manufacturing technology are affecting the process. The behavior of other cost elements with respect to production rate changes is also an important research topic.

APPENDIX A

```
C THE CUMULATIVE PRODUCTION AND PRODUCTION RATE COST MODEL
C PROGRAMMER IS LARRY L. SMITH - MARCH 1976.
C
C THIS PROGRAM IS TAILORED TO EVALUATE VARIATION IN UNIT DIRECT LABOR
C REQUIREMENTS AS A FUNCTION OF CUMULATIVE PRODUCTION AND PRODUCTION RATE.
C THE COST MODEL IS Y = B0 * (X1**B1) * (X2**B2) * (10.**E).
C Y IS THE UNIT OR AVERAGE DIRECT MAN HOURS PER POUND.  X1 IS THE
C CUMULATIVE PRODUCTION PLOT POINT AS IN THE LEARNING CURVE MODEL.
C X2 IS THE PRODUCTION RATE PROXY SUCH AS DELIVERIES PER MONTH.
C E REPRESENTS THE ERROR TERM.  CASES ARE USUALLY DISTINGUISHED BY
C PRODUCTION LOT OR MONTHLY PRODUCTION.  EACH TERM IN THE MODEL IS
C  TRANSFORMED TO LOGARITHMS TO FACILITATE REGRESSION ANALYSIS.
C
C THE PROGRAM FIRST TRANSFORMS THE DATA TO COMMON LOGARITHMS.
C REGRESSION RESULTS FOR A TWO VARIABLE MODEL WITHOUT THE PRODUCTION RATE
C ARE CALCULATED.  THIS MODEL IS EQUIVALENT TO THE UNIT LEARNING CURVE
C MODEL.  THEN THE RESULTS FOR A TWO VARIABLE MODEL WITHOUT THE CUMULATIVE
C PRODUCTION VARIABLE ARE CALCULATED.  THE STATISTICS ARE OF TWO TYPES.
C FIRST, STANDARD REGRESSION STATISTICS ARE CALCULATED.  THEY ARE IN
C LOGARITHMS BECAUSE OF THE TRANSFORMED DATA AND MODEL.  IN ADDITION,
C ACTUAL RESIDUALS BASED ON OBSERVATIONS IN THEIR ORIGINAL FORM ARE
C CALCULATED.  R**2 ACTUAL AND MSE ACTUAL STATISTICS ARE CALCULATED
C BASED ON THESE RESIDUALS.  FOR THE FULL MODEL, THE OBSERVED AND
C PREDICTED VALUES FOR EACH CASE ARE PRINTED OUT IN BOTH THEIR ORIGINAL
C FORM AND IN LOGARITHMS.  RESIDUALS AND PER CENT DEVIATION FROM
C THE OBSERVED VALUE ARE ALSO PRINTED.
C
C DATA ARE INPUT AS FOLLOWS:
C THE FIRST CARD IS THE NUMBER OF CASES RIGHT JUSTIFIED IN COLUMNS 1-3.
C THE PROGRAM IS NOW SET TO ACCEPT ONE CARD FOR EACH CASE.  THE VARIABLES
C ARE INPUT IN THE ORDER X1,X2 AND Y WITH 10 COLUMNS FOR EACH VARIABLE
C IN F FORMAT.  I OFTEN ALTER THE READ(1,30) AND 30 FORMAT STATEMENTS
C TO FIT THE DATA ON THE CARD.  THE PROGRAM IS SET TO ACCEPT A MAXIMUM
```

```
C OF 99 CASES BUT ONE COULD ENLARGE THE MATRICES IN THE DIMENSION
C STATEMENT TO ACCEPT MORE.
C
C
      DIMENSION PLOT(99),RATE(99),HRS(99),Y(99),X(2,99)
      DATA  SUMHRS,SUMX1,SUMX2,SUMY,SSX1,SSX2,SUMX1Y,SUMX2Y,SMX1X2,SSER,
     1      SSE,SSTO,SSERL,SSTOL,SSEL,SSERRL,SSERR/17*0./
C
C   INPUT DATA AND TRANSFORM THE VARIABLES TO LOGARITHMS.
C
      READ(1,10)NCASES
10    FORMAT(I3)
      DO 50 I = 1,NCASES
         READ(1,30)PLOT(I),RATE(I),HRS(I)
30       FORMAT(5X,2F5.0, 5X,F5.0)
         X(1,I) = ALOG10(PLOT(I))
         X(2,I) = ALOG10(RATE(I))
         Y(I)   = ALOG10(HRS(I))
