RONALD W. SHEPHARD

# Theory of Cost and Production Functions

PRINCETON LEGACY LIBRARY

*THEORY OF COST*
*AND PRODUCTION FUNCTIONS*

PRINCETON STUDIES IN MATHEMATICAL ECONOMICS

Edited by David Gale and Harold W. Kuhn

1. Spectral Analysis of Economic Time Series, by C. W. J. Granger and M. Hatanaka
2. The Economics of Uncertainty, by Karl Henrik Borch
3. Production Theory and Indivisible Commodities, by Charles Frank, Jr.
4. Theory of Cost and Production Functions, by Ronald W. Shephard

# *THEORY OF COST AND PRODUCTION FUNCTIONS*

*BY RONALD W. SHEPHARD*

*PRINCETON UNIVERSITY PRESS*

*PRINCETON, NEW JERSEY*

*1970*

LCC 75-120762
ISBN 0-691-04198-9

This book is composed in Fotosetter Times Roman

Printed in the United States of America
by Princeton University Press

*To*
*Hilda Maloy Shephard*

# *PREFACE*

Fifteen years have passed since my original monograph on cost and production functions† was published by Princeton University Press. Until recently there has been little if any reference to the work. The monograph has long been out of print and for some time I have been aware that individuals were seeking copies of the apple green booklet. Some of the ideas and conceptions of the monograph seem to have percolated to the surface of theoretical and econometric studies, renewing my interest in the subject.

About three years ago, my friend Oskar Morgenstern, whose interest in my early work on cost and production functions was largely responsible for the publication of the first monograph, began urging me to rewrite the booklet, and I set myself the task of doing so. It soon became evident that considerable modernization and extension of the subject matter was desirable, and in this book I have tried to develop the theory of cost and production functions in a more complete and systematic way. The subject matter is essentially mathematical and, although there is a predilection in mathematical economics for the use of symbolism in the place of words, I have not hesitated to use words when the precision of the discussion is not lost. The mathematical arguments are simple and direct, although perhaps inelegant, but minimally invoking theorems which disconnect the reasoning.

One may ask: why devote a book to the theory of cost and production functions? In a narrow sense, the mathematical economic theory of production is a theory of cost and production functions, with the central topic being an understanding of the possibilities of substitution between the factors of production to achieve a given output. Optimization in production planning is yet another topic now largely being pursued in Operations Research, where the models reflect the peculiarities of the individual firm and the difficulties are mainly computational and algorithmic. Econometric studies of capital expansion, returns to scale and factor substitution lean heavily upon a clear understanding of cost and production functions, complicated by problems of aggregation which are still unsolved. Realistically, one may hope to advance the economic theory of production by concentrating upon the core of this subject, i.e., cost and production functions.

Discussions of this subject are at best confusing. I have not tried to reference comprehensively the work of others, this being a distracting

† Ronald W. Shephard: Cost and Production Functions, Princeton University Press 1953.

chore. The references used in this connection have been chosen at my convenience to contrast viewpoints.

The material for this book has been developed in a series of preliminary reports issued at the Operations Research Center, College of Engineering, University of California, Berkeley. These reports have not been referenced because of their limited distribution.

I take this opportunity to express my gratitude to Oskar Morgenstern for his supporting interest in the research from which this book has evolved. I also wish to express my indebtedness to Dr. Stephen Jacobsen for his reading of my manuscript as it developed and the many helpful suggestions which he has made.

I gratefully acknowledge the financial support of Professor Morgenstern's Econometric Research Project at Princeton University, supported by the Office of Naval Research, and the support of the Office of Naval Research and the National Science Foundation research grants to the Operations Research Center at the University of California, Berkeley, both of which assisted the research which has led to the publication of this book. Also, I take this means of expressing my appreciation to Mrs. Linda Betters for typing the manuscript.

*June 1969*
*Berkeley, California*

RONALD W. SHEPHARD

# *TABLE OF CONTENTS*

# *THEORY OF COST AND PRODUCTION FUNCTIONS*

# CHAPTER 1

## *INTRODUCTION*†

In economic theory the production function is a mathematical statement relating quantitatively the purely technological relationship between the output of a process and the inputs of the factors of production, the chief purpose of which is to display the possibilities of substitution between the factors of production to achieve a given output. The distinct kinds of goods and services which are usable in a production technology are referred to as the factors of production of that technology and, for any set of inputs of these factors, the production function is interpreted to define the maximal output realizable therefrom.

The more or less traditional treatments of the production function exclude free goods as inputs and require that the function express the variable, substitutional and limitational character or other qualifications of the factors of production peculiar to some hypothetical production unit. Sune Carlson [5] states: "As regards the productive services which constitute the input to the technical unit during the period we shall only consider the services which are limited in supply." He makes a distinction between fixed and variable productive factors and asserts: "The production function, it must be remembered, is defined relative to a given plant; that is certain fixed services." Erich Schneider [25] distinguishes between "substitutional" and "limitational" factors and expresses the production function in terms of substitutional inputs with side equations between output and each limitational factor. Samuelson [24] explains that: "The production function must be associated with a particular institution (accounting, decision making, etc.), and must be drawn up as of any unique circumstances pertaining to this unit." All of these qualifications are made in the context of a general theory of production!

