ECONOMIES OF SCALE IN COMPUTER CENTERS

Thesis for the Degree of Ph. D.
MICHIGAN STATE UNIVERSITY
GERALD L. MUSGRAVE
1973

This is to certify that the

thesis entitled

ECONOMIES OF SCALE IN COMPUTER CENTERS

presented by

Gerald L. Musgrave

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Economics

Major professor

Date

ec

pl

th

th

le

th

co

eq

in

ac

ad

of

of

es

th

ABSTRACT

ECONOMIES OF SCALE IN COMPUTER CENTERS

By

Gerald L. Musgrave

The objective of this study is to determine if economies of scale exist in the production of computer output. A model of the production process was constructed on the assumption that the computer centers attempt to minimize the cost of producing an exogenously determined expected level of output.

Stochastic disturbances are assumed to be present in the production process due to the unpredictable nature of computer hardware and software failure. The factor demand equations are also assumed to be non-deterministic because of imperfections in management. The adjustment of inputs from actual to desired levels is modeled as a stochastic stock adjustment process where the adjustment rate is a function of the cost of adjustment.

Indirect least squares estimates of the parameters of the production function were obtained from least squares estimates of the reduced form cost function coefficients. The data, a 1965 cross-section of 115 college and university

computer centers, indicate that the null hypothesis should be rejected in favor of the alternative of increasing returns to scale.

ECONOMIES OF SCALE IN COMPUTER CENTERS

By

Gerald L. Musgrave

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Economics

1973

G81346

ACKNOWLEDGMENTS

I am pleased to acknowledge the help and guidance given to me by my thesis chairman, Jan Kmenta. His care and patience in reading drafts and revisions was more than could be expected and without his efforts the thesis would not have been completed. Robert Rasche contributed generously of his time and his counsel improved the thesis. Professor Gustafson gave the thesis a careful reading. Partial financial assistance was provided by the Institute of Public Utilities at Michigan State University.

Over many years, two of my teachers contributed greatly to my development. My Principles teacher and friend, Paul Smith, has been a source of encouragement throughout my economics training. Professor Thomas Saving instilled an interest in economic theory in his students. His ability to communicate economic ideas set a high standard for his students.

My colleague Wally Creighton provided moral support and confidence in my work. My association with him is a pleasure and honor. Finally, in this time of consumerism, I remain responsible for all defects in the thesis.

TABLE OF CONTENTS

		Page
ACKNOWLEDGMENTS		ii
LIST OF TABLES		v
LIST OF FIGURES		vi
Chapter		
I.	STATEMENT OF THE PROBLEM	1
	1. Introduction	1
	2. Nature of the Computation Process	1
	3. Nature of Input Factors	5
	4. Operational Charter of Computer Centers	7
II.	REVIEW OF LITERATURE	8
	1. Introduction	8
	2. Notes on Material Related to Production Functions	10
	3. Literature Pertaining to Economies of Scale in Computers	14
	4. Summary	19
III.	THEORETICAL MODEL	21
	1. Introduction	21
	2. Stochastic Nature of the Production Process	21
	3. Output from the Production Process	23
	4. Factors of Production	36
	5. Output and Input Price	39
	6. Behavioral Constraint	41
	7. Specification of the Cost Minimization Model	43
	8. Summary	56

Chapter | Page

IV.	DATA AND EMPIRICAL RESULTS.		58
	1.	Introduction	58
	2.	Sources of Data	58
	3.	Method of Data Collection.	61
	4.	Cost Function.	63
	5.	Omission of Relevant Explanatory Variables	68
	6.	Empirical Results	73
V.	CONCLUDING REMARKS		81
BIBLIOGRAPHY			88
APPENDIX A			94
APPENDIX B			112
APPENDIX C			114

Tabl

I

II

A-I

C-I

C-II

LIST OF TABLES

Table		Page
I.	Test Statistic with $W_A = W_C = .5$	74
II.	Output Elasticities	79
A-I.	The Data	105
C-I.	Test Statistic with $W_A = .9$, $W_C = .1$.	114
C-II.	Test Statistic with $W_A = .1$, $W_C = .9$.	115

LIST OF FIGURES

Figure Page

A-1. Run Diagram for Generalized File Processing . 95

A-2. Flow Chart for Generalized File Processing, Problem A. 96

A-3. Run Diagram for Random Access File Processing Problem. 98

A-4. Flow Chart for Random Access File Processing Problem. 100

ies

To d

sis

proc

sis

to t

pria

are

be i

proc

Scie

Whic

not

ener

the

stud

CHAPTER I

STATEMENT OF THE PROBLEM

1. Introduction

The purpose of this study is to determine if economies of scale exist in the production of computer output. To do that, it is necessary to formulate a testable hypothesis by constructing an economic model of the production process. On the basis of the model we can test the hypothesis of economies of scale using the available data. Prior to the construction of the theoretical model it is appropriate to delineate what aspects of the computation process are to be considered and what factors of production are to be included in the analysis.

2. Nature of the Computation Process

The entire process of computation and information processing by the computer is so complex that computer scientists have not yet developed the technical relations which are available for other engineering processes. We do not have a measure of the "horse-power" or the potential energy of a computer. It is therefore necessary to define the production of the computer on an ad-hoc basis for this study as output from the central processing unit of the comput

a firs

The ce

tions,

from v

inform

might

nomer

on eco

justif

total

the CP

the up

This u

betwee

above

These

units,

comput

is the

sis to

import

by the

hold th

computer, it is the most important part to be considered as a first step in understanding the whole computer system. The central processing unit (CPU) performs arithmetic operations, logical tests or branches, interprets instructions from various peripheral devices, stores and retrieves information. This list of tasks resembles what a layman might call "thinking" and is one explanation for the misnomer of "Electronic Brain."

Restricting the analysis to the CPU is justified on economic as well as on engineering grounds. Engineering justification is based upon the importance of the CPU to the total computing system. All information must pass through the CPU or be controlled by it. The CPU's rate of output is the upper-bound or capacity of the computer to process data. This upper-bound is a result of the technologic relation between the speed of the CPU to execute the tasks listed above and the data transmission devices connected to the CPU. These data transmission devices, sometimes called peripheral units, are much slower than the CPU. Because of measuring computer output in terms of CPU output, the resultant measure is the capacity of the computer system.

The economic justification of restricting the analysis to CPU output is based on four issues. First, it is important to obtain an output measure which is not influenced by the quality of the computer output. It is desirable to hold the quality, those variables which allow one to say this

is go
quali
the e
of pr
put i
machi
skill
tative
analys
of the
which
progra
grammi

with t
centra
cost f
parame
sis, i
the pr
of the

physic
motion
inputs
cost fu

is good output and that is bad output, constant. In holding quality constant it is necessary to exclude from the output the effect of programming, systems analysis and availability of pre-programmed applications. The needed measure of output is designed to account for the physical output of the machine and not be influenced by differences in programmers' skills, efficiency of applications programs or other qualitative influences on computer output. By restricting the analysis to the CPU, it is possible to measure the output of the computer system in terms of machine computations which are not influenced by the qualitative differences in programmers, applications programs or availability of programming systems.

Second, a major component of the cost associated with the physical computer system can be allocated to the central processing unit. Since it will be shown that the cost function plays an important role in estimating the parameters necessary to test the economies of scale hypothesis, it is desirable to include in the model that part of the production process which accounts for a large portion of the cost.

Third, it is desirable that output be measured in physical units. This is in keeping with the neoclassical notion of a production function which relates a vector of inputs to the maximum quantity of production. Since the cost function is derived from the production function, it

is
the

fur
eff
imp
pre
sar
out
the
out
ter
(In
and
ert
out

ter
ope
par
par
itse
capa
the
mate
prog

is necessary that the output measure is consistent with the theory of production.

Fourth, because of the definition of the production function, output is considered to be the set of technically efficient outputs from the system. "Technically efficient" implies that the product has economic value and one would prefer more of the product rather than less, given that the same factors of production could produce both quantities of output. This property of the function is consistent with the formulation of the problem in terms of central processor output. Output from the central processor is measured in terms of maximum possible output given the factor inputs. (In Chapter IV a discussion of output measures is presented and it is demonstrated that the measure chosen has the property of measuring the computer's capacity to produce output.)

In summary, the production process of computer centers has been restricted to the production of machine operations performed by the central processing unit. This part of the computing system is responsible for a major part of the costs of hardware or the physical machine itself. The central processing unit also determines the capacity of the system to produce output. By restricting the analysis to machine operations, it is possible to eliminate the influence of quality differentials due to programmers or application programs. The physical units of cor

spe

out

a s

thi

sid

sho

out

duc

Sin

sys

exi

dema

Beca

sing

inpu

sets

nel.

the

class

input

The m

computer computation are consistent with the objective of specifying an economic model for the production of computer output.

3. Nature of Input Factors

In the previous section the rationale for choosing a single measure of computer output was discussed and in this section the inputs to the production process are considered. The objective here is to explain which factors should be considered in the production process for computer output.

Labor service is the most important factor of production which can be varied by the center's administration. Since the center has control over usage of this factor, a systematic relation between output and factor usage should exist if the center is adjusting the factor usage as output demand varies. Unfortunately, labor is not a single input. Because of the large diversity of jobs in the center a single measure of total labor input is not justified. This input must be separated into components which represent sets of tasks which can be performed by specialized personnel. It is also important to separate these groups because the management of these facilities adjusts the various classes of employees at different rates. Different factor input ratios exist between these classes of employees also. The most appropriate differentiation of these groups is betw

full

III w

the c

diffe

input

and c

on ca

facto

the l

rathe

from

betwe

employ

more p

syster

we are

of di

Becaus

tial e

influe

to inc

measur

all sec

between operations, administration and programming staff. A full explanation of this differentiation is given in Chapter III when the model is explained, and in Chapter IV concerning the data and measurement of these inputs. The problem is to differentiate the center's labor force and to relate these inputs to the center's output production.

One might argue that once the machine is turned on and operating properly the quantity of labor has no effect on capacity output. It should be remembered that both factor prices and the level of output jointly determine the level of factor usage. Input variables are endogenous rather than exogenous to the conceptual model. It appears from casual observation that a systematic relation exists between the output of computers and the labor services employed at the installations. Larger facilities require more personnel. This relation may be highly variable within systems of similar output capacities. But in this analysis we are interested in determining the relation between CPU's of different output capacities and factor inputs.

Another important input is the quantity of capital. Because of a lack of engineering theory explaining differential effects of various electronic components and their influence upon computer performance, no attempt will be made to include technical differences between machines in the measure of capital. Also, the machines in the sample are all second generation computers. They are batch processing,

n
i
s
t
al
v

ar
pr
pr
tu
ti
er

dir
to
the
pri
dat
beh
cos

atte
cost
tion
dire

non-time sharing machines which have transistors but not integrated circuit electronics. Because the machines are similar with regard to technology, they are all considered to be of the same vintage. This assumption allows us to abstract from the problem of technological change and innovation.

4. Operational Character of Computer Centers

Computer centers under consideration in this study are operated by institutions of higher education. This fact presents an interesting analytic problem in that the direct price of the product is zero. Indirect prices or opportunity costs to the user such as waiting and programming time are not zero. These indirect prices are not considered in the study.

Since the center does not charge the consumer a direct price for the output, it does not seem appropriate to consider the center as a profit maximizing firm. Even if the center charged some of the users (on funded research or private consulting, etc.) the ersatz calculation of revenue data would be questionable. Fortunately, an alternative behavioral assumption is available and this assumption is cost minimization.

The computer center is assumed to be a firm which attempts to meet the output demands placed on it at minimum cost. These demands are generated by the day-to-day operations of the educational institution and are not under the direct control of the computer center.

CHAPTER II

REVIEW OF LITERATURE

1. Introduction

Unfortunately, the literature concerning estimation of production functions for computer centers is sparse, but work related to production functions has been done in the computer engineering field. None of the studies reviewed considered any objective of the organization, say, profit maximization or cost minimization. It should be emphasized that this study is concerned with output or production from computers and not the production or manufacturing of computers. No attempt, in this study, is made to analyze the computer manufacturing industry. An outstanding study and thorough bibliography of work done in this area is available in Billings and Hogan [1970].[1] This chapter does not contain a review of the general literature of econometric estimation of production or reduced form cost functions. The reader who is interested in production function literature is referred to the following ordered list: deterministic

---

[1] H. Billings and R. Hogan, "A Study of the Computer Manufacturing Industry in the United States" (unpublished Master's thesis, U.S. Naval Postgraduate School, Monterey, California, 1970).

pr
ri
19
me
[M
Km
co
ma
de
ti
fo
co
Ch
of
Ne

lit
com
mar
sci
are
com
stu
The
put
to t
and

production functions [Ferguson, 1959, part I]; econometric background for estimating production functions [Kmenta, 1971, especially Chapter 11 on non-linear models]; econometric specification and estimation of production functions [Marschak and Andrews, 1944; Walters, 1963; Zellner, Kmenta, and Dréze, 1966; Kmenta, 1967; and Mundlak, forthcoming]. The reader interested in the cost function material is referred to the following ordered list: deterministic cost functions and their relation to production functions [Shephard, 1953; Uzawa, 1964; McFadden, forthcoming]; statistical background for estimation of cost functions [Johnson, 1960; and Malinvaud, 1968, Chapter 16 on simultaneous equation models]; estimations of reduced form cost functions [Merewitz, 1972, and Nerlove, 1964].

A number of papers have appeared in the computer literature concerning alternative methods of "evaluating" computer systems. The work in this area originates primarily from two fields--operations research and computer science. From the economist's point of view, these studies are about the definition or description of the nature of computer output. In Chapter IV, the method used in this study is discussed in the context of the literature. The reader who is interested in pursuing the study of output evaluation in the operations research field is referred to the following: stochastic processes [Schneidewind, 1966, and Schwab, 1967]; computer simulations [Huesmann and g

re

p

a

b

fu

by

an

ma

el

de

Pr

Ma

Pre
Ma

"Ca
of

Tes
duct
Ass

Goldberg, 1967; Ihrer, 1967; and Knight, 1963]. The reader interested in computer systems evaluation from the perspective of the computer scientist is directed to Arbuckle [1966], and an outstanding review article of batch processing systems by Calingaert [1967].

2. Notes on Material Related to Cost and Production Functions

The Cobb-Douglas functional form first appeared in 1928 in the *American Economic Review*.[2] Douglas used this function to analyze the share of national income received by various sectors of the economy--the labor market being his prime concern.[3] Later microeconomic data were analyzed and the restrictive assumptions of Douglas were dropped; namely, the sum of the output elasticities equalling one,[4] elasticity of technical substitution equalling one,[5] and the development of other less restrictive functional forms.[6]

---

[2] C. Cobb and P. Douglas, "A Theory of Production," *American Economic Review*, Vol. 18 (March, 1928), pp. 139-65.

[3] P. Douglas, *The Theory of Wages* (New York: Macmillan, 1934).

[4] P. Douglas, "Are There Laws of Production?" Presidential Address, *American Economic Review*, Vol. 38 (March, 1948), pp. 1-41.

[5] K. Arrow, H. Chenery, B. Minhas and R. Solow, "Capital-Labor Substitution and Economic Efficiency," *Review of Economics and Statistics* (August, 1961), pp. 225-50.

[6] J. Ramsey and P. Zarembka, "Specification Error Tests and Alternative Functional Forms of the Aggregate Production Function," *Journal of the American Statistical Association*, Vol. 66 (September, 1971), pp. 471-77.

Th

co

me

de

cu

(t

pu

pr

ti

it

ti

one

one

Thi

of

cus

proc

tech

func

---

tion
1967

ton:

The cost one pays for the greater generality has been more complicated estimation procedures, but fortunately ingenious methods have been devised to minimize these difficulties.[7]

Shephard [1953] demonstrated that, with the usual definition of the production function (the set of maximum outputs for given quantity of inputs) and the cost function (the set of minimum costs of producing given levels of outputs) plus a restriction on the production function, the production function can be determined from the cost function.[8] The restriction on the production function is that it is convex. In terms of geometry, the convexity restriction on the production surface means that "a decrease in one coordinate (input value) without increasing at least one other coordinate results in a lower output rate."[9] This restriction is equivalent to diminishing marginal rate of technical substitution.

Uzawa [1964] extended Shephard's results and discussed the structure of cost functions derived from given production sets. He found that if the marginal rates of technical substitution are non-increasing the total cost function, as defined above, is determined for positive input

---

[7]J. Kmenta, "On Estimation of the C.E.S. Production Function," _International Economic Review_, Vol. 8 (June, 1967), pp. 180-89.

[8]R. Shephard, _Cost and Production Functions_ (Princeton: Princeton University Press, 1953), p. 22.

[9]_Ibid_., p. 4.

prices and output quantities, is continuous, non-negative, homogeneous of degree one in input prices, and concave with respect to the output level.[10]

Uzawa and Shephard assume pure competition in both input and output markets. The impact of the Shephard-Uzawa duality theorem is that the information about the production technology is contained in either the production or the cost function.

Nerlove [1963] used the duality principle to estimate the parameters of a Cobb-Douglas production function for electric power generation by single equation ordinary least squares method. It was assumed that power producers minimize the cost of producing an exogenously determined level of output. Nerlove found that under these conditions, the cost function was linear in the logarithms of output, factor prices, and a stochastic term in the production function which allowed for "neutral" variations in efficiency among firms.[11]

Using a cross-section of 145 privately owned electric utilities for the calendar year 1955, Nerlove concluded ". . . there is evidence of a marked degree of

---

[10] H. Uzawa, "Duality Principles in the Theory of Cost and Production," International Economic Review, Vol. 5 (May, 1964), p. 217.