C
C   CALCULATE THE SUMS OF THE VARIABLE PRODUCTS.
C
         SUMHRS = SUMHRS + HRS(I)
         SUMX1  = SUMX1   +X(1,I)
         SUMX2  = SUMX2  + X(2,I)
         SUMY   = SUMY   + Y(I)
         SSX1   = SSX1   + X(1,I)**2


         SSX2   = SSX2   + X(2,I)**2
         SUMX1Y = SUMX1Y + X(1,I)*Y(I)
         SUMX2Y = SUMX2Y + X(2,I)*Y(I)
         SMX1X2 = SMX1X2 + X(1,I)*X(2,I)
```

```
50    CONTINUE
C
C  CALCULATE B0 AND B1 FOR THE REDUCED MODEL.
C
      B1 =(SUMX1Y-((SUMX1*SUMY)/NCASES))/(SSX1-(SUMX1**2/NCASES))
      YBAR = SUMY/NCASES
      HRSBAR = SUMHRS/NCASES
      X1BAR = SUMX1/NCASES
      X2BAR = SUMX2/NCASES
      B0 = YBAR- B1*X1BAR
      AB0 = 10.**B0
C
C  CALCULATE THE RESIDUALS FOR THE REDUCED MODEL.
C
      DO 150 I = 1,NCASES
         YHATL = B0+B1*X(1,I)
         RESIDL = Y(I) - YHATL
         SSERL = SSERL + RESIDL**2
         SSTOL = SSTOL + (Y(I) - YBAR)**2
         YHAT = 10.**YHATL
         RESID = HRS(I)-YHAT
         SSER = SSER + RESID**2
         SSTO = SSTO +  (HRS(I)-HRSBAR)**2
150   CONTINUE
C
C  CALCULATE STATISTICS FOR THE REDUCED MODEL.
C
      R2LOG = (SSTOL-SSERL)/SSTOL
      FMSEL = SSERL/(NCASES-2)
      FLOG = (SSTOL-SSERL)/FMSEL
      RSQR =  (SSTO-SSER)/SSTO
      FMSER = SSER/(NCASES-2)
C
      WRITE(3,5)
5     FORMAT('1')
      PRINT,'RESULTS OF THE REDUCED MODEL WITHOUT PRODUCTION RATE.'
```

```
      PRINT, '********************************************************'
      PRINT,' '
      PRINT,' '
      PRINT,'YHAT = ',ABO,'*(X1**',B1,')'
      PRINT,' '
      PRINT,'STATISTICS IN LOGARITHMS'
      PRINT, ' '
      PRINT,'LOG R2 = ',R2LOG,'LOG MSE = ',FMSEL,'LOG F = ',FLOG
      PRINT,' '
      PRINT,'STATISTICS DEVELOPED FROM ORIGINAL VARIABLE FORM.'
      PRINT, ' '
      PRINT,'R2 ACTUAL = ',RSQR,'MSE ACTUAL = ',FMSER
      PRINT,' '
C
C  CALCULATE BO AND B2 FOR THE REDUCED HRS V RATE MODEL.
C
      B2 =(SUMX2Y-((SUMX2*SUMY)/NCASES))/(SSX2-(SUMX2**2/NCASES))
         BO = YEAR - B2*X2BAR
         ABO = 10.**BO
C




C CALCULATE AND PRINT RESIDUALS FOR THE REDUCED HRS V RATE MODEL.
C
      DO 200 I = 1,NCASES
         YHATL = BO + B2*X(2,I)
         RESIDL = Y(I)-YHATL
         SSERRL = SSERRL + RESIDL**2
         YHAT = 10.**YHATL
         RESID = HRS(I) - YHAT
         SSERR = SSERR + RESID**2
200   CONTINUE
```

```
C
C   CALCULATE AND PRINT STATISTICS FOR THE REDUCED HRS V RATE MODEL.
C
      R2LOG =(SSTOL- SSERRL)/SSTOL
      FMSEL = SSERRL/(NCASES-2)
      FLOG = (SSTOL-SSERRL)/FMSEL
      RSQR =(SSTO-SSERR)/SSTO
      FMSER = SSERR/(NCASES-2)
C
      PRINT,' '
      PRINT,'RESULTS OF A REDUCED MODEL WITHOUT CUMULATIVE LEARNING.'
      PRINT, '*******************************************************'
      PRINT,' '