Sometimes the productive factor inputs are classified as to whether they are flow or stock quantities, the former referring to labor services, raw materials, energy, etc. and the latter designating real capital goods such as plant, machinery and equipment. Krelle [17] makes this distinction and introduces stock variables in the production function along with flow variables for the inputs of consumable factors. In order to exhibit the structure of investment planning, V. L. Smith [27] puts forth the notion of a "stock-flow production function" in which capital stock inputs are freely variable along with current input flows for treating

† The preliminary paragraphs of this chapter have appeared in *Unternehmensforschung*, Band 11, Heft 4, 1967, and are used here with the permission of Physica Verlag, Würzburg.

hypothetical alternative production plans, but once the physical configuration has been chosen the real capital inputs can no longer be varied like current inputs.

Also, it is common in the economic theory of production to distinguish between "short run" and "long run" production functions, the form of the function being essentially different in the two cases. But the production function is ideally a statement of purely technological alternatives, without regard to their execution, and one need not define different production functions for these two situations. In doing so, institutional conditions of specific economic planning are brought into the definition, confusing the purely technological (engineering) character of the production function. The significance of the short run is that there are constraints on the amounts and kinds of factor inputs and these qualifications are best kept in this form, leaving the production function as a statement of unconstrained technological alternatives relative to some horizon of planning, encompassing arrangements which have not yet been realized as well as those which have been put into operation.

The viewpoint taken in this study is that neither the exclusion of free goods nor the requirement that the production function express the variable, substitutional, consumable character or the limitational, fixed stock character of the productive factors, as qualifications peculiar to a particular production unit, are logically necessary for the definition of the production function.

The production function is regarded here as a mathematical construction for some well defined production technology. This technology consists of a family of conceivable and feasible engineering arrangements, not restricted necessarily to particular realizations found in practice and possibly spanning historical changes in the application of the technology. Once defined, the technology implies a certain set of factors of production and no limitations will be put upon the inputs of these factors both as to type and amount available. Thus the production function will be taken to describe the unconstrained technical possibilities of a technology without limitation to any existing or realized production units.

The productive factors are not restricted to economic goods and services, i.e., those with a positive market price, because this implies some particular resource availabilities relative to demand in an exchange economy, which is irrelevant to the technical alternatives defined by the production function. However, the situations of interest in economics are those for which not all factors of production are free.

No limitations will be put upon the available amounts of the factors of production, because this implies reference to some particular production unit which confounds the notion of a production function with some implicit economic decisions or production plan, the variety of which is unlimited, preventing a clear, unambiguous and generally applicable definition of the production function.

Both the input and output variables will be defined as time rates. The unconstrained service flows from real capital (plant, machinery and equipment) imply freely variable physical counterparts, in whatever units and capacity they arise in the technology, and unutilized capacity of a physical item is merely excess input flow of the related capital service which does not hinder output.

If the production function is to define purely technological possibilities, the available means of a firm or other production unit are not relevant. Such limitations merely prescribe a particular realization of the technology which may be considered by imposing constraints on the input flows which restrict the analysis to a particular subset of the factor input space. For example, if the production unit uses only certain machinery and equipment, then the input flows of these factors can be bounded by the positive capacities involved, while the input rates of other real capital conceived for the technology but not available to the production unit may be bounded by zero. In such circumstances the substitution possibilities of one factor for another are likely to be limited, i.e., specific realizations of the technology will have a high degree of factor complementarity, whereas the substitutability of factors sought in economics will arise for a broadly defined technology not constrained to particular realizations, which is the kind of production structure most interesting for economic planning. These matters will become clear when we consider such constraints as defining subsets of input vectors available to the firm.

In Chapter 2, the foregoing conception of the production function is developed in some detail. From an engineering viewpoint, the structure of production may be conceived as a family of production possibility sets, specifying for each nonnegative output rate the set of input vectors which yield at least the given output rate. On this structure the production function may be defined as the maximum output rate obtainable for any given nonnegative input vector, giving to it the traditional meaning in economic theory. Conversely, one may postulate the existence of a production function with certain properties and determine the production possibility sets as the level sets of this function, and the *uniqueness of the production function is a question of some interest.* These ideas are developed at length in Chapter 2.