[11] M. Nerlove, "Returns to Scale in Electricity Supply," in Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld, ed. by C. Christ (Stanford: Stanford University Press, 1963), p. 106 and footnote number 11, p. 128.

increasing returns to scale at the firm level in U.S. Steam electricity generation." He also found, first, that

> The appropriate model at the firm level in this industry is a statistical cost function which includes factor prices and which is uniquely related to the underlying production function. . .

and, secondly, that

> . . . at the firm level, it is appropriate to assume a production function which allows substitution among factors of production. When a statistical cost function based on a generalized Cobb-Douglas production function is fitted to cross-section data on individual firms there is evidence of such substitution possibilities.[12]

In all, 28 regressions were run using various restrictions on the coefficients of the cost function, assumptions concerning the nature of technical change, and the homogeneity of the production function. The majority of the estimates of the output elasticities for capital were in the range of 0.0 to 0.25, output elasticity for fuel was between 0.50 and 0.75, and the majority of the point estimates for the labor output elasticity were between 0.50 and 0.75 with 4 being greater than one. As mentioned earlier, the sum of these output elasticities was greater than one with 9 of the 28 estimates being greater than 2.00.

Using the same methodology as Nerlove but extended to multiple outputs, Merewitz [1969] estimated cost functions for small- and medium-sized post offices. The operations were intra-office operations which represent processing of mail and retailing. Using a 1966 cross-section of 156 offices, he found moderately increasing

---

[12] _Ibid._, p. 126

returns to scale. The output elasticities of General Services Administration floor space, inside and platform space and capital were positive and less than one. The output elasticity of labor was greater than one.[13]

It should be noted that in the Nerlove and Merewitz studies cited the objective of the organization was to minimize cost for an exogenously determined output. If the behavioral constraint were profit maximization with competitive input and output markets, the economies of scale results would be inconsistent with the second order stability conditions.[14] An additional inconsistency with the first order extrema condition would result if any output elasticity were greater than one.[15] It is shown in Chapter III that these inconsistencies do not hold for the cost minimization case.

3. Literature Pertaining to Economies of Scale in Computers

In the late 1940's, Herbert Grosch believed the average cost of computation to be a decreasing function of the size of the computer. Solomon [1966] reported that Grosch said output would increase as the square of cost.

---

[13] L. Merewitz, *The Production Function in the Public Sector: Production of Postal Services in the U.S. Post Office* (Berkeley: Center for Planning and Development, 1969), p. 62.

[14] J. Henderson and R. Quandt, *Microeconomic Theory* (New York: McGraw-Hill, 1971), p. 95.

[15] *Ibid.*, p. 97.

Because this idea was formulated in the 1940's, it is not clear if the hypothesis is about retail price of the computer or its manufacturing cost. In fact, Grosch may have been discussing short-run average total cost and the influence of fixed versus variable cost. The assertion of Grosch could be interpreted as if we compared two computers, one twice as costly as the other, the former would have four times the capacity output rate of the latter. Grosch never published his belief and it is part of the oral tradition of early engineering work on computers.[16]

Knight [1963] was concerned with the technical engineering changes which occurred in computers produced in the period 1950 to 1962. Knight also considered the relation of computer power to average computer rental cost. His measure of power is the quantity of a set of instructions which could be performed in one second multiplied by a factor to account for memory size. He attempted to develop experimentally a representative set of instructions by examining programs which were classified as either scientific or commercial. After examining the programs, different sets of instructions were used for the two classifications.

The measure of cost was the number of seconds of computer processing time that could be purchased for one

---

[16] William F. Sharpe, The Economics of Computers (New York: Columbia University Press, 1969), p. 315.

dollar on the various machines. If Knight's measure of cost is K, it is convenient to express his results by using $C = \frac{1}{K}$, cost in terms of dollars per second. Presumably to obtain a U-shaped average cost curve, Knight hypothesized

$$\frac{C}{Q_p} = e^{\alpha_0} Q_p^{\alpha_1 + (\alpha_2 \ln Q_p) - 1}$$

where

C = computer system cost (\$1/second)

Q_p = Quantity of computer power (operations/second)

α's = parameters to be estimated.

i. If $\alpha_1, \alpha_2 > 0$ and $(\alpha_1 + \alpha_2 \ln Q_p - 1) < 0$

then $\frac{d(C/Q_p)}{dQ_p} < 0$, average cost decreases as capacity output increases.

ii. If $\alpha_1, \alpha_2 > 0$ and $(\alpha_1 + \alpha_2 \ln Q_p - 1) > 0$

then $\frac{d(C/Q_p)}{dQ_p} > 0$, average cost increases as capacity output increases.

iii. If $\alpha_1 < 1$ and $\alpha_2 = 0$ then $(\alpha_1 + \alpha_2 \ln Q_p - 1) < 0$

and we have case i, above.

iv. If $\alpha_1 < 1$ and $\alpha_2 > 0$, we have the U-shaped average cost function.

Since the technical change was the concern of the paper, Knight hypothesized any change in technology which occurred would cause a shift in the cost function. His single equation became

$$\text{(a)} \quad \ln C = \alpha_0 + \alpha_1 \ln Q_p + \alpha_2 (\ln Q_p)^2 + \sum_{j=1}^{n} B_j w_j$$

where the w_j terms are binary variables (1 if the machine was first produced in year j, 0 otherwise), j = 51, ..., 62. Knight ran a second equation omitting the $(\ln Q_p)^2$ term

$$\text{(b)} \quad \ln C = \alpha_0 + \alpha_1 \ln Q_p + \sum_{j=1}^{n} B_j w_j .$$

He found the R^2's of equations (a) and (b) to be "close" and thus "very little additional explanatory power was gained by allowing for a U-shaped average cost curve."[17] In addition, the original regression results were not presented but the second regression was presented after some of the observations were excluded. The sample points were deleted if "actual cost exceeded that predicted by more than one-half the standard deviation (error) of predicted cost."[18] This process of eliminating observations which do not fit is more prevalent in engineering than in other

---

[17]K. Knight, "A Study of Technological Innovation--The Evaluation of Digital Computers" (unpublished Doctoral dissertation, Carnegie Institute of Technology, 1963).

[18]*Ibid.*

fields. It seems the objective was to eliminate "over-priced" systems, which means ones from a different population. Once this process of elimination was completed, it was assumed the equation would hold exactly except for possible measurement errors. Thus the influence of other explanatory variables, other than Q_p, was supposedly held constant--like a laboratory experiment. The reader interested in the econometric approach to such issues is referred to Kmenta [1971, Chapter 10 and the section on omission of relevant explanatory variables]. Knight estimated α_1 at 0.519 for scientific computation and α_1 equalling 0.459 for commercial data processing. These results suggest economies of scale exist for computers produced between 1950 and 1962.

The only paper which dealt directly with the issue of economies of scale is that by Solomon [1966] who considered International Business Machines Systems 360 models 30, 40, 50, 65 and 75. The author assumed that no technical differences existed between these machines and believed that any differences in production rates could be explained by differences in the size of computers. This paper was concerned with what was described as "cost vs. performance" and since the paper was engineering in nature, no behavioral model was constructed.[19] The regression equation was

[19] M. Solomon, "Economies of Scale and the I.B.M. System/360," Communications of the ACM, Vol. 5 (June, 1966), pp. 435-40.

$$\ln(C_i) = a + b \ln (T_i), \quad i = 1, \ldots, 5$$

where C is monthly rental and T is the time to compute a set of computer instructions. Note that T is the inverse of output, the quantity of instructions processed in a unit of time, e.g., $(T = \frac{1}{Q})$, and we have

$$\ln(C_i) = a - b \ln(Q_i), \quad i = 1, \ldots, 5.$$

Output was defined in four ways: matrix multiplication, floating point square root, field scan, and a scientific instruction mix suggested by Arbuckle [1966]. Using ordinary least squares, he found

b	Instruction Type	R^2
-0.494	Matrix Multiplication	.989
-0.478	Floating Point Square Root	.999
-0.632	Arbuckle's Scientific Mix	.977
-0.682	Field Scan	.969

He concludes, from the sample of five, that economies of scale exist for the 360 series.

4. Summary

The available evidence suggests that the parameters of the production function can be estimated from the cost function. Two studies have been reviewed which use this reduced form cost function technique on regulated industries with some success.

Evidence concerning economies of scale of computer centers is less convincing. Since the work thus far is in the area of engineering curve fitting, the construction of an economic production process may be interesting and is the subject of the next chapter.

CHAPTER III

THEORETICAL MODEL

1. Introduction

In this chapter, the problem of specifying a model of the cost minimizing computer center operating with a Cobb-Douglas production function is analyzed. It is assumed that the productive process is stochastic and that the managers of the computer center include this fact in their cost minimizing actions. What follows is an analysis and justification of the assumption of the nondeterministic nature of the productive process of computer output. In Sections 3 and 4, output and factor inputs are discussed. Section 5 deals with the unit prices of the factors of production and Section 6 is a presentation of the behavioral constraints which influence the computer center management. In Section 7, the specification of the cost minimization model is presented. The final section is a summary of the work in this chapter.

2. Stochastic Nature of the Production Process

The specification of computer production developed in this chapter is based on the assumption that the production function is of the Cobb-Douglas type. This assumption seems at least partially justified on the basis of the Nerlove [1965] and Merewitz [1971] studies on regulated industries as indicated in the previous chapter.

Thus, for each computer center j we have

$$Q_j = A \prod_{i=1}^{n} X_{ij}^{\alpha_i} e^{U_{oj}}$$

where Q_j is the output of computer center j and X_{ij} represents the input of factor i to the process of the j^{th} center. U_{oj} is a normally distributed random variable, with assumed expectation $E(U_{oj}) = 0$ and constant variance σ^2, indicating unpredictable variations in computer output. The disturbance is the result of unpredictable hardware, software and operator error. It is also assumed that $E(U_{oj}\ U_{ok}) = 0$ for all $j = k$.

If the assumption of homogeneous non-human capital is dropped and we accept the idea that technical change is embodied in successive vintages of computers which become more efficient over time, the inclusion of Solow-neutral technical change is possible. While the first delivery date of the type of computer is known, the actual installation date is not and numerous field modifications occur which make the actual vintage of the machine uncertain. Since all the machines in the sample are second generation computers, they are assumed to be of the same vintage.

3. Output from the Production Process

It is assumed that the computer produces a single homogeneous product. This product will be measured in physical units which will be called "units of machine computation." These physical units can be thought of as the amount of "computation" a representative computer would complete in a specified unit of time. Alternatively, it is possible to conceive of this unit of output as a flow of standardized tasks performed in a specified unit of time.

This study evaluates the "through-put" of the computer. The through-put is the amount of work the computer performs during a given period of time. Various definitions have been proposed to measure computer through-put. These methods include actual job comparisons, benchmarks, program kernels, and instruction mixes.

Actual job comparisons comprise the most fundamental method of defining the output of a computer. This method involves using particular tasks on which the computer center operates. These tasks, in the form of actual programs and the data which are operated upon, are processed on the computers which are under investigation. Each program is run on the computer and the amount of time required to complete the task is determined.

This method is often used in business data processing applications such as payroll. Since tasks of this form are often coded or programmed in a standard language--CØBØL--it is possible to compare directly the number of actual payroll runs which could be completed in a unit of time. It is also possible to do some scientific tasks such as matrix inversion in this same fashion, if the program were written in the FORTRAN or ALGOL programming language. However, even these standard high level languages have inconsistencies. (For example, some instructions are more efficient on one machine than another, and in some cases certain instructions are available on particular computers only.) These differences stem from variations in design characteristics and thus the standard language programs operate with various degrees of efficiency, or in some cases do not operate at all on various computers.

If one were to use this method of actual job comparison, one would be faced with three basic problems. The most important problem would be the cost of actually running the programs on the several computer systems. A second problem is to model the influence of the particular set of instructions which comprise the program in the output measure; for example, some machines allow the use of buffering, which is the ability of the computer to read input information and store it while operating upon data previously stored in the computer. If a test program were written using this instruction, the machine which does not contain the feature in its instruction set could not in general execute any of the program's instructions. The output would be zero. To attempt optimal programming for all machines would be costly. The third difficulty is to determine the generality or applicability of a particular job to the whole range of jobs processed on different computers. Since the specified job may not be representative of the whole class of jobs performed, the measure of output may be suspect. In any event, further analysis of the generality of the particular program would be necessary. For these reasons, the actual job comparison technique was not used in the study.

A benchmark is a carefully defined problem which is coded and then timed on various machines. This benchmark program becomes the numéraire for computer output. The number of these benchmarks which a computer can process in a unit of time is the output of the computer system. Appendix A contains basic examples of five separate components of a benchmark. These could be combined or weighted by the relative frequency of occurrence of each component. These frequencies could be determined via professional judgment or statistical analysis of past jobs on a given computer system. This benchmark method involves the specification, in minute detail, of the actual task to be performed, the data to be processed, programming of the task, and its execution on the various computers. This method is therefore costly to employ on an ad-hoc basis. One firm, Auerbach Information, Inc., publishes data on a set of benchmarks which the company has specially prepared.[1] Unfortunately, the Auerbach firm does not have benchmarks for the majority of computers in the sample.

Program kernels are programs or parts of programs such as: comparison of detail transactions with master file and sequence-checking of both files, table look-ups, and block data transfers. As a practical matter, in defining either benchmarks or program kernels, one can be as precise about the methods used to complete the task one chooses, but as more constraints are placed on such methods, the specialized features of various computers have less importance. Thus, in both kernels and benchmarks, the design of the output measure includes the subjective valuation of the analyst as to the importance of special machine capabilities and specialization. In addition, both methods not only measure the computer's production but also the ability of the programmer to produce efficient code. As with benchmark programs, no standard kernel programs are avilable which have been processed on machines of the type included in this study. The cost of such an undertaking was considered prohibitive.

Instruction mixes are combinations of individual instructions, such as: add, subtract, transfer or branch.

---

[1] Standard EDP Reports (New York: Auerbach Information, Inc., 1970).

Each instruction performs a specific task in a predictable way. Each of these discrete tasks can be timed (usually by the manufacturer of the computer or the times are available in the professional literature) and the speed of particular instructions is available. Each of these instruction times can be weighted by their relative frequency of occurrence. Such as $S = \sum_{i=1}^{n} W_i I_i$ where S is the speed of the instruction mix, $\sum_{i=1}^{n} W_i = 1$, and each W_i = the relative fequency of occurrence of instruction i, n is the number of instructions under consideration, and I_i is the time to complete the $i^{\underline{th}}$ instruction. This method would yield a measure of the speed of the individual computer under consideration. To find a measure of output of the computer, one would use $Q = 1/S$, which yields the number of such mixes per unit of time. If the instruction times were in units of microseconds, the output would be the number of instruction mixes performed in one microsecond.

It should be noted that the instruction mix technique is somewhat analogous to the previous methods of actual jobs, kernels, and benchmarks, since these are in effect collections of instructions. The primary difference between the instruction mix technique and actual jobs or kernels or benchmarks is that the instruction mix does not have a specified purpose. That is, the mix, if actually performed on a computer would not produce an inverted matrix or a transfer of a block of data. One could examine a benchmark or kernel or job and determine the relative frequency, W_i, of each instruction or class of instructions. Then, by using the weight W_i multiplied by I_i, one would have an instruction mix.

One advantage of constructing an instruction mix is that one need not actually run the mix on the computer. It is possible to obtain the times, as indicated earlier, from engineering data. Given the timings of instructions and the weights or relative frequencies of each of the instructions, the speed and output per unit of time can be determined for any individual machine. This method is less costly than the previous methods because of the saving in computer time. This method is easy to apply to a wide range of machines which may be unavailable for physical test, such as machines no longer in general use or those in the development stage and not available to the public. In these cases, the engineering data would be the only available information the economist has at his disposal. Another advantage of the instruction mix technique is that it eliminates the confounding influence of quality differentials in programming. Since computer programming skill varies, even a specific task can be done with various instructions being executed or the problem solved in a shorter time period. This confounding influence of confusing computer output and programming efficiency is suspect in either the kernel or benchmark, or actual job method. Since the instruction mix is not designed to perform a specific task, the programming ability of the analyst does not influence the output measure.

Some well-known difficulties do exist with the use of instruction mixes. The first difficulty lies in the fact that some of the actual execution times for various instructions, e.g., floating-point instruction is not constant, in practice an arithmetic mean of many sample times is typically used as the time for the instruction. A second difficulty is that some instructions are specialized in nature and not directly comparable across machines. Thus, one finds himself using a "generalized" or "representative" set of instructions. Because the weights used for each instruction are commonly based upon some form of dynamic trace of instructions on an individual machine, the appropriateness of the set of weights, W, for alternative computers is assumed to be relatively unimportant up to the class of jobs categorized as business data processing or scientific computation. A fourth problem arises since machines are designed in quite different ways. That is, different number of registers, word sizes, fixed and variable length words, and single- and multiple-address logic, exist across machines.[2] Even though these problems

[2]R. A. Arbuckle, "Computer Analysis and Thruput Evaluation," Computers and Automation (January, 1966), pp. 12-19.

exist, in this study a method analogous to the instruction mix technique has been chosen because this method uses data on machine performance which are available and the other methods require operational data which are not available.

The method selected includes the use of the fixed point addition instruction time, storage-retrieval cycle time, and word size. The complete add time is the time required to acquire from memory and execute one fixed point add instruction using all features such as overlapping memory banks, instruction look-ahead and parallel execution. The add is either from one full work in memory to a register, or from memory to memory; but not from register to register. Thus, the add time is the number of microseconds normally required to perform one addition, of the type a = b + c, upon fixed point operands at least 5 decimal digits (or an equivalent number of bits) in length. All the execution times include the time required to access both operands from working storage and store the result in working storage. This insures valid comparisons between computers with one address, and multiple-address formats.[3]

Storage cycle time in the context of this study is related to internal core storage only. For core storage, cycle time is the total time to read and restore one storage

---

[3] Auerbach Computer Technology Reports (New York: Auerbach Information, Inc., 1969), Chapter 11, p. 2.

word. It is the minimum time interval between the starts of successive accesses to a storage location. This measure must not be confused with access time. Access time is the interval of time between the instant when the computer or control unit calls for a transfer of data to or from a storage device and the instant when this operation is completed. Thus, access time is the sum of the time interval

1. when the computer or control unit calls for transfer of data and the beginning of transmission, and
2. the time it actually takes to transfer the data.