      PRINT,' '
      PRINT,'YHAT = ',AB0,'*(X2**',B2,')'
      PRINT,' '
      PRINT,'STATISTICS IN LOGARITHMS'
      PRINT,' '
      PRINT,'LOG R2 = ',R2LOG,'LOG MSE = ',FMSEL,'LOG F = ',FLOG
      PRINT,' '
      PRINT,'STATISTICS DEVELOPED FROM ORIGINAL VARIABLE FORM.'
      PRINT, ' '
      PRINT,'R2 ACTUAL = ',RSQR,'MSE ACTUAL = ',FMSER
C
C   CALCULATE B0,B1, AND B2 FOR THE FULL MODEL.
C
      DENOM = ((SSX1-X1BAR*SUMX1)*(SSX2-X2BAR*SUMX2) -(SMX1X2-X1BAR*
     6    SUMX2)**2)
      B1 =((SSX2-X2BAR*SUMX2)*(SUMX1Y-X1BAR*SUMY) -
     7    (SMX1X2-X1BAR*SUMX2)*(SUMX2Y-X2BAR*SUMY))/DENOM
      B2 = ((SSX1-X1BAR*SUMX1) *(SUMX2Y-X2BAR*SUMY) -
     8    (SMX1X2-X1BAR*SUMX2)*(SUMX1Y-X1BAR*SUMY))/DENOM
      B0 = YBAR-B1*X1BAR-B2*X2BAR
      AB0 = 10.**B0
C
C   CALCULATE RESIDUALS FOR THE FULL MODEL.
```

```
C
      PRINT,' '
      PRINT,'THE FOLLOWING DATA COME FROM THE FULL MODEL.'
      PRINT, '***************************************************************'
      PRINT,' '
      WRITE(3,210)
210      FORMAT('0','CASE  OBSERVED  PREDICTED  RESIDUAL  PER CENT  ',
     1   '  LOG        LOG          DIFFERENCE')
      WRITE(3,211)
211   FORMAT(' ',6X,'VALUE      VALUE',16X,'DEVIATION   OBSERVED',
     2   '  PREDICTED')
      DO 250 I= 1,NCASES
         YHATL = B0 + B1*X(1,I) + B2*X(2,I)
         RESIDL = Y(I) - YHATL




         SSEL = SSEL + RESIDL**2
         YHAT = 10.**YHATL
         RESID = HRS(I)-YHAT
         SSE = SSE + RESID**2
         PERCEN = 100.*RESID/HRS(I)
         WRITE(3,115)I,HRS(I),YHAT,RESID,PERCEN,Y(I),YHATL,RESIDL
115      FORMAT(' ',I3,2X,3(F6.3,5X),F6.2,5X,3(F11.5))
250   CONTINUE
C
C  CALCULATE SOME STATISTICS FOR THE FULL MODEL.
C
      R2LOG = (SSTOL-SSEL)/SSTOL
      FMSEL = SSEL/(NCASES-3)
      FINCL = (SSERL-SSEL)/FMSEL
      FINCL2=(SSERRL-SSEL)/FMSEL
      FLOG = (SSTOL-SSEL)/2./FMSEL
```

```
      SSEST = SQRT(FMSEL)
      SEB1 = SQRT((FMSEL*(SSX2-X2BAR*SUMX2))/DENOM)
      SEB2 = SQRT((FMSEL*(SSX1-X1BAR*SUMX1))/DENOM)
      TB1 = B1/SEB1
      TB2 = B2/SEB2
      RSQR = (SSTO-SSE)/SSTO
      FMSEA = SSE/(NCASES-3)
C
      PRINT, ' '
      PRINT,'YHAT = ',ABO,'*(X1**',B1,')*(X2**',B2,')'
      PRINT,' '
      PRINT,'STATISTICS IN LOGARITHMS'
      PRINT,' '
      PRINT,'LOG R2 = ',R2LOG,'LOG MSE = ',FMSEL,'LOG F = ',FLOG
      PRINT,' '
      PRINT,'LOG F INCREMENTAL FROM ADDING X2 = ',FINCL
      PRINT,' '
      PRINT,'LOG F INCREMENTAL FROM ADDING X1 = ',FINCL2
      PRINT,' '
      PRINT,'TB1=',TB1,'TB2=',TB2
      PRINT,' '
      PRINT,'STD ERROR OF EST IS ',SSEST
      PRINT, ' '
      PRINT,'STATISTICS DEVELOPED FROM ORIGINAL VARIABLE FORM.'
      PRINT,' '
      PRINT,'R2 ACTUAL = ',RSQR,'MSE ACTUAL = ',FMSEA
      WRITE(3,5)
      STOP
      END
```