Of particular significance to the theory of production for study of returns to scale is the discovery of the circumstances under which cost data may be "deflated" to real terms by an index function of the prices of the factors of production. To pursue this matter, a class of production functions was defined in the first Princeton monograph [26] and named homothetic. There it was shown that the cost function factors into a function of output rate and a linear homogeneous function of the prices of the factors of production (an index function of prices), if and only if the production function is homothetic. Interestingly, the ACMS production function [2] and Uzawa's extension of this function [29], the Cobb-Douglas production function and its modifications, used for the study of returns to scale, are all very special cases of homothetic production functions. One might speculate that these endeavors could profit from a conscious use of the general definition of a homothetic production function, taking some special mathematical form for the linear homogeneous function of the prices into which the cost function factors, but not forcing any special form for the other term (i.e., the function of output rate), the inverse function of which defines the returns to scale.

In Chapter 2, a slightly more general definition of homotheticity† will be given, with a discussion of the properties of the corresponding production possibility sets. Also, a brief discussion of a classification of the factors of production is presented which seems to be more useful than the traditional notion of complementarity, and a discussion of the production function of a limited unit or firm is given. The chapter is closed with a discussion of the law of diminishing returns which provides a proof of a form of the law without assumptions on the fine structure of production, and the implications for commonly used production functions like the Cobb-Douglas and CES are developed.

In Chapter 3 the distance function of a production structure is introduced as an alternative to the production function. The properties of this function are determined and the special form of the distance function for homothetic production structures is deduced. At first it may seem strange that the distance function is considered. But, as will be seen in the subsequent chapters, the minimum cost function is a distance function of a price-output cost structure and the duality between cost and production function is naturally formulated in terms of these distance functions. When production correspondences are considered in Chapter 9, the significance of the distance function will become further apparent, because it affords a means of investigating the possibilities for a joint production function.

Chapter 4 is devoted to the factor minimal cost function, i.e., the traditional cost function defining the minimum total cost rate for any

† To avoid an unnecessary assumption of continuity.

output rate u and vector p of the prices of the factors of production, with the input rates of the factors of production adjusted to yield minimum total cost. This function is called the "factor minimal" cost function in order to distinguish it from another cost function introduced in Chapter 7 for discussion of the duality between cost and production function. The properties of the factor minimal cost function are stated and proved in Chapter 4, after which the special form and properties of this function for homothetic production structures are developed.

In Chapter 5 it is observed that the factor minimal cost function has the properties of a distance function and it is shown that indeed it is a distance function for a certain cost structure consisting of a family of subsets of price vectors for the factors of production. The properties of the price sets in this cost structure are determined, and the special properties of the cost structure corresponding to homothetic production structures are developed. Then as a dual (to be shown later) to the production function (factor maximal output function) a cost limited output function† is defined on the cost structure, which provides for any nonnegative price vector of the factors of production the supremal output which can be obtained for any positive cost rate. When differentiable, this function enables a calculation of the marginal productivity of money capital to supply the cost of production, and, in the case of the cost structure for a homothetic production structure, simple formulas are given.

Chapter 6 is addressed to the aggregation problem for the theory of cost and production functions. Certain criteria are set fortn for aggregating the input variables and prices of the factors of production. An aggregation of homothetic production, cost and cost limited output functions in terms of one variable for inputs and one variable for prices is determined and shown to satisfy the criteria. The usual aggregation for Cobb-Douglas production and cost functions is then shown to satisfy the criteria, and an aggregation of the ACMS production and cost function is given which yields the same aggregate form as that for the Cobb-Douglas function. Then it is shown that for a certain class of homothetic production and cost functions the aggregate form is a Cobb-Douglas production and cost function. The chapter is closed with a demonstration by construction that a generalization of homothetic cost and production functions can be aggregated to satisfy the criteria.

As preparation for the duality between cost and production functions, a price minimal cost function defined on the cost structure is defined in Chapter 7. It is shown that this function has the same properties as the distance function of the production structure from which the cost

† Also known as the Indirect Production Function.

structure was derived, and, when treated as a distance function in the factor input space, it defines a production structure identical to the parent production structure.

Thus in Chapter 8, where the duality between cost and production functions is discussed, the production possibility sets of the production structure and the price sets of the cost structure are shown to be duals, derivable from each other by dual cost minimizations which determine the factor minimal and price minimal cost functions as dual distance functions. A duality between the production function and the cost limited output function is then developed by showing that they may be determined in terms of each other by dual maximum problems. After which, the elegant geometric relationship between the dual cost and production structures is demonstrated, and a theorem is proved establishing homotheticity as an if and only if property for the factorization of the cost function into a function of output rate and a linear homogeneous function of the prices of the factors of production. The chapter is closed with a discussion of dual expansion paths in the cost and production structures.