Cycle time is composed, in part, of access time. In addition, it includes the time to restore the original data read. Cycle time in general will be longer than access time by the amount of time needed to rewrite the work just read before another read operation can be initiated.[4]

The length of each computer word will be defined as the word size. Word size is expressed in terms of the maximum number of binary digits or decimal digits or alphanumeric characters the computer word can accommodate. It is similar to the maximum number of numbers or letters which can be typed on a single line on a typewriter. Just as some typewriters have longer or shorter carriages, so

---

[4]Computer Characteristics Review (Watertown, Massachusetts: Key Data Corporation, 1969).

computers have different size words. In some computers the word size is not fixed but is variable. For variable word length computers, data is usually presented in the form of the number of bits, digits, or characters which comprise a byte. The timing of these variable length instructions is on the basis of bytes.

The method used to determine computer output can be presented by:

$$OUTM_{ij} = \{[CT_i^{-1}w_1 + AT_i^{-1}w_2]*WS_i^{-1}\}*HRS_j$$

where $OUTM_{ij}$ is the output of the $i^{\underline{th}}$ type of computer at the $j^{\underline{th}}$ installation where CT_i is the cycle time in microseconds for the $i^{\underline{th}}$ computer, AT_i is the add time in microseconds for the $i^{\underline{th}}$ computer and WS_i is the word size in bits for the $i^{\underline{th}}$ computer, and HRS_j is the number of hours the $j^{\underline{th}}$ installation operates its computer per month. The term CT_i^{-1} or AT_i^{-1} indicates the number of cycles or additions the computer could process in one millionth of a second. The number of cycles in one hour would be $CT_i^{-1}*(3.6*10^9)$, the elimination of the constant term $(3.6*10^9)$ in the output measure does not harm our results since it is simply a scaling of the output variable. As stated earlier, the constant term in the production function includes this constant scaling factor.

The output measure is a combination of three influences in the design of computers. We are concentrating on the output of the central processing unit, CPU, of the computer. Since much of the total system (CPU, storage modules, peripheral units such as printers, types, drums, and disks) is dependent upon the output of the central processing unit, we are justified in considering this somewhat restricted view of the computer system. The method used produces numerical values which are similar to measures produced by the alternative techniques of actual jobs, kernels, and benchmarks. For example, a group of computer specialists using a method of actual job comparisons found the relative output of two IBM systems (370/155 and 360/75) to be in the ratio of 1.60 to 1. That is, the IBM 370/155 would do 1.6 units of work in the time it would take to produce 1 unit of work on the IBM 360/75 system.[5] Using the adopted measure of output, the IBM 370/155 would produce 1.345 units of computation and the IBM 360/75 would produce .75 units in one microsecond. Thus the relative ratio of speeds would be 1.79 to 1. Since the adopted method was derived for second generation computers and the systems in question are third generation, a problem of comparability might exist but the instruction mix method appears to be acceptable on the grounds of producing measures which are consistent with other methods in current use.

[5]M. Sewald, M. Rauch, L. Rodick, and L. Wertz, "A Pragmatic Approach to Systems Measurement," Computer Decisions (July, 1971), p. 39.

The three influences of add-time, cycle time, and word size are used to measure computer output. Add-time measures the ability of the machine to process a unit of data. This processing of data is what a computer accomplishes. To the extent that a single instruction does not represent a whole set of possible processing functions, the output measure will not reflect what work is actually done; this is the basic objection to the use of any instruction mix as stated earlier. It is assumed that the add-time is representative of the set of instructions and that the divergence from a "representative" mix is small.

Cycle time is included to account for transfer, acquisition, and distribution of data internal to the computer. It is not enough to process the instruction (addition); the machine must also be able to acquire the data to operate upon and return or transfer the data prior to future operations. The cycle time measures this function and is more than a single instruction. The cycle time indicates what limits are placed upon the speed of the machine to process instructions, since the data are usually manipulated during the other processing operations.

Word size enables us to compare machines which have different word lengths. Since add time and cycle time are in units of microseconds per word, comparisons of machines with different word lengths is unacceptable. The unacceptability arises because computer words which are longer contain more information and bigger words require more time, other things equal.[6] The word size is measured in bits and the result of the multiplication of the weighted sum of add time and cycle time by WS_j^{-1} yields the speed in terms of bits per microsecond. In this way, we have removed the objection to the use of add or cycle time in computers with nonstandard word sizes.

The terms W_1 and W_2 are included to weight the importance of cycle time and add-time. Because computer engineers have not produced a scientific measure of output, it was considered interesting to see how sensitive the model would be to various sets of weights. In the empirical chapter, results are presented assuming various weights for add and cycle time.

Thus far we have determined the unit of output for a particular computer,

$$OUTM_i = \{[CT_i^{-1}W_1 + AT_i^{-1}W_2]*WS_i^{-1}\}.$$

This indicates the units of output which would be produced in one-millionth of a second; the number of units produced in one hour would be $OUTM_i*(3.6 \times 10^9)$. For a particular computer center, say, the $j^{\underline{th}}$ center, the number of units of output would be the output unit measure, $OUTM_i$,

---

[6] Gordon Raisbuk, Information Theory (Cambridge: Massachusetts Institute of Technology Press, 1966), p. 8.

multiplied by the number of hours the machine is operating, HRS_j. The number of units of output produced by the $j^{\underline{th}}$ center during the sample period of one month is given as

$$OUTM_{ij} = \{[CT_i^{-1}W_1 + AT_i^{-1}W_3]*WS_i^{-1}\}.$$

4. Factors of Production

The basic input factors are assumed to be labor and capital. These factors of production are combined to produce "computer output," which was defined in the previous section.

Labor input is probably the most important factor of production under management control which can alter the level of costs. This is the factor which management can adjust upon relatively short notice in order to meet the operational objectives of cost minimization. The labor service required in the production process is clearly differentiated into three groups.

The first group is the administrative or management group. This group is responsible for the long-term planning, day-to-day management of the facility, and the achievement of productive objectives. This first group generally is comprised of a director and several others who are responsible for functional areas such as machine operations, business matters, programming, clerical duties, teleprocessing and on-line data acquisition and control, etc.

The second group is the programming staff. This group includes those personnel who are responsible for the generation of general purpose or application programs. General purpose programs are used numerous times, usually with different sets of data and control cards. Examples of these are statistical packages, mathematical programming, circuit analysis, and numerical analysis algorithms. Also in this group are systems programmers. Systems programmers are responsible for generation and maintenance of the set of programs which control the operation of the computer. These programs control the flow of work to the computer, assign various hardware components of the computer system to specific tasks, call standard routines for programs, and maintain accounting records of the computer's operation. Systems programmers also are typically responsible for the various language processors (compilers/assemblers) which convert user written programs to a form the computer can "understand." Because the maintenance of these programs requires a knowledge of the operating system and the machine language used to produce parts of the compilers, the system programmers are in charge of this task. Programmers also act as consultants to those who are writing their own programs and need assistance.

The third category includes the operations staff. Operations staff include such tasks as key-punching, general tab-room card preparation, clerical work (both administrative and those tasks related to physical handling of card inputs to and outputs from the facility), computer machine operators, and maintenance personnel. These three groups of employees--administration, programming, and operations--are the components of the labor service input. They are directly controllable and their use rate can be altered to meet the center's objective.

Capital usage includes the service of the electronic digital computer and the associated equipment such as the tape disk and drum devices, the on- and off-line input/output units, and communication modules. In addition, the capital usage includes the physical building, air conditioning, office space, tab-room and keypunch equipment.

Since computers are often rented or leased under agreements of an original 24 month lease and open termination on 6 months to 30 days' notice, facilities often have the opportunity of changing machine models or manufacturers within a short planning horizon. Although it is possible to adjust the use of capital equipment by switching to alternative manufacturers, such a procedure is less likely to occur in practice. A locked-in effect results after a particular manufacturer is first selected. This locked-in effect occurs because programs are specialized factors and cannot easily be changed to function on an alternative manufacturer's machine. Further, specialized program packages usually take advantage of special features of one machine which are not available on other machines. This locked-in effect is not directly considered in the model and it is assumed that such influences do not drastically influence management planning.

5. Output and Input Price

The previous discussion centered on the output and inputs of the stochastic production process; this section is about the unit price of these variables. Prices of the factors of production are included in this study but the price of computer output is not.

In the context of university computer centers, it is appropriate to assume that factor prices are determined exogenously from the production model. The wage of the factor is assumed to be constant for any one computer center and not related to the rate of use of that factor. It is also assumed that wage levels for various categories of labor service, administration, programming and operations are determined in the local geographic area where the center is located. These prices vary from center to center because of geographic immobility of these employees.

The next price to consider is the cost of capital for the university computer center. We are concerned with the unit cost of capital or the unit price of capital. The most probable assumption seems to be that cost is determined on a national level in terms of a general interest rate structure, given the risk of default. This price of capital is assumed constant for all universities and colleges in the sample. If any differences exist between universities, it would be due to special tax treatment of the bonds issued in the states, market imperfections, or the risk of the issuing agent and not systematically related to the operation of the computer centers in question.

Computer centers in institutions of higher education do not directly charge individual users for computer output. The center could use a scheduling algorithm or priority system which gives better service, in terms of faster production, to various users. By adjusting the length of time between submission of execution of a program, the center can encourage or discourage computer usage. This is equivalent to charging an implicit price of the output to the user since waiting time is a cost. This implicit price influence is assumed to be small with respect to total demand for computer output and such scheduling systems only alter the relative number of programs in each job category. Such a change in mixture of jobs is assumed to be independent of the total production of the facility. Under these assumptions, output price is not controlled by the computer center. In this respect, the installation is in the same position as the firm in a regulated industry where the rates are determined by the regulator's administrative fiat. In the next section, we will discuss the effect that the lack of output price has on the construction of a model for the production process. It will be shown that it is possible to model the process without the use of output price information.

6. Behavioral Constraint

The computer center is regarded as a firm with the objective of minimizing the cost of producing an expected level of output. The assumed objective of cost minimization is preferable to profit maximization because the computer center does not charge users for computer output and thus does not produce revenues. Cost minimization is also preferred to the alternative assumption of output maximization for a given level of cost. The demand for computer output is generated by classroom problems, computer aided instruction, professional research, and administrative data processing. These demand-generating influences are exogenous and not controllable by the management of the computer facility.

Computer processing cannot be stored or held in inventory. It is not possible to produce in anticipation of future output or peak demand periods. Because of this fact, it is not necessary to consider any behavior constraints related to product inventory management.

It is unreasonable to assume a manager will maximize a quantity which is undefined in the case of profits, or is not under his control in the case of output. Since the demand for output and factor prices are determined outside the production process, the manager has one tenable behavior objective which is to adjust the rate of usage of the factors of production such that costs are minimized.

If cost minimization is selected as the appropriate behavioral objective for the computer center management, then we will not require any information regarding the price of computer output. Cost minimization for a given level of expected output will result in the same utilization of inputs independent of the selling price of the output. Factor usage is dependent on production function and resource costs which are assumed to be independent of output price or the relation between price and marginal revenue. The next section is the specification of the cost minimization model and one should note that both the first and second order minimization conditions are not influenced by a change in product price if the level of expected output is unchanged. If the expected level of output changes, only the absolute level of resource use is altered and input ratios are constant for homogeneous production functions.

7. Specification of the Cost Minimization Model

This section contains a model of the technical relation between the inputs and output and the behavioral constraints presented in previous sections. It should be remembered that the objective of the computer center is to minimize cost subject to a given level of expected output.

The basis of the cost model can be seen from the following analysis of the production system. Given

$$R_t : Q = f(X_i) \qquad i = 1, \ldots n \text{ inputs},$$

where the operator f is a set of rules to obtain the maximum output from various inputs.

If we assume pure competition in the factor markets, we have

$$R_e : \frac{P_i}{P_j} = g_f \ (X_j, X_i) \ i > j; \qquad \begin{array}{l} i = 2, 3, \ldots n \\ j = 1, 2, \ldots, n\text{-}1. \end{array}$$

R_e is the economic rule which indicates how to obtain the lowest cost for a given level of output. The g_f operator, which restricts our use of inputs to those which minimize cost for a given level of output, has the subscript f to note that the cost rule must incorporate the production rule. These n - 1 equations determine factor demands. The rule D_c is the definition of total cost and can be written as

$$D_c : C = \sum_{i=1}^{n} P_i X_i \qquad i = 1, \ldots n.$$

The cost equation and the notation used above allows us to demonstrate the relationship between cost, predetermined output, and prices:[7]

$$C = f\,(Q, P_1, P_2, \ldots, P_n).$$

The cost function includes the definition of cost, the technical information from the production function, and the cost minimization rule. The next step is to define the operations in specific functional forms.

Beginning with the Cobb-Douglas production function which was introduced in Section 2, we have

$$Q_j = A \prod_{i=1}^{n} X_{ij}^{\alpha_i} e^{U_{oj}}$$

where the subscript j refers to the $j^{\underline{th}}$ computer center and U_o is a random variable normally distributed with zero mean

---

[7]The following procedure leads to the Shephard-Uzawa duality theorem discussed in Chapter II. (1) Rewrite R_e for each input r in terms of one of the remaining inputs as $X_r = g_f^r\,(P_i X_i)$ which is the constant output factor demand schedule for X_r. (2) Rewrite R_t as $Q = f^*\,(g_f^r)$, the f* maps the factor demand schedules to output. (3) Solve the cost definition D_c in terms of output and factor prices. Transposing the previous relation and solving for g_f^r, we have here $g_f^r = h^r(Q)$ for all inputs. Using the definition of D_c, we have

$$C = H[h^r(Q), P_1, P_2, \ldots P_n] \qquad \text{or} \qquad C = f(Q,p).$$

and variance σ^2 and $E(U_{oj} U_{or}) = 0$.[8] The remaining symbols have been defined earlier. The rule for obtaining the lowest cost for a given level of expected output can be obtained from the Lagrange technique as follows. From the Lagrangian function

$$L = \sum_{i=1}^{n} P_i X_i - \lambda [E(Q) - \overline{Q}_o].$$

Since output is stochastic, as explained in Section 2, the mathematical expectation of output is entered in the constraint as E(Q). The first order conditions for constrained cost minimization are

$$\frac{\partial L}{\partial X_i} = P_i - \lambda \frac{\partial E(Q)}{\partial X_i} \qquad i = 1, \ldots n.$$

Recall that the inputs have a perfectly elastic supply for each center and the prices are known with certainty as stated in Section 5. We will assume for the moment that the marginal productivity conditions are deterministic and

---

[8] The assumption of normality of the production function disturbance may create some estimation problems. The multiplicative log normal disturbance used in the transformation of the production relation may cause attention to be shifted to the conditional median rather than conditional mean which is of interest here. Goldberger states that in practice the minimum variance unbiased estimators which account for this fact may not differ detectably from those which do not, and we assume the simpler specification in this case. For more information, see J. Goldberger, "Interpretation and Estimation of Cobb-Douglas Functions," Econometrica, Vol. 36 (July-October, 1968), pp. 464-72; and A. Zellner, J. Kmenta, and J. Dréze, "Specification and

and that adjustments to optimal levels of inputs are instantaneous. After eliminating λ, we obtain (n-1) independent relations

$$\frac{P_k}{P_i} = \frac{\alpha_k}{\alpha_i} \cdot \frac{X_i}{X_k} .$$

The whole system then consists of n + 1 equations

$$\frac{P_k}{P_i} = \frac{\alpha_k}{\alpha_i} \frac{X_i}{X_k} \qquad i \neq k,\ k = 1, \ldots (n-1)$$

$$C = \sum_{i=1}^{n} P_i X_i$$

$$Q = A \prod_{i=1}^{n} X_i^{\alpha_i} e^{U_o}$$

with unknowns X_i, i = 1, ... n and c.

We have completed Re, the first-order condition for a relative extrema point, assuming that this is also the global extrema in the constrained feasible region. Now we proceed to the second-order conditions. To economize on exposition, only the two factor case will be considered in detail, but the generalization to n inputs is also discussed. Defining f as $Q = f(x_1, x_2)$, the second-order condition for constrained minima is that

---

Estimation of Cobb-Douglas Production Function Models," *Econometrica*, Vol. 34 (July-October, 1966), pp. 784-95.

$$\begin{vmatrix} f_{11} & f_{12} & -P_1 \\ f_{21} & f_{22} & -P_2 \\ -P_1 & -P_2 & 0 \end{vmatrix} > 0 .$$

Where $f_i = \frac{\partial f}{\partial X_i}$, $f_{ij} = \frac{\partial f}{\partial X_i \partial X_j}$ and

since $P_i = \lambda f_i$, we can multiply column 3 by $1/\lambda$, row 3 by $1/\lambda$ and the determinant by λ^2, and we preserve the initial value of the determinant which is

$$\lambda^2 \begin{vmatrix} f_{11} & f_{12} & -f_1 \\ f_{21} & f_{22} & -f_2 \\ -f_1 & -f_2 & 0 \end{vmatrix} > 0 .$$

We need only examine the value of the bordered Hession determinant, since $\lambda^2 > 0$. Thus,

$$2f_{12}f_1f_2 - f_{11}f_2^2 - f_{22}f_1^2 > 0 .$$

Given positive marginal products from the first-order conditions, $f_{ii} < 0$ (marginal products decreasing) is a sufficient condition for second-order stability, but not a necessary condition. What is necessary is that the isoquant is concave from above in all directions, which is equivalent to establishing diminishing marginal rate of technical substitution. This is demonstrated next.