SELECTED REFERENCES

Alchian, Armen A. and Allen, William R. University Economics. Belmont, California: Wadsworth Publishing Company, Inc., 1964.

Asher, Harold. Cost-Quantity Relationships in the Airframe Industry. Santa Monica, California: The Rand Corporation, 1956.

Beals, Ralph E. Statistics for Economists. Chicago, Illinois: Rand McNally and Co., 1972.

Cochran, E. B. Planning Production Cost: Using the Improvement Curve. San Francisco, California: Chandler Publishing Company, 1968.

Dean, Joel. Managerial Economics. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1961.

Dunne, William E. "Microeconomic Theory Applied to Parametric Cost Estimation of Aircraft Airframes." GOR/SM/75D-3, M.S. Thesis, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio, 1975.

Fazio, Peter F. and Russell, Stephen H. "An Analytical Approach To Optimizing Airframe Production Costs As A Function Of Production Rate." SLSR 30-74A, M.S. Thesis, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio, 1974.

Johnson, Gordon J. "The Analysis of Direct Labor Costs for Production Program Stretchouts." National Contract Management Journal (Spring 1969): 25-41.

Kroeker, Herbert R. and Peterson, Robert. A Handbook of Learning Curve Techniques. Columbus, Ohio: The Ohio State University Research Foundation, 1961.

Large, Joseph P.; Hoffmayer, Karl; and Kontrovich, Frank. Production Rate and Production Cost. R-1609-PA&E, Santa Monica, California: The Rand Corporation, December 1974.

Neter, John and Wasserman, William. Applied Linear Statistical Models. Homewood, Illinois: Richard D. Irwin, Inc., 1974.

Noah, J. W. "Resource Input vs. Output Rate and Volume in the Airframe Industry." Draft Technical Report Number TR-204-USN, Contract Number N00014-73-C-0319, Washington, D.C.: J. Watson Noah Associates, Inc., December 1974.

Noe, Charles G.; Smith, Larry L.; and Jacobs, Grady L. Defense Cost and Price Analysis. 2 Vols. Gunter Air Force Station, Alabama: Extension Course Institute, 1974.

Orsini, Joseph A. "An Analysis of Theoretical and Empirical Advances in Learning Curve Concepts Since 1966." GSA/SM/72-12, M.S. Thesis, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio, 1970.

Spurr, William A. and Bonini, Charles P. Statistical Analysis for Business Decisions. Homewood, Illinois: Richard D. Irwin, Inc., 1967.

U.S. Department of Defense. Armed Services Procurement Regulation, The 1973 Edition. Washington, D.C.: Government Printing Office, 1973.

U.S. Department of Defense. Armed Services Procurement Regulation Manual for Contract Pricing. Washington, D.C.: Government Printing Office, February 1969.

Wonnacott, Thomas H. and Wonnacott, Ronald J. Introductory Statistics. 2nd Ed. New York, New York: John Wiley and Sons, Inc., 1972.

Larry L. Smith