All of the previous considerations apply to production structures with a single output, but they are extendable to technologies with multiple or joint outputs. In Chapter 9, the concept of a production correspondence P is introduced for joint outputs by treating the production relationship as a mapping of each input vector into a subset of output vectors which can be realized with the given input vector. Certain well defined properties of this mapping are assumed, which are a natural extension of those assumed in Chapter 2 for a production technology with single output. The inverse correspondence L of P is a mapping of each output vector into a subset of input vectors which yield at least the given output vector, analogous to the level sets of the production function. Thus, for the production relationship of a technology with multiple outputs, we have point to set mappings defining outputs realizable with each input vector and, inversely, point to set mappings defining for each output vector a subset of input vectors yielding at least the given output vector. The assumptions made for the mapping P imply certain properties for the inverse mapping L which are analogous to the properties of the level sets of the production function for a technology with single output.

Except for modifications to permit nondisposable outputs, the properties taken for the production correspondence P follow those used by my erstwhile student Dr. Stephen Jacobsen in his doctoral thesis [15], where he extended my duality between cost function and distance function for the level sets of the production function to the cost function and distance function of the map sets of the inverse correspondence L, applying my notions of distance function and cost structure induced by the cost function and giving a somewhat more general definition of homotheticity of input structure for the production correspondence.

These formulations of the production relationship for joint outputs include those of the production function for a single output as a special case. Indeed, the production correspondence with single output defines a production function with the properties assumed in Chapter 2, and conversely the production function induces a suitable production correspondence with one component output vectors.

After developing these notions for the production correspondence, the property of homotheticity is considered. Both the input structure and the output structure of the production correspondence may independently have a homothetic property. My definition of homothetic input structure is equivalent to that used by Jacobsen, although stated in different terms. Symmetrically, the output structure may also have a homothetic property and a definition of his property is given. If both the input structure and the output structure of the production correspondence are homothetic, the production relationship is homogeneous of degree one.† This discussion of homotheticity of structure contains certain propositions concerning the representation of the input sets, and also the output sets of the production correspondence when they are homothetic.

The next topic of Chapter 9 is the definition of two distance functions, one for the input sets of the inverse correspondence L, like that used by Jacobsen, and another for the output sets of the correspondence P. The properties of these two distance functions are developed in some detail, and demonstration is given that they may be used to define the mappings P and L of the production correspondence. In terms of these two distance functions, the question of the existence of the frequently used joint production function is examined. It is found that the joint production function does exist if outputs are disposable and the production correspondence is continuous. However, the two distance functions can always be used separately to define the boundary substitutable output vectors for a given input vector (i.e., the output isoquants) and the boundary substitutable inputs to yield a given output vector (i.e., the input isoquants).

Chapter 9 is concluded with a discussion of the special forms which the two distance functions take when the related structure is homothetic.

Chapter 10 is initiated by consideration of the cost function defined on the input sets of the production correspondence as the minimal cost

† The extended definition of homotheticity, given in Section 10.5 and used following, avoids this result.

of attaining an output vector u with a price vector p for the factors of production, showing that it has properties analogous to those set forth in Chapter 4 for the cost function related to a production function with single output. Then a formulation is given for the revenue (or benefit) function defined on the output sets of the correspondence P as the maximal revenue (or benefit) obtainable with an input vector x and prices (or unit values) r for the outputs, depending upon whether the outputs are disposable (or nondisposable). The term benefit is used to indicate coverage of situations where the outputs do not have market prices but may have unit values (positive or negative) in accordance with some social weighting system, the negative values applying to undesirable outputs. Just as in the case of the cost function for a production relationship with single output, the cost function for the production correspondence induces a cost structure. This cost structure is definable by a correspondence which maps an output vector into a subset of price vectors for the inputs which yield at least unit minimal cost in attaining the output vector; or any level of cost for that matter, since the cost function is homogeneous of degree one in the price vectors for the factors of production. Similarly, the revenue (or benefit) function induces a revenue (or benefit) structure which may be taken as a correspondence mapping an input vector into a subset of output price (or unit value) vectors which yield at most unit revenue (or benefit) for the maximal revenue (or benefit) which may be obtained by using the input vector; and the unit level is not restrictive for defining this structure, since the revenue (or benefit) function is homogeneous of degree one in the price (unit value) vector for the outputs. The properties of these two correspondences are considered in some detail, and it is shown that the cost function and the revenue (or benefit) function are distance functions for the map sets of the correspondence which they induce.

By way of a digression in order to show the connection with the cost limited maximal output function introduced in Chapter 5, two additional correspondences are defined. One is a cost limited output correspondence mapping a price vector for the factors of production into a subset of output vectors for which the minimal cost of attaining the output vector with the given price vector for the inputs is less than unit value. Similarly, the other correspondence is a mapping of price (or unit value) vectors for outputs into a subset of input vectors for which the maximal revenue (or benefit) obtainable with the input vector is greater than unit value. The first correspondence enables one to determine for any input price vector the output vectors which can be attained at cost less than any given level of cost, and the second provides a determination for any output price (or unit value) vector the input vectors which will yield more than any given level of maximal revenue (or benefit).

After showing the special forms taken by the cost and benefit functions if the input and output structures of a production correspondence are homothetic, the final section of Chapter 10 is devoted to a consideration of returns to scale and for this purpose an extended definition of homothetic structure is given. With this extended definition, the alterations of previous propositions for homothetic structures are straightforward.