Define MRTS $= \left[\frac{f_1}{f_2}\right]$. To demonstrate that $\frac{\partial MRTS}{\partial X_i} < 0$ for all i is equivalent to the second-order condition, we write

$$\frac{\partial\left[\frac{f_1}{f_2}\right]}{\partial X_1} = \frac{\left[f_2\left(f_{11} + f_{12}\left[\frac{\partial X_2}{\partial X_1}\right]\right) - f_1\left(f_{12} + f_{22}\left[\frac{\partial X_2}{\partial X_1}\right]\right)\right]}{f_2^2} < 0 .$$

Since $\frac{\partial X_2}{\partial X_1} = \left[-\frac{f_1}{f_2}\right]$, i.e., the slope of the isoquant,

$$\frac{\partial\left[\frac{f_1}{f_2}\right]}{\partial X_1} = \frac{1}{f_2^2}\left[f_2\left(f_{11} + f_{12}\left[-\frac{f_1}{f_2}\right]\right)\right] - f_1\ f_{12} + f_{22}\left[-\frac{f_1}{f_2}\right]\Bigg] < 0 .$$

Multiplying by $\left[\frac{f_2}{f_2}\right]$, we have

$$\frac{\partial\left[\frac{f_1}{f_2}\right]}{\partial X_1} = \frac{1}{f_2^3}\left[f_2^2 f_{11} - 2f_{12}f_1f_2 + f_{22}f_1^2\right] < 0 .$$

With $f_2^3 > 0$, the bracketed term is the negative of our second-order condition and we have demonstrated our objective. The necessary condition for second-order stability is diminishing marginal rate of technical substitution. It is also notable that this condition is the necessary condition for the Shephard-Uzawa duality theorem. The reader interested in comparing the cost minimization with the profit maximization case is referred to Appendix B.

Considering the Cobb-Douglas production function, the second-order condition is

$$[f_2^2 f_{11} - 2f_{11}f_1f_2 + f_{22}f_1] < 0$$

which is equivalent to

$$\frac{Q^3\theta^2}{X_2^4}[(\alpha_1^2 - \alpha_1)\alpha_2^2 + (\alpha_2^2 - \alpha_2)\alpha_1^2 - 2(\alpha_1^2\alpha_2^2)] < 0$$

where $\theta = \left[\frac{P_1}{P_2} \cdot \frac{\alpha_2}{\alpha_1}\right]$. The quantity $\frac{Q^3\theta^2}{X_2^4}$ is positive when α_1 and α_2 have the same sign and we need only examine

$$[(\alpha_1^2 - \alpha_1)\alpha_2^2 + (\alpha_2^2 - \alpha_2)\alpha_1^2 - 2\alpha_1^2\alpha_2^2] < 0$$

which reduces to

$$[\alpha_1\alpha_2^2 + \alpha_1^2\alpha_2] > 0 .$$

The second-order condition is satisfied for the Cobb-Douglas production function if $\alpha_1\alpha_2 > 0$. No restriction is placed on the sum of the output elasticities as compared to profit maximization where the sum is restricted.

It should be noted that no allowance has been made for institutional or other restraints on the cost minimizing activity of the computer center. This seems justified since little if any union activity or other administrative intervention in the management of the centers existed. This clearly is not the case for other regulated industries and in fact is the basis of the Averch-Johnson (1962) model of the firm under rate of return regulation.

Examining the cost minimizing rule more closely, we note that we have implicitly assumed the desired amounts of inputs to be those which are actually used. The rule R_e becomes more realistic and complicated when the desired and actual quantities of inputs are not identical. We could expand the nature of the g_f operator, but g_f would lose some of its economic interpretation as a deterministic procedure for management to follow. A much more interesting approach is to represent the cost minimizing conditions in terms of desired levels of inputs.

A partial adjustment relation is assumed to exist between the actual and the desired levels of inputs. The desired values are denoted X_i^* and are not directly observable, but the actual values X_i are presumably being adjusted to X_i^*. The explanation of the incomplete adjustment of X_i is that adjusting inputs has a cost and it takes time to adjust factors [see Griliches (1967)]. It is possible to write the partial adjustment model as:

$$X_{i,t}/X_{i,t-1} = (X_{i,t}^*/X_{i,t-1})^{\gamma_i} e^{U_{i,t}}$$

where $U_{i,t}$ is a random influence in adjusting the actual to the desired levels of input. It is due to randomness in the availability of extra or overtime labor or unknown variability in transaction time to sell or acquire inputs. $U_{i,t}$ is normally distributed with zero mean and constant variance σ_i^2, and $E(U_{i,t}U_{j,t}) = 0$, $i \neq j$. The $U_{i,t}$ terms are assumed to be statistically independent of $U_{o,t}$, the production function disturbance. This independence may be justified on the grounds that the production function disturbance originates from computer hardware and operator failure and is technical in nature but the $U_{i,t}$ terms are the result of stochastic nature of factor availability and transaction time. The adjustment coefficients are defined in the bounded region where $0 < \gamma_i \geq 1$ and γ_i is the rate of adjustment of X_i to X_i^*. If we solve the equation for X_i^*, we have

$$X^*_{i,t} = X_{i,t}^{\left[\frac{1}{\gamma_i}\right]} X_{i,t-1}^{\left[\frac{\gamma_i - 1}{\gamma_i}\right]} e^{\left[\frac{-U_{i,t}}{\gamma_i}\right]}$$

which expresses the unobservable $X^*_{i,t}$ in terms of current and lagged observable X_i's. The marginal productivity conditions in terms of observable quantities then become

$$\frac{P_i}{P_j} = \frac{\alpha_i}{\alpha_j} \frac{X_{j,t}^{\left[\frac{1}{\gamma_j}\right]} X_{j,t-1}^{\left[\frac{\gamma_j-1}{\gamma_j}\right]} e^{\left[\frac{U_{i,t}}{\gamma_i}\right]}}{X_{i,t}^{\left[\frac{1}{\gamma_i}\right]} X_{i,t-1}^{\left[\frac{\gamma_i-1}{\gamma_i}\right]} e^{\left[\frac{U_{j,t}}{\gamma_j}\right]}} \qquad i \neq j$$

The inputs are three types of labor, and it is reasonable to assume that the costs of adjustment are equal. This assumption will result in γ's which are similar. Consider the case where $\gamma_i = \gamma_j$, the marginal productivity condition

$$(1) \qquad \frac{P_i}{P_j} = \frac{\alpha_i}{\alpha_j} \frac{X_j^*}{X_i^*}$$

$$(1') \text{ becomes } \qquad \frac{P_i}{P_j} = \frac{\alpha_i}{\alpha_j} \left[\frac{X_{j,t} X_{j,t-1}^{\gamma-1} e^{U_{i,t}}}{X_{i,t} X_{i,t-1}^{\gamma-1} e^{U_{j,t}}}\right]^{\frac{1}{\gamma}} .$$

Let us return to equation (1) and note that the marginal productivity conditions are deterministic in desired inputs. If managerial error, inertia, resistence to change and the like are present, we find that the equation determining the desired factor input has an additional disturbance, $U_{i,t}^{+} \sim N(0,\sigma_{+}^{2})$, for the (n-1) marginal productivity equations. Equation (1') becomes

$$(1'') \quad \frac{P_i}{P_j} = \frac{\alpha_i}{\alpha_j} \left[\frac{X_{j,t} X_{j,t-1}^{\gamma-1}}{X_{i,t} X_{i,t-1}^{\gamma-1}} \right]^{\frac{1}{\gamma}} e^{\left(\frac{U_{i,t} + U_{j,t} + \gamma U_{i,t}^{+}}{\gamma}\right)}$$

which is the general case. The special case of instantaneous adjustments yields

$$(2) \quad \frac{P_i}{P_j} = \frac{\alpha_i}{\alpha_j} \left[\frac{X_{i,t}}{X_{j,t}} \right] e^{(U_{i,t} + U_{j,t} + U_{i,t}^{+})}$$

With instantaneous adjustments the equations remain stochastic. Since we are concerned with the possible bias introduced via the stock adjustment of inputs, the cost functions will be derived using both equations (1") and (2). In the data analysis chapter the possible bias introduced by the omission of relevant lagged explanatory variables will be considered.

First, considering the instant adjustment case the marginal productivity condition can be rewritten as

$$X_{i,t} = \left[\frac{P_j}{P_i} \right] \left[\frac{\alpha_i}{\alpha_j} \right] X_{j,t} \, e^{(U_{i,t}^{*} + U_{j,t})}$$

where $U_{i,t}^{*} = U_{i,t} + U_{i,t}^{+}$. Deriving the fixed output demand function for X_i using the production function, we have

$$X_{i,t} = Q^{\frac{1}{r}} A^{-\frac{1}{r}} \left(\frac{\alpha_j}{\alpha_i}\right)^{\frac{-\alpha_j}{r}} \left(\frac{P_i}{P_j}\right)^{\frac{-\alpha_j}{r}} e^{\left[\frac{\alpha_j (U^*_{i,t} + U_{j,t}) - U_{o,t}}{r}\right]}$$

where $r = \sum_{i=1}^{n} \alpha_i$. This equation is homogeneous of degree zero in prices, which is the desired result. The final procedure is to solve for total cost in terms of input prices and output. For the instantaneous adjustment case, we have

$$C_t = \left[r(A \prod_{i=1}^{n} \alpha_i^{\alpha_i})^{-\frac{1}{r}}\right] Q_t^{\frac{1}{r}} \prod_{i=1}^{n} P_i^{\frac{\alpha_i}{r}} \{e^{\frac{-U_{o,t}}{r}} (\sum_{i=1}^{n} e^{(U^*_{i,t})})\} .$$

Taking logarithms of both sides of the equation and expanding the combined stochastic disturbance term around the mean of each disturbance, via Taylor series, results in

$$(3) \quad \ln C_t = \ln \left[r(A \prod_{i=1}^{n} \alpha_i^{\alpha_i})^{-\frac{1}{r}}\right] + (\tfrac{1}{r}) \ln Q_t$$

$$+ \sum_{i=1}^{n} \left[\frac{\alpha_i}{r}\right] \ln P_i + \{ \sum_{i=1}^{n} U^*_i - \frac{U_o}{r} + \ln (2)\}$$

For the non-instantaneous adjustment where $0 < \gamma < 1$ the constant output factor demand function is

$$X_{i,t} = Q_t^{\frac{1}{r}} A^{\frac{-1}{r}} \left[\left(\frac{\alpha_j}{\alpha_i}\right)^{\gamma} \left(\frac{P_i}{P_j}\right)^{\gamma} \left(\frac{X_{i,t-1}}{X_{j,t-1}}\right)^{\gamma-1}\right]^{\frac{\alpha_2}{r}} e^{\frac{-U_o}{r}} e^{\frac{\alpha_j (U^*_i + U_j)}{r}}$$

where $U^*_{i,t} = u_{i,t} + \gamma U^{+}_{i,t}$.

The cost function is

$$C_t = \left[r^{\gamma} \left(A \prod_{i=1}^{n} \alpha_i^{\alpha_i}\right) \right]^{\frac{-\gamma}{r}} Q_t^{\frac{1}{r}} \, Q_{t-1}^{\left(\frac{\gamma-1}{r}\right)} \prod_{i=1}^{n} P_i^{\frac{\alpha_i}{r}} \{ e^{\frac{-U_o}{r}} \left(\sum_{i=1}^{n} e^{U^*_i} \right) \}$$

$$\cdot \left(\sum_{i=1}^{n} X_{i,t-1}^{(1-\gamma)} \right) \left(\sum_{i=1}^{n} P_i^{(1-\gamma)} \right) .$$

Taking logarithms of both sides of the cost function handling the error term as in (3) and approximating the lagged input and price summation terms via Taylor expansion around $\gamma = 1$ and dropping higher order than on terms of γ, we have

$$\text{(4)} \quad \ln C_t = \{ [(n+1) \cdot \ln(2)] + [\ln r^{\gamma} \left(A \prod_{i=1}^{n} \alpha_i^{\alpha_i} \right)^{\frac{-\gamma}{r}} \}$$

$$+ \left(\frac{1}{r}\right) \ln Q_t + \left(\frac{\gamma-1}{r}\right) \ln Q_{t-1}$$

$$+ \sum_{i=1}^{n} [(\frac{\alpha_i}{r} + \frac{\gamma-1}{2}) \ln P_i] + \sum_{i=1}^{n} [(\frac{\alpha-1}{2}) \ln X_{i,t-1}]$$

$$+ \{ \sum_{i=1}^{n} U^*_i - \frac{U_o}{r} \} .$$

Equation (3) is the single period cross-section model and equation (4) includes observations on the previous period inputs and output. Of course, if we have an equilibrium situation where $\gamma = 1$, equation (3) would be equivalent to (4). The apparent difference in the intercept is influenced by the remainder in the Taylor series which is not zero. The effect of the omission of possibly relevant explanatory lagged variables is examined in the next chapter on data and estimation.

8. Summary

This chapter considered the problem of specifying a model, the production of computer output. The maintained hypothesis of cost minimization subject to the Cobb-Douglas production function was presented. Stochastic disturbances were assumed in the production function, the factor demand schedules, and the stock adjustment process for desired versus actual level of factor employment. Various inputs to the production process were discussed in terms of their relevance to the process. In addition, a method of measuring the output from the computer center has been developed. This method appears to be satisfactory in comparison to the high cost alternatives.

It has been demonstrated that a reduced form cost function can be derived from the production function and marginal productivity conditions which result from the cost minimization objective. This function is linear in the logarithms of cost and the explanatory variables of output and factor prices.

It was also shown that this cost function is consistent with the Shephard-Uzawa duality principle and thus contains all the information from the production function. Thus we can test the hypothesis of economies of scale via empirical estimation of the parameters of the model. This estimation is the subject of the next chapter.

CHAPTER IV

DATA AND EMPIRICAL RESULTS

1. Introduction

This chapter deals with the availability of data and procedures used to relate the model to the data. In the light of the needs for data on physical output, dollar costs and factor prices, it is shown that such data exists for a cross section of computer centers. Section 2 contains a discussion of the sources of data. A description of the methods by which these data were obtained is in Section 3, since these methods may influence the interpretation of the empirical results. Section 4 contains a discussion of the cost function and the statistical procedures employed to estimate the parameters of the model and test the hypothesis of economies of scale. The fifth section is an analysis of the statistical bias which is introduced by the omission of relevant lagged explanatory variables. The empirical results of the estimation procedure and hypothesis test are presented in the sixth section. Section 7 is a summary of the chapter.

2. Sources of Data

In the first quarter of 1966, the National Science Foundation contracted with the Southern Regional Education

Board to collect data on the Computer Sciences Project. This project concerned the very rapid expansion of the computer facilities of institutions of higher education. Some government officials became aware of the increased demand for trained computer scientists and technicians as complementary factors of production in the defense and space industries. The CSP entailed the development and testing of a questionnaire "which could be used to provide the kind of information needed for future planning of relevant Government agencies."[1] The study was to determine the sources and uses of funds for instructional activities and research in colleges and universities in the United States. An inventory of computers was also prepared. The fiscal year 1965 was used as the sample point for all actual expenditures which are the concern of this study.

The data available are the 1964-1965 expenditures for equipment rentals, rental or amortized cost for building space to house computer activities, and maintenance costs not included in the previous two categories. These three categories are summed for each installation to form the total capital cost of operating the center. Added to this are the salaries and wages of all personnel, other direct costs of materials and supplies. Unit wage rates for three

---

[1] I. Hamblin, Computers in Higher Education (Atlanta: Southern Regional Educational Board, 1967), p. 21.

categories of employees--administrative and other professions, systems and utility programmers, and all others (keypunch, machine operators, clerical, and technicians)--are available for each installation in the fiscal year 1964-1965. Each institution reported the number of hours the facility was operated in a typical month.

Since it is assumed that the cost of capital for all computer centers is constant, the true cost function can be written as

$$(4.0)\quad \ln C_t = A^* + \left(\frac{1}{r}\right) \ln Q_t + \left(\frac{\gamma-1}{r}\right) \ln Q_{t-1} + \sum_{i=1}^{3} \left[\left(\frac{\alpha_i}{r} + \frac{\gamma-1}{2}\right) \ln P_{i,t}\right] + \sum_{i=1}^{4} \left[\left(\frac{\gamma-1}{2}\right) \ln X_{i,t-1}\right] + V_t$$

where $V_t = \{\sum_{i=1}^{3} U_i^* - \frac{U_o}{r}\}$ and

$$A^* = \left(\frac{\alpha_n}{r} \ln \bar{P}_n\right) + 2.76 + \left[\ln r^{\gamma} \left(A \prod_{i=1}^{4} \alpha_i^{\alpha_i}\right)^{\frac{-\gamma}{r}}\right.$$

where capital is the n^{th} factor and $\bar{P}_n$ is the constant cost of capital.

Technical engineering data on the machine characteristics of various computers are available. The source of these data includes the trade journals for data on new machines.[2] Specialized publications contain detail performance data for a wide range of computers.[3] The manufacturer of the computer publishes performance data.[4] It is possible to perform experimentation on computer systems and measure performance on a case-by-case basis.[5] Fortunately, all the computers in the sample were listed in the specialized publications. Data on word size, add time, and cycle time for each machine were obtained from Computer Characteristics Review. With the questionnaire data on the number of hours the computer was in operation, it was possible to construct an output measure for all computer centers using the method presented in Chapter III.

3. Method of Data Collection

Since the data on cost, hours of operation, and factor price were collected via questionnaire, it seems appropriate to discuss the methods used. In early 1966,

---

[2]Such as Electronics News (New York: Fairchild Publishing Co., 1972); or Computer Components Review (Norwood: Commander Publishing, 1969); or Data Products News (New York: Data News, Inc., 1970).

[3]Auerbach Computer Reviews and Key Data (New York: Auerbach Information, Inc., January, 1968).

[4]Such as IBM Technical Publications (White Plains: Technical Documentation Center, International Business Machines, 1970).