The final chapter of the book takes up various dualities, starting with the extension for production correspondences of the duality between cost function and distance function for the input sets of a production relationship with single output (developed in Chapter 8). This extension was made by Jacobsen [15] in his doctoral thesis. Another duality between the revenue (or benefit) function and the distance function for the output sets of the production correspondence is also developed. The first duality consists of dual cost minimization problems which enable one to determine the cost function and distance function for the input sets of the production correspondence in terms of each other, and the second duality provides a determination of the revenue (or benefit) function and the distance function for the output sets of the production correspondence in terms of each other by dual revenue (or benefit) maximization problems. As a consequence of these two dualities, it is shown that the cost function and the revenue (or benefit) function factor in a certain way if and only if the input sets and the output sets of the production correspondence respectively have homothetic structure. For the cost function, the factorization is into a product of a function of the output vector and a function of the price vector for inputs, and the benefit function factors into a product of a function of the input vector and a function of the price (unit value) vector for outputs. These factorizations provide simple forms for cost and revenue (or benefit) function which enable one to express price deflated minimal costs as a function of output vector alone and price deflated maximal revenue (or benefit) as a function of input vector alone. Correspondingly, the distance functions for the input sets and the output sets of the production correspondence also take simple factored forms if and only if these two structures are homothetic.

The next three dualities considered are directed to the determination of accounting (shadow) prices for input vectors (given prices (or unit values) for outputs), for output vectors (given prices for inputs) and simultaneously for both output and input vectors. These dualities are obtained by recombination of the two dual cost minimization problems and the two dual revenue (or benefit) maximization problems described above. The connection between these three dual problems and the duality arising in mathematical programming is only incidental, because the dual program in mathematical programming is merely one aspect of a saddle point problem for the Lagrangian of the primal problem, a device which is not of immediate economic significance except in the case of the simple linear model of production.

In the final section of Chapter 11, the constant coefficient model of production is discussed as a linear production correspondence. There it is shown that, although differently motivated, the dual problems in linear programming, one for imputing input prices and another for imputing output prices, are special forms of the first two of the three dualities introduced for determination of shadow prices with more restrictive constraints for the prices imputed.

A discussion of the aggregation problem for production correspondences is not given, because, due to the symmetry of definition for homotheticity of input structure and output structure, the arguments of Chapter 6 on the problem of aggregation may be applied for either the input structure or the output structure of the production correspondence.

Throughout the exposition to follow no strict attempt has been made to avoid repetition of argument in different contexts or restatement of properties, for the purpose of enabling the reader to follow the discourse without frequent referral to the text of previous sections. Propositions are numbered consecutively as statements which summarize arguments and they are used when applicable as parts of subsequent arguments.

Some frequently used mathematical notions are included in Appendix 1 and Appendix 2. Since the results of the first eight chapters carry over for utility functions, they are summarized in Appendix 3.

# CHAPTER 2

## *THE PRODUCTION FUNCTION*†

### 2.1 Definition of a Technology

A production technology consists of certain alternative means, arrangements of these means and uses of materials and services by which goods or services may be produced. The distinct goods and services which may be used as inputs to the technology are called factors of production. Free goods or services are not excluded as factors of production, since market prices have no bearing upon the technical roles of these inputs. The technology exists independently of the political and social structure in which it may operate and also of the scarcity of inputs, i.e., it is a blueprint for production.

At this point of our study, it is assumed that a single good or service is obtainable as an output of the technology.†† Let $u \in [0, +\infty)$ denote the output rate and take $x = (x_1, x_2, \ldots, x_n)$ to denote the input rates of the factors of production. The input vector x ranges over the non-negative domain D of a Euclidian space $R^n$, i.e., $x \in D$, where

$$D = \{x \mid x \geqq 0, x \in R^n\}. \qquad (1)$$

It is not assumed that x must be strictly positive for u to be positive, i.e., some of the factors of production may be substituted completely for others.

*Definition:* A production input set L(u) of a technology is the set of all input vectors x yielding at least the output rate u, for $u \in [0, +\infty)$.

Obviously not all input vectors x belonging to an input set L(u) are technologically efficient. The efficient subset E(u) of a production input set L(u) is given by the following definition:

*Definition:* $E(u) = \{x \mid x \in L(u), y \leq x \Rightarrow y \notin L(u)\}$.†††

A production technology is defined as follows:

*Definition:* A production technology is a family of input sets T: L(u), $u \in [0, +\infty)$ satisfying:

† The contents of this chapter, except for Section 2.7, have appeared in Unternehmensforschung, Band II, Heft 4, 1967, and are used here in modified form with the permission of Physica Verlag, Würzburg..

†† See Chapter 9 for extension to multiple outputs.

††† $y \leqq x \Rightarrow y_i \leqq x_i$, $i = 1, 2, \ldots, n$.
$y \leq x \Rightarrow y_i \leqq x_i$, $i = 1, 2, \ldots, n$, $y \neq x$.

P.1 $L(0) = D, 0 \notin L(u)$ for $u > 0$.

P.2 $x \in L(u)$ and $x' \geqq x$ imply $x' \in L(u)$.

P.3 If (a) $x > 0$, or (b) $x \geq 0$ and $(\bar{\lambda} \cdot x) \in L(\bar{u})$ for some $\bar{\lambda} > 0$ and $\bar{u} > 0$, the ray $\{\lambda x \mid \lambda \geqq 0\}$ intersects $L(u)$ for all $u \in [0, +\infty)$.

P.4 $u_2 \geqq u_1 \geqq 0$ implies $L(u_2) \subset L(u_1)$.

P.5 $\bigcap_{0 \leqq u \leqq u_0} L(u) = L(u_0)$ for $u_0 > 0$.

P.6 $\bigcap_{u \in [0, +\infty)} L(u)$ is empty.

P.7 $L(u)$ is closed for all $u \in [0, +\infty)$.

P.8 $L(u)$ is convex for all $u \in [0, +\infty)$.

P.9 $E(u)$ is bounded for all $u \in [0, +\infty)$.

The Properties P.1, . . . , P.9 are taken as valid for any technology. Property P.1 states merely that any nonnegative input vector yields at least zero output (a truism), and positive output cannot be obtained from a null input vector. Property P.2 implies disposability of inputs. For example, if chemical fertilizer is used as an input with land to produce a crop and excessive amounts of fertilizer have been provided, one merely disposes of the surplus. Fortuitous events, such as floods supplying excess water, are not encompassed. Excess capacity of machinery and equipment imply merely that the services of such capital are foregone. Thus, the technology is regarded as a rational, controllable arrangement.

Property P.3 states first that any output rate $u \in [0, +\infty)$ can be realized by scalar magnification of a positive input vector x, although not necessarily in an efficient way, and second that, if a positive output rate can be obtained by scalar magnification of a semi-positive input vector x, any null inputs of x are not required for production and the same attainability of all output rates holds by scalar magnification of the semi-positive input vector x. Divisibility of output rate is not implied.

Property P.4 is clearly appropriate, since an input vector yielding at least an output rate $u_2 \geqq u_1$ also yields at least $u_1$, and Property P.6 is merely a precise way of stating that an unbounded output rate cannot be attained by a bounded input vector.

Properties P.5 and P.7 have only mathematical significance. Property P.5 is imposed in order to guarantee the existence of the production function $\Phi(x)$ as the maximum output rate attainable with x. Property P.7 is imposed in order to be able to define the production isoquant for an output rate u as a subset of the boundary of the input set L(u) relative to $R^n$.

Property P.8 is valid for time divisibly-operable technologies. For example, if $x \varepsilon L(u)$, $y \varepsilon L(u)$ and $\theta \varepsilon [0,1]$, the input vector $[(1 - \theta)x + \theta y]$ may be interpreted as an operation of the technology a fraction $(1 - \theta)$ of some unit time interval with the input vector x and a fraction $\theta$ with y, assuring at least the output rate u.† Nothing is implied about the efficiency of such an operation.

Property P.9 is imposed as an obvious physical fact that no output rate is attained efficiently (in a technological sense) by an unbounded input vector.

In the foregoing definition of a technology, nothing is assumed which is peculiar to any particular physical system of production. Substitutions between the factors of production are permitted, both as alternative and complementary means of production. The family of sets L(u) defines the input unconstrained technical possibilities.

From the definition of the efficient subset E(u) of a production input set L(u), it is clear that the technologically efficient input vectors belong to the boundary of L(u). Because, suppose $x \varepsilon$ interior L(u). Then there would exist a spherical neighborhood $S_\delta(x)$, centered at x, composed entirely of points of L(u), implying $y \varepsilon L(u)$, $y \leq x$, a contradiction.

The set of efficient points E(u) need not be closed, because there is a counterexample to closure. See [1]. Even so, it is sufficient for our purposes to use the closure $\bar{E}(u)$ of E(u). Note that $\bar{E}(u) \subset L(u)$, since L(u) is closed.