[5]An example of a study of this nature is the Mobile Oil Study prepared by Sewald, et. al., op. cit., comparing various IBM computers in the Mobile Oil Corporation.

the National Science Foundation let contract NSF-C465 which required the Southern Educational Board to finalize the questionnaire, disseminate it to the institutions, process the returns, and summarize the results. This questionnaire had already been developed by the Mathematical Sciences Section of NSF. The Bureau of the Budget recommended that the National Center for Education Statistics of the Office of Education draw a "stratified systematic" random sample of approximately 700 of the 2,219 institutions of higher education which existed in fiscal year 1965. The questionnaires were mailed in mid-July 1966 and follow-up reminders were sent in September, December, and the end of January, 1967. Eventually, 669 institutions responded. Since this was the first survey of sources and uses of funds for computers operated by institutions of higher education, it posed problems for those administrators who had to complete the questionnaire. Hamblin states, ". . . though the temptation to use a random number generator might have been strong at times, a high percentage of institutions made an honest effort to obtain and report accurate figures."[6] His team edited, cross-checked, and in some cases phoned the various respondents to check accuracy. He stated that machine

---

[6]I. Hamblin, "Expenditures, Sources of Funds, and Utilization of Digital Computers for Research and Instruction in Higher Education, 1964-1965 with Projections for 1968-1969," Communications of the ACM, Vol. 7 (April, 1968), pp. 257-262.

rental and salaries were within the range of plus or minus 10 per cent of the true values. Unfortunately, we have no information on the accuracy of this estimate.

Even though Hamblin checked the returns carefully of the 669 institutions responding, only 133 at most were available for this study. The other responses had to be dropped because they lacked at least one value for type of computer, operation time, salaries, or cost. In a number of cases, the center was operating first generation computers and these centers were also excluded from the study. The institutions which did not fully respond to the questionnaire may systematically cause an unknown bias to enter the estimation of the parameters. We have no way of determining if the institutions which were unwilling to submit salary, cost or utilization data are systematically related to each other or to those who gave full information. Of course, these data are of the questionnaire variety and all the customary caveats should be observed.

4. Cost Function

The cost function (4.0) can be written as

$$(4.1) \quad \ln C_j = A^* + (\tfrac{1}{r}) \ln Q_j + (\tfrac{\gamma-1}{r}) \ln Q_j^{**} + \sum_{i=1}^{3} [(\tfrac{\alpha_i}{r} + \tfrac{\gamma-1}{2}) \ln P_{i,j}] + \sum_{i=1}^{4} [(\tfrac{\gamma-1}{2}) \ln X_{ij}^{**}] + V_j \ .$$

Here the double asterisks represent lagged values of the variables for the j computer center. The remaining symbols are the same as before. Some points about the methodology are in order. First, recall that the maintained hypothesis includes the Cobb-Douglas production function and its formulation via cost minimization into the cost function (4.0) and it also includes the assumption with regard to the disturbance term.[7]

(4.2) V_j is normally distributed

(4.3) $E(V_j) = 0$

(4.4) $E(V_j^2) = \sigma^2$

(4.5) $E(V_j V_k) = 0 \qquad (i \neq k)$

(4.6) Non-stochastic explanatory variables

---

[7]It is assumed that the logarithm of the disturbance term V_j is normally distributed with zero mean and σ^2, which of course implies that the distribution of (eV_j) is log-normal in the multiplicative cost model.

In some applications the influence of this positive skewed disturbance might be undesirable, but not in this case. It is assumed that the median of the distribution is

(4.7) No exact linear relation exists between any of the explanatory variables and of course the number of observations is greater than the number of coefficients to be estimated.

Under these conditions, it is well known that the OLS estimators of A*, $(\frac{1}{r})$ and (α_i/r) have the classical desirable properties, while the estimator of A* and r obtains all those asymptotic properties of the estimator, while the small sample properties (such as unbiasedness) do not "carry over."[8]

Since the data were collected via a questionnaire, measurement errors may be present. If errors are present only in the dependent variable, cost we would have observed C_j', the true value would be C_j, and they would be related as

$$(4.8) \quad C_j' = C_j + v_j^*$$

where v_j is the error in measurement of cost. If $v_j \sim N(0,\sigma_v^2)$ and

$$(4.9) \quad E(v_j^*,v_k^*) = 0 \quad (j \neq k)$$

---

less than the mean because of the influence of some exogenous demands which place heavy or peak demands upon computer systems. For any individual system, a few peak periods such as end of terms or quarters will be reflected in skewed output distributions. These two influences reinforce the acceptability of the maintained hypothesis of positive skewedness of the output distribution in the multiplicative cost model.

[8] J. Kmenta, Elements of Econometrics (New York: Macmillan, 1971), p. 458.

(4.10) $E(V_j, v_j^*) = 0$

we would have

$$(4.11)\quad \ln C_j' = A^* + \left(\frac{1}{r}\right) \ln Q_j + \left(\frac{\gamma-1}{r}\right) \ln Q_j^{**} + \sum_{i=1}^{3} \left[\left(\frac{\alpha_i}{r} + \frac{\gamma-1}{2}\right) \ln P_{i,j}\right] + \sum_{i=1}^{4} \left[\left(\frac{\gamma-1}{2}\right) \ln X_{i,j}^{**}\right] + \eta_j$$

where $\eta_j = V_j + v_j^*$, with $\eta_j \sim N(0,\sigma_N^2)$ and $\sigma_n^2 = \sigma^2 + \sigma_{v*}^2$. Thus (4.11) would be equivalent to (4.0), estimators having the desirable properties as given, with the note that the disturbance (η_j) contains disturbances which influence the economic relations in the cost function (V_j) and those due to measurement errors (v_j^*). Because the measurement of total cost includes many components, some error is likely to occur; it is assumed that no error in measurement of the independent variables exists.

It seems justifiable to assume that the independent variables are free of measurement error on two grounds. First, output has been defined in an engineering or scientific way. This methodology of "operationism" is accepted in engineering.[9] Of course, the answer to the

[9] P. W. Bridgeman, The Logic of Modern Physics (New York: Macmillan, 1928); and The Nature of Physical Theory (New Jersey: Princeton University Press, 1936).

question of the "correct" definition of computer output is unknown. Second, the data on salaries are quite good because of the institutional requirements such as federal tax and internal auditing requirements. If the independent variables included significant measurement errors, OLS estimation would lead to inconsistent estimators of the parameters. We could use the method of instrumental variables to obtain consistent estimators, but no instruments are available. It seems best to assume no errors in measurement or errors only in the cost data.

The concern of the next two sections is to present an analysis of omitted variables and the empirical results of the statistical estimation of the output elasticities and to test the hypothesis of economies of scale. It should be noted before we present these results that the maintained hypothesis includes all the assumptions which have been made but not subject to test and the results are conditioned on these assumptions. The appropriateness of using indirect least squares estimation and the properties of the estimators which result are also dependent on the maintained hypothesis. The results of Section 6 present the multiplicative inverse of the sum of the output elasticities ($\frac{1}{r}$) and the individual output elasticities. In addition, the confidence interval for the sum of the output elasticities is also presented. The hypothesis of economies of scale is tested using a t-test as follows.

The null hypothesis is that constant returns to scale exist, which occurs when the inverse of the sum of the output elasticities ($1/\Sigma a_i$) equals unity. The alternative hypothesis is that economies of scale exist. That is, the sum of the output elasticities is greater than one; only large values [$(1/\Sigma a_i) < 1$] would constitute evidence against the null hypothesis.

5. Omission of Relevant Explanatory Variables

Using the logarithmic version of the cost function (where the asterisk represents natural logarithms of the variables), we have

$$(1) \quad C^* = X^*\beta + Z^*\gamma + \varepsilon^* \ .$$

Here C^* is the (n x 1) vector of observations on total cost, X^* is the (N x K) matrix of observations of included independent variables, β is the (K x 1) vector of coefficients of included variables, Z^* is the (N x L) matrix of observations on excluded variables, γ is the (L x 1) vector of coefficients of the excluded independent variables, and ε^* is the stochastic disturbance discussed in the previous chapter. The cost model would be the classical normal linear multiple regression model if the Z variables were included. The estimators of β would be unbiased but, because the Z^* variables are omitted, specification error is present in the cost equation resulting from the omission of relevant explanatory variables. The error is due to the fact that no observations are available on the omitted variables and because of specification error the possibility of bias must be considered. When

$$(2) \quad C^* = X^*\beta + \varepsilon^{**}$$

is estimated but the true model is equation (1), we have

$$(3) \quad E(\hat{\beta}) = (X^{*\prime}X^*)^{-1} X^{*\prime}(X^*\beta + Z^*\gamma)$$

$$= \beta + (X^{*\prime}X^*)^{-1} X^{*\prime}(Z^*\gamma) \; .$$

The Z matrix can be partitioned as $[Z^*, Z_2^* \ldots Z_i^*]$ for the L omitted variables, which results in

$$(4) \quad E(\hat{\beta}) = \beta + \sum_{j=1}^{L} (X^{*\prime}X^*)^{-1} X^{*\prime}Z_j^*\gamma_j$$

and the coefficient of the $i^{\underline{th}}$ included variable is

$$(5) \quad E(\hat{\beta}_i) = \beta_i + \sum_{j=1}^{L} d_{ji}\gamma_j$$

where

$$(6) \quad Z_j = \sum_{i=1}^{K} d_{ji}X_i^* + R_j \qquad j = 1, \ldots L$$

and R_j is the residual in the least squares regression (6) and since the X's are non-stochastic the equation is purely descriptive. The bias of $\hat{\beta}_i$ is equal to $\sum_{j=1}^{L} d_{ji}\gamma_j$ and its sign and magnitude depend on d_{ji} [the empirical relation between the included variable X_1 and the excluded variables Z_j in equations (6)] and γ_j [the relation between the excluded variable and the dependent cost variable in equation (1)]. A sufficient condition to establish the direction of bias would require all the signs of the (d_{ji}, γ_j) terms to be the same for positive and opposite for negative bias.

Consider the cost model developed in the previous chapter which can be written

$$(7)\quad C_t^* = b_0^* + b_1Q_t^* + b_2P_{1,t}^* + b_3P_{2,t}^* + b_4P_{3,t}^* + b_5P_{4,t}^*$$

$$+ \gamma_1Q_{t-1}^* + \gamma_2X_{1,t-1}^* + \gamma_3X_{2,t-1}^* + \gamma_4X_{3,t-1}^*$$

$$+ \gamma_5X_{4,t-1}^* + \varepsilon_t^* .$$

Since P_4 is assumed to be the same for all computer centers, we have $b_0^{*'} = \{b_0 + b_5\overline{P}_4\}$ and equation (7) can be rewritten as

$$(8)\quad C_t^* = b_0^{*'} + b_1Q_t^* + b_2P_{1,t}^* + b_3P_{2,t}^* + b_4P_{3,t}^* + \gamma_1Q^*_{t-1}$$

$$+ \gamma_2X_{1,t-1}^* + \gamma_3X_{2,t-1}^* + \gamma_4X_{3,t-1}^* + \gamma_5X_{4,t-1}^* + \varepsilon_t^* ,$$

we estimate

$$(9) \quad C_t^* = b_0^{*\prime} + b_1 Q_t^* + b_2 P_{1,t}^* + b_3 P_{2,t}^* + b_4 P_{3,t}^* + \varepsilon_t^* .$$

Since we are concerned with economies of scale and $b_1 = (\frac{1}{r})$ where r is the sum of the output elasticities, the b_1 term is examined first.

$$(10) \quad E(\hat{b}_1) = b_1 + d_{11}\gamma_1 + d_{21}\gamma_2 + d_{31}\gamma_3 + d_{41}\gamma_4 + d_{51}\gamma_5$$

where

$$(11) \quad Q_{t-1}^* = d_{11}Q_t^* + d_{12}P_{1,t} + d_{13}P_{2,t} + d_{14}P_{3,t} + R_{1,t}$$

$$(12) \quad X_{\ell,t-1}^* = d_{\ell+1,1}Q_t^* + d_{\ell+1,2}P_1 + d_{\ell+1,3}P_{2,t} + d_{\ell+1,4}P_{3,t} + R_{\ell+1,t}\,, \qquad \ell = 1, \ldots 4.$$

The signs of the γ coefficients are presumed to be positive. The centers with large past period output and inputs have large current costs. Centers which are large are consistently large, and those that are relatively small in the current period were small in the previous period. The signs of the $d_{\ell+1,1}$ terms are assumed to be positive. This conjecture is reasonable since there is evidence that larger computer centers pay higher wages than the smaller ones.[10]

A clarifying digression may be important before the empirical results are presented. The assumption of a

---

[10]"Annual Salary Survey," Computers and Automation (January, 1972), p. 43.

positive relation between current prices and both lagged output and inputs need not be a functional relation. Equations (5) are descriptive; all the variables in those equations are non-stochastic. Also, it should be remembered that each facility has its own geographically insulated labor market. Since it has been postulated that centers are relatively consistent in their output, it is possible to have a direct relation between observed current input prices and lagged input quantities and still retail downward sloping factor demand schedules. If we consider the traditional factor demand schedule in the factor price and quantity space, we find the schedule shifts to the right for centers with larger expected output. With a positive sloped supply schedule, we trace a positive sloped locus in factor price and quantity space. This is what one concludes from the model and the data.

We see that $E(\hat{b}_1) > b$, when b_1 is estimated from (9) and the expectation of the estimator of the sum of the output elasticities may have a negative bias because $\hat{b}_1 = (\frac{\hat{1}}{r})$.[11] The expectation of the estimator of the returns to scale parameter is possibly less than would be the result if the model were correctly specified. It should be noted that the estimators of the price variables

---

[11] It should be noted that the small sample properties do not "carry over" so that $1/E(\hat{b}_i) \neq E(1/\hat{B}_i)$.

have a positive bias. Following the same method as used in current output, we have

$$(13) \quad E(\hat{b}_k) = b_k + \{\sum_{j=1}^{5} d_{j,k}\gamma_j\} \qquad k = 2, \ldots 4$$

where the d's were defined in equations (11) and (12). The bias is positive since both the d_{ji} and γ_i terms are to remain positive as developed in the output case.[12]

6. Empirical Results

The form of the regression results are presented as follows:

$$\hat{C}^*_j = \hat{A}^* + \underset{(S_{\hat{Q}})}{\left(\widehat{\frac{1}{r}}\right)} Q^*_j + \underset{(S_{\hat{P}_1})}{\left(\frac{\hat{\alpha}_1}{r}\right)} P^*_1 + \underset{(S_{\hat{P}_2})}{\left(\frac{\hat{\alpha}_2}{r}\right)} P^*_2 + \underset{(S_{\hat{P}_3})}{\left(\frac{\hat{\alpha}_3}{r}\right)} P^*_3 , \qquad R^2 = \ldots$$

where the asterisk represents the natural logarithms of the variable, n is the number of observations, the "hat" (^) indicates parameter estimates and on C it is a reminder that equation holds for the fitted values of the variables. R^2 is the coefficient of determination, $S_{(\hat{Q})}$ is the estimated standard error of the coefficient $(\widehat{1/r})$ and the asterisk indicates natural logarithm of the variables.

---

[12] It is also known that the estimated standard error of the estimator contains an upward bias. The customary test of significance will tend to reject the null hypothesis less frequently than is appropriate at the given level of significance. See Kmenta, Econometrics, op. cit., p. 314.

It is clear that the coefficients of the production function-structural equation can be determined from the estimates in the reduced form cost function. We have exact identification of all the structural equations. Thus the estimators of the production function parameters obtained via indirect least squares on the cost function are maximum likelihood estimators.

The regression used output as defined in Chapter III with the weight for add-time and cycle-time being equal ($w_1 = w_2 = .5$). This resulted in the estimated relation for the cost function

$$\hat{C}^*_j = \underset{}{2.174} + \underset{(0.044)}{0.428Q^*_j} + \underset{(0.178)}{0.676P^*_{1j}} + \underset{(0.161)}{0.308\ell nP^*_{2j}} + \underset{(0.143)}{0.289P^*_{3j}}$$

$$R^2 = .608$$

$$n = 115$$

where input 1 is systems programmers, input 2 is administration, and input 3 is operational personnel, as previously defined.

TABLE I.--Test Statistic with $W_A = W_C = .5$

	Value	t-ratios	Significance Level
Output	Q	9.637	1%
Systems Programmers	P_1	3.807	2%
Administration	P_2	1.911	10%
Operations	P_3	2.014	5%

It is clear from the computed t-ratios that we can reject the simple null hypothesis at the 10% significance level, that each coefficient is equal to zero against the alternative hypothesis that the coefficients are not equal to zero. The size of the coefficient of determination and the F-statistic equaling 43.672 reinforces the maintained hypothesis that the functional form of the production function is satisfactory, but of course is not a test of specification error.

It is possible to test the hypothesis of linearity of the logarithmic cost equation with either the Durbin-Watson test[13] or the Yule-Kendal[14] normality test. The Durbin-Watson test statistic is

$$d = \frac{\sum_{j=2}^{115} (e_j - e_{j-1})^2}{\sum_{j=1}^{115} e_j^2}$$

which was originally designed to test autoregression; i.e., disturbances uncorrelated over time against the alternative hypothesis of one-period autoregression. If we order the residuals (e_j) with respect to increasing values of the dependent variable, we can test whether the residuals of

[13] J. Durbin and G. S. Watson, "Testing for Serial Correlation in Least Squares Regression I," Biometrica (June, 1950), pp. 409-28.

[14] U. Yule and M. Kendall, An Introduction to the Theory of Statistics (London: Charles Griffin, 1950).

the regression are random as would be true for the disturbances if the population regression equation is linear. The calculated value for d is 1.69 and the critical region upper bound at the 1% level is 1.63. We cannot reject the null hypothesis of linearity.

A less restrictive test is the normality test. This "turns" test involves counting the number of terms (P) defined as when $e_{j-1} < e_j > e_{j+1}$ or $e_{t-1} > e_t < e_{t+1}$, where the residuals are again in order of increasing output levels. Under the null hypothesis of randomness the mean of (P) would be

$$H_0 : E_{(P)} = \frac{2(n-2)}{3} \quad \text{and Var (P)} = \frac{16n - 29}{90} .$$

The alternative hypothesis of non-linearity is

$$H_A : E_{(P)} \neq \frac{2(n-2)}{3}$$

and the test statistic is $\dfrac{P-2(n-2)/3}{\sqrt{(16n-29)/90}} \sim N(0,1)$

P = 70 and the test statistic equals 2.20. Using a two-tailed test, at the 1% level of significance, we cannot reject the null hypothesis of randomness and thus linearity. These tests are not specification error tests of the function form of the production function but they tend to justify the maintained hypothesis.

The parameter estimates are used to investigate the economies of scale question. The null hypothesis is that the sum of the output elasticities is equal to one or $(1/\Sigma a_i)$ equals one. We will only accept values of $\Sigma a_i > 1$ as evidence rejecting the null hypothesis and thus the alternative hypothesis is that $\Sigma a_i > 1$ or $(1/\Sigma a_i) < 1$. Using a one-tailed t-test, we find

$$H_0 : (\frac{1}{r}) = 1$$

$$H_Z : (\frac{1}{r}) < 1$$

$$\frac{(\frac{1}{r}) - 1}{S_{(\hat{Q})}} = \frac{0.428 - 1.00}{.044} = -13.0$$

and the critical region begins at 2.326 at the 1% significance level. I have no idea about the relative cost of Type I and Type II error. The 1% level in this and the previous linearity test was adopted and can only be justified on "popularity" grounds. We reject the hypothesis of constant returns to scale. We estimated the sum of the output elasticities (sometimes called the production function coefficient)[15] to be 2.34 which indicates substantial economies of scale.

---

[15] C. E. Ferguson, The Neoclassical Theory of Production and Distribution (London: Cambridge University Press, 1969), pp. 158-63.

Since $(1/\Sigma a_i)$ was estimated rather than (Σa_i), the confidence interval for Σa_i must be approximated. An approximation developed by Klein (1953) was used to obtain the large sample variance of the sum of the output elasticities. The general form of the approximation of the variance of $\hat{\alpha} = f(\hat{\beta}_1, \hat{\beta}_2, \ldots \hat{\beta}_k)$

$$\text{var}\ (\hat{\alpha}) \approx \sum_{i=1}^{k} \left[\frac{\partial f}{\partial \hat{\beta}_i}\right]^2 \text{var}\ (\hat{\beta}_i) + 2 \sum_{j<i}^{k} \left[\left(\frac{\partial f}{\partial \hat{\beta}_j}\right)\left(\frac{\partial f}{\partial \hat{\beta}_i}\right)\right] \text{Cov}\ (\hat{\beta}_j, \hat{\beta}_i)$$

$$j,\ k = 1,\ \ldots\ k;\ j < k\ .$$

In our case,

$$\Sigma \hat{a}_i = f\left(\widehat{\frac{1}{\Sigma a_i}}\right)$$

$$\text{var}\ (\Sigma \hat{a}_i) = (\Sigma \hat{a}_i)^4\ \text{var}\ \left(\widehat{\frac{1}{\Sigma a_i}}\right)$$

$$\text{var}\ (\Sigma \hat{a}_i) = (2.34)^4\ (0.0019)$$

$$\text{var}\ (\Sigma \hat{a}_i) = 0.0570$$

$$S_{(\Sigma \hat{a}_i)} = 0.2387$$

and the confidence interval for $\Sigma \hat{a}_i$ at the 95% level is

$$\Sigma\hat{a}_i - t_{n-k,.025} \cdot s_{(\Sigma\hat{a}_i)} \leq \Sigma a_i \leq (\Sigma\hat{a}_i) + t_{n-k,.025} \cdot s_{(\Sigma\hat{a}_i)}$$

$$1.866 \leq a_i \leq 2.794 .$$

The output elasticities for the individual factors are given in Table II.

TABLE II.--Output Elasticities

Systems Programmer	$\frac{\partial LN(Q)}{\partial LN(X_1)}$	1.581
Administration	$\frac{\partial LN(Q)}{\partial LN(X_2)}$	0.720
Operations	$\frac{\partial LN(Q)}{\partial LN(X_3)}$	0.675

The possible bias introduced by the omission of relevant lagged explanatory variables is a possible explanation for the calculated sum of the individual output elasticities being larger than the inverse of the coefficient of the output variable. Appendix C contains a discussion of the influence of the relative weights of cycle time and add time. All of the estimated coefficients are positive as expected and the rather large output elasticity for systems programmers is consistent with economic theory when a factor such as systems programming is used in relatively small amounts and the production function exhibits economies of scale.[16]

7. Summary

In this chapter we outlined the method of data collection and indicated some of the procedures used in this process. The sources of empirical data were described. The cost function and its logarithmic transformation were discussed and some of the important consequences were noted. The empirical results indicate that these data reject the hypothesis that the production function for computer output exhibits constant returns to scale when confronted with the alternative of economies of scale. We have no evidence to support the hypothesis that the economies of scale coefficient is changed due to weights in the measure of computer output. The measure of output proposed in this work does not produce significantly different empirical results than more expensive procedures used to measure output of computers.

---

[16] Milton Friedman, Price Theory (Chicago: Adline Publishing Co., 1962).

CHAPTER V

CONCLUDING REMARKS

The purpose of the study was to examine the hypothesis of economies of scale in the production of computer output. The model which was developed using the Shephard-Uzawa duality principle contains the assumption that the objective of the computer center is to minimize the cost of producing an expected level of output subject to a Cobb-Douglas production function. Engineering studies, cited in the literature review, have attempted to discover a relation between cost and output. These studies did not systematically develop models or test hypotheses but were restricted to analytic curve fitting. In contrast to the engineering studies, the model developed in this study specifically accounts for the optimization behavior of the facility and the technical relation between inputs and output from the production process.

The model has five equations (three marginal productivity equations, the definition of cost, and the production function), and five unknowns (systems programming, administration, operations service, capital, and total cost). The parameters of these structural equations can be estimated by indirect least squares regression on the reduced form cost function. The reduced form cost function is linear in the logarithms of cost and the explanatory variables output and prices of the factors of production. In deriving the model, the stochastic nature of the production process was noted and the disturbance due to unpredictable machine malfunction and operator error were explicitly included. Also, the factor demand equations were assumed to be stochastic rather than deterministic since it seems reasonable to assume that the computer center management can make mistakes in factor employment. The final stochastic element of the model is the adjustment process of the desired versus actual level of factor usage. The optimizing conditions only determine the desired level of factor usage. Since the level of factor usage cannot be instantaneously altered, a stock adjustment model represents the relation between actual and desired levels of the inputs.

The introduction of the stock adjustment process creates a difficulty. If we had instantaneous adjustment, we would only need data on current prices, cost, and output. But with non-instantaneous adjustment we need information on lagged output and inputs. The cross section of data available from the National Science Foundation contains only 1965 data. Thus we have a specification error resulting from omission of relevant lagged explanatory variables. The possible bias which is introduced because of the specification was considered, and its influence on the parameter estimates was discussed. The most important finding was that the bias was positive with respect to the multiplicative inverse of the sum of the output elasticities. The estimates of the sum of the output elasticities has a negative bias.

Under the specification of the model, we have exact identification, and the indirect least squares estimates are equivalent to maximum likelihood estimates. Indirect least squares estimation was performed on the cost function and the results are discussed next.

The null hypothesis of constant returns to scale was rejected in favor of the alternative hypothesis of increasing returns to scale. The sum of the output elasticities was estimated to be 2.33. This seemingly high result was demonstrated to be consistent with economic theory in the cost minimization case. In addition, the output elasticities of systems programming is greater than one. It was noted that this is consistent with economic theory when increasing returns to scale exist and the factor is used in relatively small quantities. This is likely to be the case in our study. Because the elasticity was large and on general principles, we considered testing the hypothesis of log-linearity of the cost function. Using the Durbin-Watson and a generalized runs test, we were not able to reject the hypothesis of log-linearity of the reduced form cost function.

Comparing the results of this study to previous results, we find they are similar. The Knight and Solomon studies and now Musgrave find increasing returns to scale at the firm level. It should be remembered that we are concerned with the production of computer output rather than the computer manufacturing industry. Other studies on the firm level, which also use engineering measures of output, have found analogous results. Engineers use the "Six-Tenths Rule" in estimating cost as capacity output increases. Symbolically, two plants are related as

$$C_2 = C_1 \left[\frac{X_2}{X_1}\right]^{.6}$$

where C_i is cost and X_i is capacity output of plant i. In terms of neoclassical production theory, the exponent is the inverse of the sum of the output elasticities. Both Moore [1959][1] and Alpert [1959][2] find similar results for the mineral and chemical industries. The computer engineers would have the ".43 rule" as a result of this study. But care should be taken not to extend these results too far. The sample data do not provide any information about the

---

[1] F. Moore, "Economies of Scale: Some Statistical Evidence," Quarterly Journal of Economics, Vol. 70 (January, 1962), pp. 138-50.

[2] S. B. Alpert, "Economies of Scale in the Metal Removal Industry," Journal of Industrial Economics, Vol. 17 (November, 1959), pp. 175-81.

production function outside the interval covered by the observed variables. One should not expect average unit cost to approach zero if he built one large computer to do the world's computation.

Some implications from this study might be drawn. The first is that the method chosen to measure computer output seems satisfactory for our purpose. This implication is especially strong when the cost of the alternative formulations are considered. The maintained hypothesis of cost minimization is clearly acceptable for university computer centers. The assumption may be correct for the majority of computer facilities in operation today. Relatively few centers are service bureaus--selling their output. Most computer centers are cooperating factors in firms and government organizations; with decentralized management, the organization of the computer facility can be thought of as a cost center which is consistent with our model of computer centers. This fact reinforces the generality of the results since the maintained hypothesis is close to reality.

One would expect to find increasing pressures to centralize computational activity to a single computer center. I believe a study of such activity would confirm this hypothesis. Also, one would expect the increased output from larger computers to be in various forms. One form would simply be more pages of output or new applications where lower unit costs make some applications feasible where previously they were too costly. An interesting form would be the improvement in the quality of the output. One would expect the accuracy of computational algorithms to improve and the presentation of results to be closer to what the human wants rather than what the computer dictates.

One final caution is in order. This study should not be interpreted as a complete justification for centralization of computer activities. Much more analysis would be needed to obtain a first approximation to the answer of centralization versus decentralization. The computer itself is only one aspect of the total computer center. A whole constellation of management issues need to be analyzed prior to centralizing any computer activities. The definition of output in time-sharing systems and the influence of software availability are major issues that need to be included in any centralization versus decentralization discussion. The whole organizational structure of the center as a component of a total firm or university should be studied on a cost benefit or effectiveness basis. None of these issues detract from the findings of this study but mention of these issues may prevent the unwary from making unwarranted conclusions.

The hypothesis tested in this study, like all hypotheses, is subject to further test. New data on mini-computers or time-sharing computers (where the user employs a small terminal to interactively communicate with the machine) may alter the results. An alternative functional form of the production function could alter the findings. Until further study of these issues is made, the results stand.

BIBLIOGRAPHY

BIBLIOGRAPHY

Alpert, S. B. "Economies of Scale in the Metal Removal Industry." _Journal of Industrial Economics_, Vol. 17 (November, 1959), pp. 175-81.

Arbuckle, R. "Computer Analysis and Thruput Evaluation." _Computers and Automation_ (January, 1966), pp. 12-19.

Arrow, K.; Chenery, H.; Minhas, B.; and Solow, R. "Capital-Labor Substitution and Economic Efficiency." _Review of Economics and Statistics_ (August, 1961), pp. 225-50.

Auerbach Computer Reviews and Key Data. New York: Auerbach Information, Inc., January, 1968.

Auerbach Computer Technology Reports. New York: Auerbach Information, Inc., 1969.

Averch, H. and Johnson, L. "Behavior of the Firm under Regulatory Constraint." _American Economic Review_, Vol. 52 (December, 1962), pp. 1,053-69.

Billings, H., and Hogan, R. "A Study of the Computer Manufacturing Industry in the United States." Unpublished Master's thesis, U.S. Naval Postgraduate School, Monterey, California, 1970.

Bridgeman, P. W. _The Logic of Modern Physics_. New York: Macmillan, 1928.

_________. _The Nature of Physical Theory_. New Jersey: Princeton University Press, 1936.

Calingaert, P. "Systems Performance Evaluation: Survey and Appraisal." _Communications of the ACM_ (January, 1967), pp. 13-16.

Cobb, C. and Douglas, P. "A Theory of Production." _American Economic Review_, Vol. 18 (March, 1928), pp. 139-65.

Computer Characteristics Review. Watertown, Mass.: Key Data Corporation, 1969.

Computer Components Review. Norwood: Commander Publishing, 1969.

"Annual Salary Survey." Computers and Automation. January, 1972, p. 43.

Data Products News. New York: Data News, Inc., 1970.

Douglas, P. The Theory of Wages. New York: Macmillan, 1934.

________. "Are There Laws of Production?" Presidential Address, American Economic Review, Vol. 38 (March, 1948), pp. 1-41.

Durbin, J. and Watson, G. S. "Testing for Serial Correlation in Least Squares Regression I." Biometrica (June, 1950), pp. 409-28.

Electronics News. New York: Fairchild Publishing Co., 1972.

Ferguson, C. The Neoclassical Theory of Production and Distribution. London: Cambridge University, 1969.

Friedman, M. Price Theory. Chicago: Adline Publishing Co., 1962.

Goldberger, J. "Interpretation and Estimation of Cobb-Douglas Functions." Econometrica, Vol. 36 (July-October, 1968), pp. 464-72.

Griliches, Z. "Distributed Lags: A Survey." Econometrica, Vol. 35 (November-March, 1967), pp. 31-48.

Hamblin, I. Computers in Higher Education. Atlanta: Southern Regional Educational Board, 1967.

________. "Expenditures, Sources of Funds, and Utilization of Digital Computers for Research and Instruction in Higher Education, 1964-1965 with Projections for 1968-1969." Communications of the ACM, Vol. 7 (April, 1968), pp. 257-62.

Huesmann L. and Goldberg, R. "Evaluating Computer Systems through Simulation." Computer Journal (August, 1967), pp. 148-52.

Henderson, J. and Quandt, R. Microeconomic Theory. New York: McGraw-Hill, 1971.

IBM Technical Publications. White Plains: Technical Documentation Center, International Business Machines, 1970.

Ihrer, I. "Computer Performance Projected through Simulation." Computers and Automation. April, 1967, pp. 21-27.

Johnston, J. Statistical Cost Analysis. New York: McGraw-Hill, 1960.

Kmenta, J. "On Estimation of the C.E.S. Production Function." International Economic Review, Vol. 8 (June, 1967), pp. 180-89.

________. Elements of Econometrics. New York: Macmillan, 1971.

Knight, K. "A Study of Technological Innovation--The Evaluation of Digital Computers." Unpublished Doctoral dissertation, Carnegie Institute of Technology, 1963.

Malinvaud, E. Statistical Methods of Econometrics. Amsterdam: North Holland Publishing Co., 1968.

Marschak, J. and Andrews, W. "Random Simultaneous Equations and Theory of Production." Econometrica, Vol. 12 (July-October, 1944), pp. 143-205.

McFadden, D. An Econometric Approach to Production Theory. Forthcoming.

Merewitz, L. The Production Function in the Public Sector: Production of Postal Services in the U.S. Post Office. Berkeley: Center for Planning and Development, 1969.

Moore, F. "Economies of Scale: Some Statistical Evidence." Quarterly Journal of Economics, Vol. 70 (January, 1962), pp. 138-50.

Mundlak, Y. Review of Production Functions. Forthcoming.

Nerlove, M. "Returns to Scale in Electricity Supply." Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld. Edited by C. Christ. Stanford: Stanford University Press, 1963.

________. Estimation and Identification of Cobb-Douglas Production Functions. Chicago: Rand McNally, 1965.

Raisbuk, G. Information Theory. Cambridge: Massachusetts Institute of Technology Press, 1966.

Ramsey, J. and Zarembka, P. "Specification Error Tests and Alternative Functional Forms of the Aggregate Production Function." Journal of the American Statistical Association, Vol. 66 (September, 1971), pp. 471-77.

Schneidewind, N. "Analytic Model for the Design and Selection of Electronic Digital Computers." Unpublished Doctoral dissertation, University of Southern California, 1966.

Schwab, B. "Economic Evaluation and Selection of Electronic Data Processing Systems." Unpublished D.B.A. dissertation, University of California at Los Angeles, 1967.

Sewald, M.; Rauch, M.; Rodick, L.; and Wertz, L. "A Pragmatic Approach to Systems Measurement." Computer Decisions (July, 1971), pp. 38-40.

Sharpe, W. The Economics of Computers. New York: Columbia University Press, 1969.

Shephard, R. Cost and Production Functions. Princeton: Princeton University Press, 1953.

Solomon, M. "Economies of Scale and the I.B.M. System/360." Communications of the ACM, Vol. 5 (June, 1966), pp. 435-40.

Standard EDP Reports. New York: Auerbach Information, Inc., 1970.

Uzawa, H. "Duality Principles in the Theory of Cost and Production." International Economic Review, Vol. 5 (May, 1964), pp. 216-20.

Walters, A. "Production and Cost Functions: An Econometric Survey." Econometrica, Vol. 31 (January-April, 1963), pp. 1-66.

Yule, U., and Kendall M. An Introduction to the Theory of Statistics. London: Charles Griffin, 1950.

Zellner, A.; Kmenta, J.; and Dréze, J. "Specification and Estimation of Cobb-Douglas Production Function Models." Econometrica, Vol. 34 (July-October, 1966), pp. 784-95.

APPENDICES

APPENDIX A

1. Generalized File Processing Problem A

The essence of most business data processing applications is the updating of files to reflect the effects of various types of transactions. This benchmark problem is a file processing run in which transaction data in a detail file is used to update a master file, and a record of each transaction is written in a report file or journal (Figure A-1). This type of run forms the bulk of the work load for many computer systems, in diverse applications such as billing, payroll, and inventory control.