It is necessary to verify that the efficient subsets E(u) are not empty for all $u \varepsilon [0, +\infty)$. Clearly, the null input vector is efficient for $u = 0$. Hence, consider $u > 0$ and let

$$B_R(0) = \{x \mid \|x\| \leqq R, x \varepsilon R^n\},\ R > 0$$

be a closed ball centered at $x = 0$ with radius R. One may choose R large enough so that $B_R(0) \cap L(u)$ is a nonempty, closed and bounded convex subset of L(u). Let

$$\|x^0\| = \text{Min}\ \{\|x\| \mid x \varepsilon B_R(0) \cap L(u)\}.$$

The vector $x^0$ exists, since it minimizes a continuous function

$$\|x\| = \left[\sum_{i=1}^{n} x_i^2\right]^{1/2}$$

over a nonempty, closed and bounded set, and $x^0 \varepsilon L(u)$. Moreover $x^0 \varepsilon E(u)$, because, if $y \leq x^0$, then $y \not\varepsilon B_R(0) \cap L(u)$ since $y \leq x^0$ implies $\|y\| < \|x^0\|$. Thus, the following proposition holds:

† Indeed the input vector $[(1 - \theta)x + \theta y]$ may have no meaning unless so interpreted.

*Proposition 1:* The efficient subset E(u) of a production input set L(u) is nonempty for all $u \varepsilon [0, +\infty)$.

Each production input set L(u) may be partitioned into the sum of the efficient subset E(u) and the set $D = \{x \mid x \geqq 0, x \varepsilon R^n\}$, and the following proposition holds:

*Proposition 2:* $L(u) = E(u) + D = \overline{E}(u) + D$.

The operator symbol + is used to denote the usual addition of sets, i.e., E(u) + D is the set of all input vectors of the form (x + y) where $x \varepsilon E(u)$ and $y \varepsilon D$.

First, we show that $(E(u) + D) \subset L(u)$. Since E(u) is nonempty, let $x \varepsilon E(u)$ and $y \varepsilon D$. $E(u) \subset L(u)$ implies $x \varepsilon L(u)$, and $(x + y) \varepsilon L(u)$ due to Property P.2 since $(x + y) \geqq x$. Thus, any input vector belonging to (E(u) + D) also belongs to L(u).

Next, we show that $L(u) \subset (E(u) + D)$. Let $y \varepsilon L(u)$ be arbitrarily chosen. The vector y belongs to the closed ball $B_{\|y\|}(0)$. Define

$$D_y = \{x \mid x \geqq 0, x \leqq y\}$$

and let

$$K(u) = \{\lambda x \mid x \varepsilon E(u), \lambda \geqq 0\}.$$

The intersection $L(u) \cap D_y$ is a bounded, closed subset of L(u). There are two cases to consider: (a) $y \varepsilon K(u)$, (b) $y \notin K(u)$ (see Figures 1 and

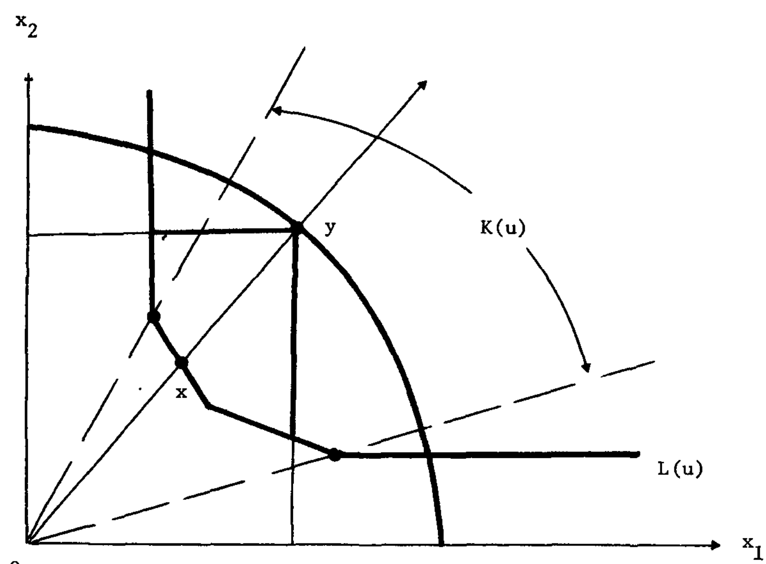


FIGURE 1: $y \varepsilon K(u)$

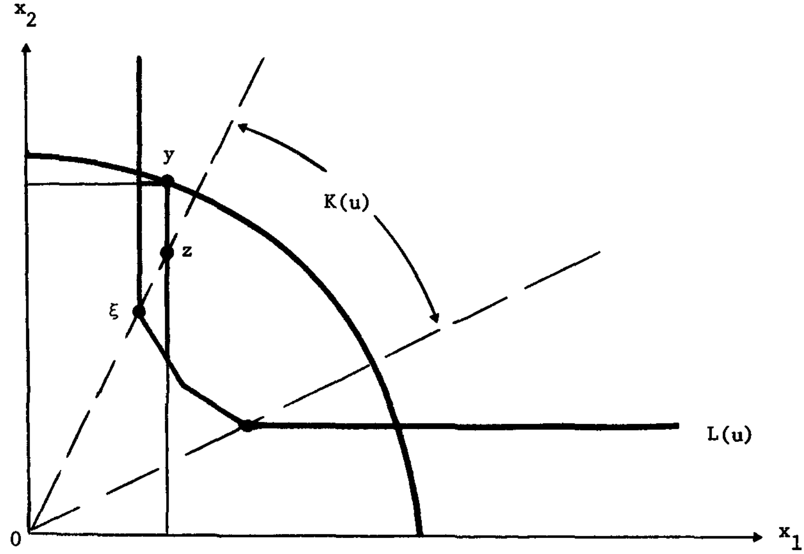


FIGURE 2: $y \notin K(u)$

2). If $y \in K(u)$, the ray $\{\theta y \mid \theta \geqq 0\}$ intersects E(u) at a point x, and $y = x + (y - x)$ with $(y - x) \in D$ since $y \geqq x$. Hence, for Case (a), $y \in (E(u) + D)$. In Case (b), consider†