The listed "activity" factors of 0.0, 0.1, and 1.0 refer to cases in which an average of 0, 0.1, and 1.0 transaction record, respectively, must be processed for each record in the master file. Low activities are characteristic of applications such as inventory control, whereas a payroll run might well have an activity factor of 1.0. All calculated processing times are reported in terms of the number of minutes required to process 10,000 master-file records.

Figure A-2 is a general flow chart that summarizes the computational process. Both the master file and detail file are sequentially arranged, and conventional batch

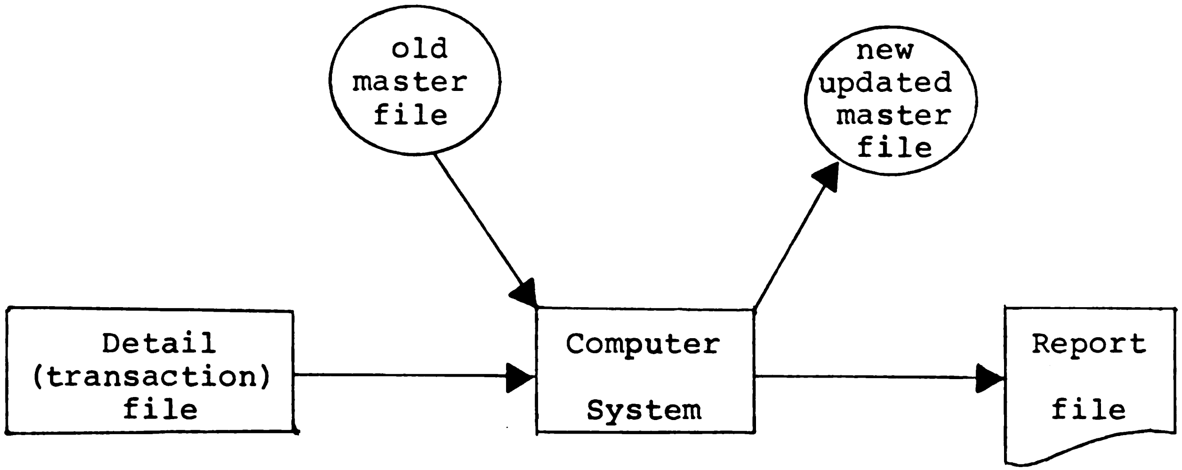

FIGURE A-1.--Run Diagram for Generalized File Processing

Problem A

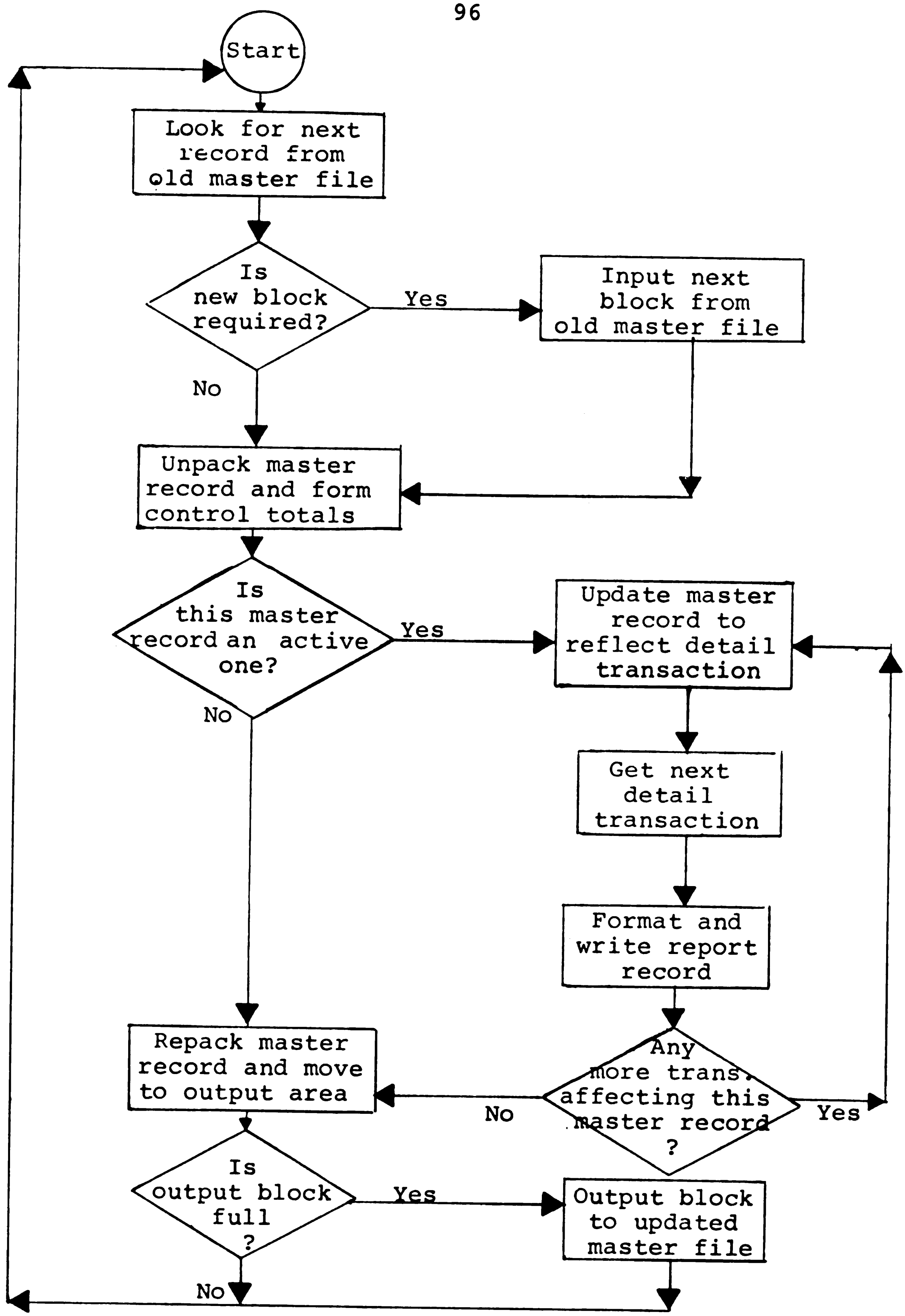

FIGURE A-2.--General Flowchart for Generalized File Processing Problem A processing techniques are employed. Record lengths are 108 characters for the master file, 80 characters (1 card) for the detail file, and 120 characters (1 line) for the report file. Record layouts are fixed for the detail and report files, but are left flexible for the master file in order to take advantage of the specific capabilities of each computer system.

Card reading and printing are performed on-line in all standard configurations except paired configurations VIIB and VIIIB, in which card-to-tape and tape-to-printer transcriptions are performed off-line, usually by a separate small-scale computer. The master file is on magnetic tape in all standard configurations except Configuration I, where it is on punched cards.

2. Random Access File Processing Problem

This benchmark problem represents a wide range of real-time computer applications in which an on-line master file is accessed to answer inquiries and/or updated to reflect various types of transactions. Figure A-3 shows the basic run diagram. Examples of this type of processing include real-time inventory control, credit checking, airline and hotel reservations, and on-line savings systems.

In contrast to Generalized File Processing Problem A, described in Item (1), this problem uses random access storage to hold the entire master file on-line, and processes all transactions as they occur, without prior sorting. All

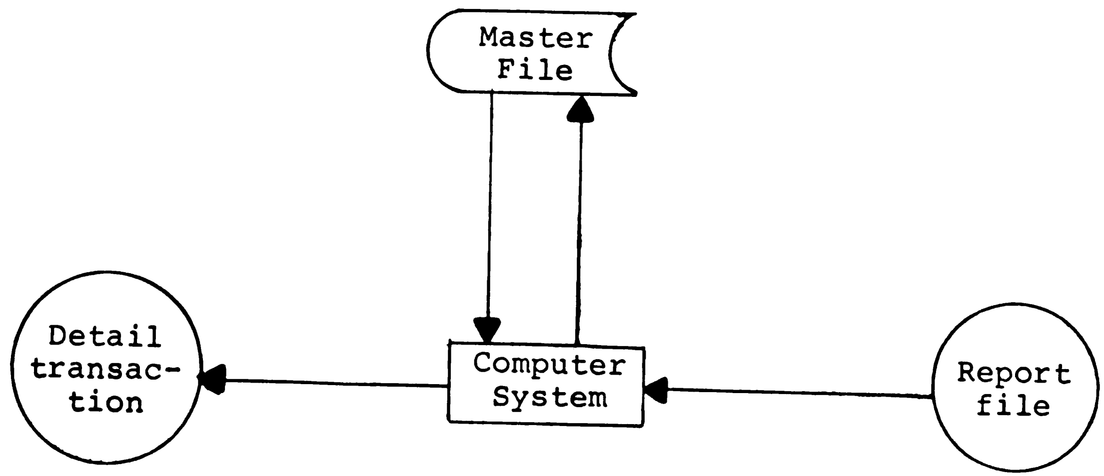

FIGURE A-3. Run Diagram for Random Access File Processing Problem calculated times are reported in terms of the time in milliseconds required to process each transaction and the total time in minutes required to process 10,000 transactions.

This problem is evaluated for one or more of the three random access standard configurations (IIIR, IVR, and VIIIR). Where there are two or more random access devices that could satisfy the specified capacity requirements, our choice is based upon considerations of economy, system throughput, software support, and reliability. Therefore, disc files will normally be chosen in preference to drums (which are relatively expensive) or magnetic strip devices (which tend to be relatively slow and less reliable).

Figure A-4 is a general flow chart that summarizes the computational process. The master file is sequentially arranged in random access storage, and a two-stage indexing procedure is used to determine the location of each master-file record that needs to be accessed. Record lengths are 108 characters for the master file, 80 characters (1 card) for the detail transactions, and 120 characters (1 line) for the report file. Record layouts are fixed for the detail and report files, but are left flexible for the master file so that the specific features of each computer system can be advantageously utilized.

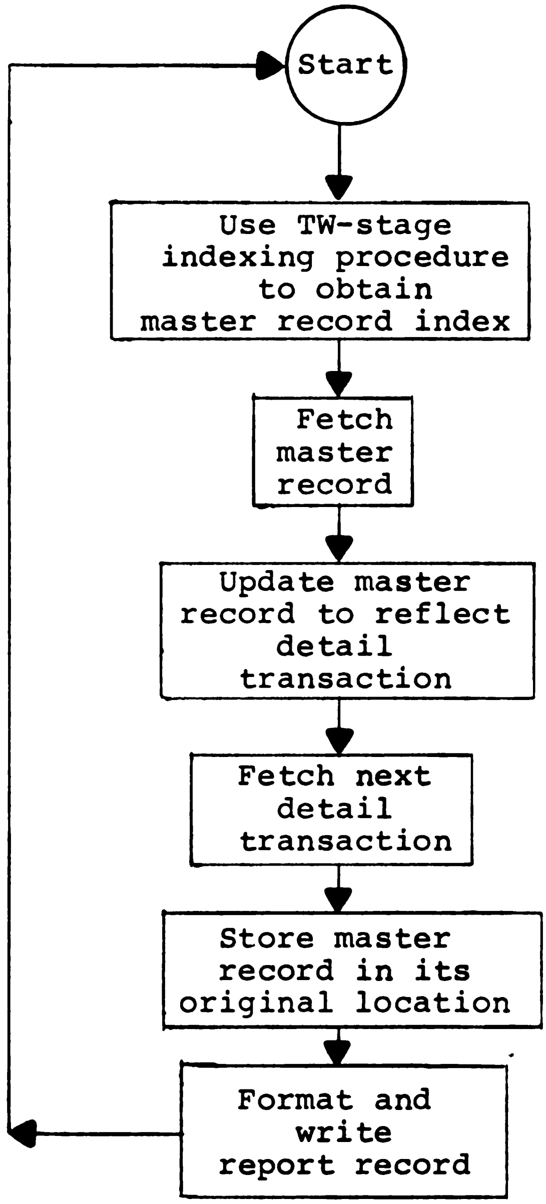

FIGURE A-4.--General Flow Chart for Random Access File Processing Problem

The detail transactions (e.g., inquiries, orders, or deposits) are assumed to be arriving in a random sequence and at a continuous rate that is high enough to ensure that one or more transactions are always waiting to be processed. Therefore, it makes no difference whether the transactions enter the system via an on-line reader, a simple remote inquire terminal, or a multiterminal data communications network. This assumption means that the Random Access File Processing Problem does not attempt the highly complex and variable task of measuring the efficiency of real-time data communications networks; it simply measures the central computer system's ability to locate and update randomly addressed master-file records.

The report file is written on either magnetic tape or a random access device, presumably for printing at some later time. Each report record is also made available for optional transmission back to the remote terminal that initiated the transaction (though the processor time required to effect this transaction is not included in the published timing figures).

3. Sorting

Because conventional data processing techniques usually require all records to be arranged in a particular sequence, sorting operations are an important and time-consuming part of the work load in most business computer installations. This benchmark problem requires that a file consisting of 10,000 records, each 80 characters in length, be arranged sequentially according to an 8-digit key, such as an account number.

The "Standard Estimate" column lists the estimated sorting times calculated by our analysts for sorting operations that use straightforward magnetic tape merging techniques. Two-way tape merging is used in the four-tape Standard Configuration II and three-way merging in all of the larger systems.

Whenever timing data is available for a standard, manufacturer-supplied sort routine, the time required to perform the same 10,000-record sort is listed in the "Available Routines" column. Because most manufacturer-supplied sort routines now use internal sorting and merging techniques which are more sophisticated than those used to prepare our estimates, the "Available Routines" sort time will often be substantially less than the "Standard Estimate" time for a given configuration. Nevertheless, the Standard Estimates provide useful, directly comparable indications of each computer system's basic capabilities to perform magnetic tape input-output operations.

4. Matrix Inversion

In many scientific and operations research applications, such as multiple regression, linear programming, and the solution of simultaneous equations, the bulk of the central processor's time is spent in inverting large matrices. This benchmark problem involves the inversion of 10-by-10 and 40-by-40 matrices. It measures the speed of the central processor on floating-point calculations; no input or output operations are involved. All matrix elements are held within the system's main storage unit in floating-point form with a precision equivalent to at least eight decimal digits.

The "Standard Estimate" columns list the matrix inversion times calculated by our analysis through a simple estimating procedure that uses the system's floating-point arithmetic speeds. Whenever timing data is available for a standard, manufacturer-supplied matrix inversion routine, it is reported in the "Available Routines" columns.

5. Generalized Mathematical Problem A

Another frequently encountered scientific problem involves the evaluation of polynomial equations of the type $Y = A + Bx + Cx^2 + Cx^3 + Ex^4 + Fx^5$. This benchmark problem includes the following basic steps:

1. Read in input record consisting of 10 eight-digit numbers,
2. Perform a floating-point calculation that consists of evaluating 5^{th} order polynomials, executing five division operations, and evaluating one square root.
3. For every 10 input records, form and print one output record consisting of 10 eight-digit numbers.

The "Computation Factors" of 1, 10, and 100 mean that the standard calculation described above is performed 1, 10, or 100 times, respectively, for each input record to show the effects of varying ratios of computation to input/output volume. Processing times are listed in terms of milliseconds per input record.

These examples are drawn from Auerbach Information, Inc.

TABLE A-I.--The Data

Identification	C	P_1	P_2	P_3	Q
11237501112341	180.00	10.00	6.00	3.00	2.36
11237501112341	5.1930	2.3026	1.7918	1.0986	0.8573
11240501271322	31.00	6.00	11.00	2.33	0.01
11240501271322	3.4340	1.7918	2.3979	0.8473	-4.6805
11261501112341	62.00	4.00	14.00	4.67	3.58
11261501112341	4.1271	1.3863	2.6391	1.5404	1.2749
12235501112341	74.00	8.00	12.00	5.00	1.08
12235501112341	4.3041	2.0794	2.4849	1.6094	0.0759
13237501112341	66.00	1.67	13.00	1.33	3.14
13237501112341	4.1897	0.5108	2.5649	0.2877	1.1453
13241501112341	175.00	10.00	11.67	6.67	18.80
13241501112341	5.1648	2.3026	2.4567	1.8971	2.9339
14253501112341	100.00	5.00	10.00	2.50	0.39
14253501112341	4.6052	1.6094	2.3026	0.9163	-0.9400
15249501521142	801.00	7.31	23.00	3.63	1.82
15249501521142	6.6859	1.9896	3.1355	1.2879	0.6016
15252501512331	28.00	3.00	5.00	5.00	0.65
15252501512331	3.3322	1.0986	1.6094	1.6094	-0.4309
15268501795342	47.00	15.00	7.00	5.00	0.76
15268501795342	3.8501	2.7080	1.9459	1.6094	-0.2808
15372501121342	1897.00	8.97	9.07	5.57	57.67
15372501121342	7.5480	2.1937	2.2049	1.7174	4.0548
15376001112341	1812.00	7.73	9.00	5.05	39.07
15376001112341	7.5022	2.0448	2.1972	1.6195	3.6653
15376501112341	183.00	5.67	6.00	5.60	0.52
15376501112341	5.2095	1.7346	1.7918	1.7228	-0.6523
15377001112341	183.00	7.88	10.00	4.62	6.83
15377001112341	5.2095	2.0637	2.3026	1.5302	1.9219
15377501112341	1003.00	7.63	9.00	4.58	31.80
15377501112341	6.9108	2.0314	2.1972	1.5220	3.4593
15378001112341	57.00	6.00	11.00	5.50	0.96
15378001112341	4.0431	1.7918	2.3979	1.7047	-0.0384

153785011112341	129.00	9.33	12.50	2.33	1.29
153785011112341	4.8598	2.2336	2.5257	0.8473	0.2521
153788011112341	34.00	3.00	7.00	3.00	0.67
153788011112341	3.5264	1.0986	1.9459	1.0986	-0.4021
153835011121342	96.00	8.20	10.50	3.71	2.90
153835011121342	4.5643	2.1041	2.3514	1.3122	1.0654
153837011121342	50.00	3.50	2.50	2.00	1.57
153837011121342	3.9120	1.2528	0.9163	0.6931	0.4485
162495012123?1	57.00	8.00	12.00	3.33	0.65
162495012123?1	4.0431	2.0794	2.4849	1.2040	-0.4349
162535011171342	49.00	5.00	10.00	5.00	0.51
162535011171342	3.8918	1.6094	2.3026	1.6094	-0.6734
172615011112341	136.00	7.00	16.00	4.25	1.20
172615011112341	4.9127	1.9459	2.7726	1.4469	0.1837
172625012211142	41.00	8.00	19.00	3.00	0.67
172625012211142	3.7136	2.0794	2.9444	1.0986	-0.3957
172645011121142	460.00	12.00	8.00	5.00	3.59
172645011121142	6.1312	2.4849	2.0794	1.6094	1.2787
192425011130342	102.00	7.50	10.00	7.50	1.29
192425011130342	4.6250	2.0149	2.3026	2.0149	0.2582
202465011112341	270.00	15.00	7.83	1.83	1.93
202465011112341	5.5984	2.7080	2.0584	0.6061	0.6595
212495011171342	177.00	10.67	15.33	5.00	4.73
212495011171342	5.1761	2.3671	2.7300	1.6094	1.5532
212525015112341	451.00	10.00	10.57	5.00	3.66
212525015112341	6.1115	2.3026	2.3582	1.6094	1.2964
212555012112341	76.00	7.00	6.00	1.00	1.88
212555012112341	4.3307	1.9459	1.7918	0.0	0.6306
212795011112341	82.00	5.67	13.75	2.50	13.21
212795011112341	4.4067	1.7346	2.6210	0.9163	2.5811

222385011112341	137.00	7.67	9.50	7.20	5.83
222385011112341	4.9200	2.0369	2.2513	1.9741	1.7636
232415011112341	90.00	7.50	9.00	3.67	0.75
232415011112341	4.4998	2.0149	2.1972	1.2993	-0.2923
242465011121332	81.00	8.00	20.00	1.14	2.16
242465011121332	4.3944	2.0794	2.9957	0.1335	0.7691
242665011130332	22.00	3.00	5.00	1.00	0.63
242665011130332	3.0910	1.0986	1.6094	0.0	-0.4564
242805015521342	273.00	9.00	10.75	4.17	0.78
242805015521342	5.6095	2.1972	2.3749	1.4271	-0.2469
243135011121342	286.00	8.67	8.25	4.67	3.40
243135011121342	5.6560	2.1595	2.1102	1.5404	1.2247
243315011112341	121.00	3.00	8.50	2.00	1.51
243315011112341	4.7958	1.0986	2.1401	0.6931	0.4130
252535014112341	40.00	4.00	13.00	4.00	1.29
252535014112341	3.6889	1.3863	2.5649	1.3863	0.2582
252555011112341	189.00	7.11	13.00	4.43	4.12
252555011112341	5.2417	1.9617	2.5649	1.4881	1.4161
252625011112341	940.00	7.92	15.00	4.00	10.53
252625011112341	6.8459	2.0690	2.7080	1.3863	2.3544
252715015521322	10.00	4.00	3.00	1.00	1.08
252715015521322	2.3026	1.3863	1.0986	0.0	0.0759
252725011130142	171.00	11.50	20.00	5.00	2.08
252725011130142	5.1417	2.4423	2.9957	1.6094	0.7340
272595011112341	149.00	9.00	8.00	2.83	14.30
272595011112341	5.0039	2.1972	2.0794	1.0415	2.6601
272775011112341	234.00	8.33	12.50	3.33	3.78
272775011112341	5.4553	5.1203	2.5257	1.2040	1.3291
282695011112341	348.00	9.13	12.40	4.33	8.29
282695011112341	5.8522	2.2110	2.5177	1.4663	2.1153
282705011113341	34.00	7.00	5.00	3.00	1.08
282705011113341	3.5264	1.9459	1.6094	1.0986	0.0759
292435011112341	234.00	10.00	15.00	6.00	4.53
292435011112341	5.4553	2.3026	2.7080	1.7918	1.5102
292555011121342	346.00	7.25	3.47	3.29	13.10
292555011121342	5.8464	1.9810	1.2432	1.1910	2.5727

31236501813311	30.00	6.00	10.00	4.00	0.32
31236501813311	3.4012	1.7918	2.3026	1.3863	-1.1281
31248501121342	148.00	3.33	7.00	4.00	2.49
31248501121342	4.9972	1.2040	1.9459	1.3863	0.9116
31266501112341	697.00	9.10	11.60	5.42	22.21
31266501112341	6.5468	2.2083	2.4510	1.6895	3.1007
32247501121342	114.00	5.00	12.67	3.60	0.54
32247501121342	4.7362	1.6094	2.5390	1.2809	-0.6172
32271501121142	777.00	8.93	11.33	4.43	10.82
32271501121142	6.6554	2.1893	2.4277	1.4881	2.3811
32322501521142	35.00	5.00	11.00	2.00	1.51
32322501521142	3.5553	1.6094	2.3979	0.6931	0.4124
33243501212331	74.00	10.00	8.00	3.50	0.43
33243501212331	4.3041	2.3026	2.0794	1.2528	-0.8404
33275501112341	363.00	7.67	14.00	6.27	7.17
33275501112341	5.8944	2.0369	2.6391	1.8352	1.9702
33291501112341	524.00	5.13	6.70	2.86	27.01
33291501112341	6.2615	1.6352	1.9021	1.0521	3.2963
34274501112341	347.00	5.30	8.70	3.19	9.57
34274501112341	5.8493	1.6677	2.1633	1.1598	2.2590
35259501112341	133.00	8.00	9.00	1.50	3.88
35259501112341	4.8903	2.0794	2.1972	0.4055	1.3568
37242501112341	15.00	2.00	5.00	3.00	0.43
37242501112341	2.7080	0.6931	1.6094	1.0986	-0.8404
38256501112341	211.00	6.25	15.00	3.00	13.19
38256501112341	5.3519	1.8326	2.7080	1.0986	2.5791
41259501121142	550.00	6.75	7.00	4.00	15.11
41259501121142	6.3099	1.9095	1.9459	1.3863	2.7152
41261501112341	239.00	7.67	10.00	4.60	8.87
41261501112341	5.4765	2.0369	2.3026	1.5261	2.1828
41264501130132	78.00	4.00	7.67	5.00	0.43
41264501130132	4.3567	1.3863	2.0369	1.6094	-0.8404
42237501412331	68.00	6.00	8.00	3.00	0.73
42237501412331	4.2195	1.7918	2.0794	1.0986	-0.3097
35259501112341	133.00	8.00	9.00	1.50	3.88
35259501112341	4.8903	2.0794	2.1972	0.4055	1.3568

362866011112341	194.00	6.75	9.75	3.29	4.26
36286601112341	5.2679	1.9095	2.2773	1.1896	1.4491
42243501112341	46.00	10.00	12.00	4.00	3.82
42243501112341	3.8286	2.3026	2.4849	1.3863	1.3399
43247501213331	101.00	9.50	7.50	3.00	1.34
43247501213331	4.6151	2.2513	2.0149	1.0986	0.2910
43249501813311	37.00	5.00	10.00	4.00	0.76
43249501813311	3.6109	1.6094	2.3026	1.3863	-0.2808
43255501113331	68.00	5.00	10.00	3.00	0.42
43255501113331	4.2195	1.6094	2.3026	1.0986	-0.8755
43256501521142	51.00	5.00	12.50	4.50	2.33
43256501521142	3.9318	1.6094	2.5257	1.5041	0.8460
43262501121342	820.00	8.00	12.50	0.85	23.21
43262501121342	6.7093	2.0794	2.5257	-0.1625	3.1445
43265501121342	202.00	7.90	12.75	5.22	6.78
43265501121342	5.3083	2.0669	2.5455	1.6529	1.9142
43278501130342	74.00	5.67	9.00	3.50	0.86
43278501130342	4.3041	1.7346	2.1972	1.2528	-0.1472
43380501512341	181.00	8.50	10.00	4.67	2.62
43380501512341	5.1985	2.1401	2.3026	1.5404	0.9613
43383501712341	18.00	8.00	6.00	4.00	1.29
43383501712341	2.8904	2.0794	1.7918	1.3863	0.2582
43397501112341	337.00	10.00	10.00	11.50	6.02
43397501112341	5.8201	2.3026	2.3026	2.4423	1.7958
43398501121342	338.00	7.50	11.33	4.00	7.58
43398501121342	5.8230	2.0149	2.4277	1.3863	2.0250
44137601121342	453.00	9.60	8.43	3.38	10.55
44137601121342	6.1159	2.2618	2.1316	1.2182	2.3559
44249501121342	154.00	4.33	9.00	2.83	11.68
44249501121342	5.0370	1.4663	2.1972	1.0415	2.4582
44287501112341	259.00	7.00	15.50	8.13	19.02
44287501112341	5.5568	1.9459	2.7408	2.0949	2.9455
44291501412331	55.00	5.00	5.50	4.00	1.04
44291501412331	4.0073	1.6094	1.7047	1.3863	0.0351
46263501112341	64.00	6.00	15.00	4.00	1.51
46263501112341	4.1589	1.7918	2.7080	1.3863	0.4124

462775011112341	811.00	6.63	9.83	4.17	1.66
462775011112341	6.6983	1.8909	2.2858	1.4271	0.5046
462895011113341	68.00	5.00	15.00	5.00	1.08
462895011113341	4.2195	1.6094	2.7080	1.6094	0.0759
462906011113341	160.00	6.17	10.50	3.40	2.01
462906011113341	5.0752	1.8192	2.3514	1.2238	0.6989
462925011113341	14.00	3.00	6.00	2.50	1.25
462925011113341	2.6391	1.0986	1.7918	0.9163	0.2209
462965011121342	192.00	6.30	14.00	4.27	1.10
462965011121342	5.2575	1.8405	2.6391	1.4523	0.0979
463005012301132	27.00	2.00	4.00	2.00	0.78
463005012301132	3.2958	0.6931	1.3863	0.6931	-0.2526
472585011112341	126.00	5.00	9.50	2.60	7.55
472585011112341	4.8363	1.6094	2.2513	0.9555	2.0213
492705017211342	56.00	7.50	18.00	3.33	0.17
492705017211342	4.0254	2.0149	2.8904	1.2040	-1.7567
492845012211142	39.00	7.50	12.00	2.67	0.85
492845012211142	3.6636	2.0149	2.4849	0.9808	-0.1642
493065011112341	322.00	6.33	9.00	3.50	52.09
493065011112341	5.7746	1.8458	2.1972	1.2528	3.9530
493405011121342	371.00	4.87	7.50	3.13	7.92
493405011121342	5.9162	1.5830	2.0149	1.1412	2.0700
493445011121342	410.00	8.33	12.00	6.29	1.56
493445011121342	6.0162	2.1203	2.4849	1.8383	0.4463
493457011121342	193.00	8.40	3.67	6.00	10.74
493457011121342	5.2627	2.1282	1.2993	1.7918	2.3744
493505011130332	40.00	6.00	4.00	1.00	1.08
493505011130332	3.6889	1.7918	1.3863	0.0	0.0759
502395012301142	226.00	8.00	17.00	2.37	2.06
502395012301142	5.4205	2.0794	2.8332	0.8638	0.7206
532785011112341	144.00	2.40	4.00	3.50	7.20
532785011112341	4.9698	0.8755	1.3863	1.2528	1.9742
532795011121342	111.00	7.50	10.00	2.75	2.50
532795011121342	4.7095	2.0149	2.3026	1.0116	0.9144
542365011112141	312.00	8.20	24.00	2.74	7.74
542365011112141	5.7430	2.1041	3.1781	1.0068	2.0458

542975011713422	58.00	4.33	12.00	4.50	1.14
542975011713422	4.0604	1.4663	2.4849	1.5041	0.1341
543115011793422	31.00	6.00	5.00	3.00	1.54
543115011793422	3.4340	1.7918	1.6094	1.0986	0.4329
543195015123331	30.00	5.00	5.00	4.00	0.01
543195015123331	3.4012	1.6094	1.6094	1.3863	-5.0356
543275011123341	321.00	5.36	13.88	3.83	43.49
543275011123341	5.7714	1.6784	2.6301	1.3437	3.7726
552405011123341	77.00	5.00	8.50	2.60	2.60
552405011123341	4.3438	1.6094	2.1401	0.9555	0.9551
572465015211132	61.00	4.50	11.50	3.00	0.35
572465015211132	4.1109	1.5041	2.4423	1.0986	-1.0635
572685011123341	49.00	6.00	15.00	3.33	0.39
572685011123341	3.8918	1.7918	2.7080	1.2040	-0.9400
572715011123341	133.00	4.00	9.50	3.33	5.13
572715011123341	4.8903	1.3863	2.2513	1.2040	1.6357
582395018133311	51.00	5.00	9.00	1.00	0.69
582395018133311	3.9318	1.6094	2.1972	0.0	-0.3704
582545011123341	619.00	11.00	18.00	6.93	4.00
582545011123341	6.4281	2.3979	2.8904	1.9363	1.3864
582565011123341	207.00	6.50	12.60	4.09	1.64
582565011123341	5.3327	1.8718	2.5337	1.4088	0.4940
592525011123341	156.00	6.50	12.00	3.40	1.59
592525011123341	5.0499	1.8718	2.4849	1.2238	0.4632

APPENDIX B

The purpose of this section is to provide a convenient reference to the second-order equilibrium conditions for profit maximization. This section is analogous to Henderson and Quandt [1958, pp. 61-62]. Given

$Q = f(X_1, X_2) = AX_1^{\alpha_1} X_2^{\alpha_2}$, the equilibrium condition under perfect competition requires

(a) $f_{11} < 0,\ f_{22} < 0$

and

(b) $\begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} > 0$

$$f_{ii} = \frac{\partial^2 Q}{\partial X_i^2} = \alpha_i(\alpha_i - 1)\frac{Q}{X_i^2}$$

which is negative if $\alpha_1 < 1$ which is the stability condition for part a. Expanding $f_{11}f_{22} - f_{12}^2$, we find

$$\left[\alpha_1(\alpha_1-1)\frac{Q}{X_1^2}\right]\left[\alpha_2(\alpha_2-1)\frac{Q}{X_2^2}\right]-\left[\frac{\alpha_1\alpha_2}{X_1X_2}\right]^2=(1-\alpha_1-\alpha_2)\left[\frac{\alpha_1\alpha_2}{X_1^2X_2^2}\right]Q^2$$

which is positive if $\alpha_1 + \alpha_2 < 1$. If $\alpha_1 + \alpha_2 = 1$ we have the Samuelson indeterminacy case and if $\alpha + \alpha_2 > 1$ the stability condition is violated.

APPENDIX C

Since the question of the influence of the weights in the output measure was introduced in Chapter III, two additional regressions were run. The first regression used .9 for add time and .1 for cycle time which yielded

$$C_j^* = 2.974 + \underset{(0.040)}{0.453Q_j^*} + \underset{(0.163)}{0.644P_{1j}^*}$$

$$+ \underset{(0.148)}{0.220P_{2j}^*} + \underset{(0.132)}{0.176P_{3j}^*} \qquad R^2 = .676$$

where the symbols are the same as in the previous equation. Again the estimates of the parameters have the expected sign but the coefficient for administration and operations personnel are not different from zero at the 10% level. Table C-I indicates the t-ratios for the estimated coefficients.

TABLE C-I.--Test Statistic with $W_A = .9$, $W_C = .1$

Value	t-ratio	Significance Level
Y	11.466	1%
P_1	3.958	1%
P_2	1.489	20%
P_3	1.330	20%

The high R^2 and F value equaling 57.225 can be interpreted as justification for the functional form of the production function where most of the variation in cost measure is attributed to variations in output and systems programming. The weight for add-time was changed to .1, cycle time to .9, and regression resulted in

$$C_j^* = 1.903 + \underset{(0.043)}{0.367Y_j^*} + \underset{(0.188)}{0.722P_{1j}^*} + \underset{(0.171)}{0.323P_{2j}^*} + \underset{(0.152)}{0.338P_{3j}^*}$$

$$R^2 = .566$$

with the elasticities all positive and significantly different from zero at the 10% level (see Table C-II) with high R^2 and F = 35.716.

TABLE C-II.--Test Statistic with $W_A = .1$, $W_C = .9$

Value	t-ratio	Significance Level
Y	8.377	1%
P_1	3.838	1%
P_2	1.891	10%
P_3	2.226	1%

The evidence does not allow us to state that the estimates of economies of scale are altered with the choice of weights in the output measure.

The study by Knight was discussed in the literature review and it was considered interesting to see what would result if we used Knight's output measure rather than OUTM. Unfortunately, Knight does not have all the machines in our sample but, using only those machines for which Knight provides data, we found the following results for n = 100.

$$C^*_j = -2.198 + \underset{(0.056)}{0.394Q^*_j} + \underset{(0.190)}{0.691P^*_{1j}} + \underset{(0.172)}{0.342P^*_{2j}} + \underset{(0.150)}{0.428P^*_{3j}}$$

$$R^2 = .591$$

where the symbols are the same except Y'* is output as defined by Knight. The elasticities are positive as expected; the negative intercept is questionable but nothing will be said about its interception.

As a comparison, the data were run using the OUTM measure which resulted in

$$C^*_j = 2.518 + \underset{(0.043)}{0.406Q^*_j} + \underset{(0.163)}{0.646P^*_{1j}} + \underset{(0.145)}{0.285P^*_{2j}} + \underset{(0.131)}{0.260P^*_{3j}}$$

$$R^2 = .679$$

If we look at the sum of the output elasticities again, we find no evidence to suspect the coefficients are different. The general fit of the equation using OUTM is better than that obtained by using Knight's measure. Because of the difficulty of obtaining measures like Knight's, OUTM appears to be the low cost method of obtaining output measure.