$$\text{Min}\left\{\sum_{1}^{n} z_i \mid z \leqq y, z \in K(u) \cap D_y \cap L(u)\right\}$$

The set $K(u) \cap D_y \cap L(u)$ is not empty and the minimum exists. Let x denote the vector yielding this minimum. Then $x \in E(u)$ and $y = x + (y - x)$ with $y \geqq x$, so that $y \in (E(u) + D)$. Consequently $L(u) \subset (E(u) + D)$ and $L(u) = E(u) + D$.

The equality between L(u) and $(\bar{E}(u) + D)$ is verified simply, as follows: $(\bar{E}(u) + D) \subset L(u)$, because, if $z \in (\bar{E}(u) + D)$, $z = x + y$ with $x \in \bar{E}(u)$, $y \geqq 0$, and $x \in L(u)$ since $\bar{E}(u) \subset L(u)$, whence $z \in L(u)$ due to Property P.2. Conversely, $L(u) \subset (\bar{E}(u) + D)$, because $(E(u) + D) \subset (\bar{E}(u) + D)$, and $L(u) = (E(u) + D) \subset (\bar{E}(u) + D)$.

Another subset of the boundary of a production input set L(u), called the production isoquant, is of use in the theory of production.

*Definition:* The production isoquant corresponding to an output rate $u > 0$ is a subset of the boundary of the input set L(u) defined by

$$\{x \mid x \geq 0, x \in L(u), \lambda \cdot x \notin L(u) \quad \text{for} \quad \lambda \in [0,1)\}$$

† This proof suggested by K. Arrow (see Math. Reviews, 5460, 1969) is simpler than the original proof given by the author in Unternehmensforschung, Heft 4, 1967.

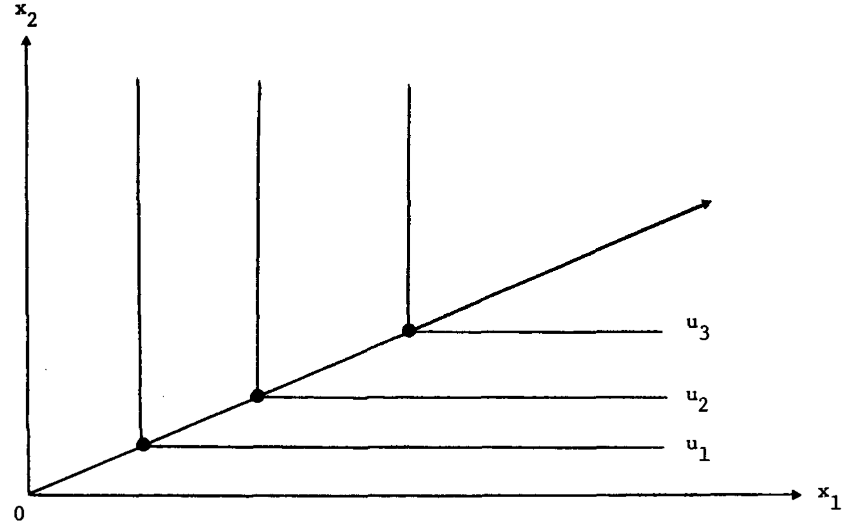


FIGURE 3: SINGLE FACTOR MIX ALTERNATIVE

The isoquant for $u = 0$ is $\{0\}$.

The production isoquant for $u \varepsilon [0, +\infty)$ is a closed subset of L(u), and the definition applies whether or not the output rate u exists as a $\text{Max}\{u \mid x \varepsilon L(u)\}$ for some $x \varepsilon D$ since the sets L(u) are defined for all $u \varepsilon [0, +\infty)$.

Various production isoquants are illustrated in Figures 3, 4, 5, 6 for two factors of production. In Figure 3, there is illustrated a technology for which the two factors of production can be used efficiently only in a fixed proportion, typical of the Leontief model of production. In Figures 4, 5, and 6, the efficient subsets of the isoquants are indicated by darkened lines. Figure 4 illustrates a technology with 4 alternatives of mixing two factors of production. Figure 5 illustrates that a factor of production need not be essential. Figure 6 shows the usual neoclassical continuity of substituting one factor for another, with both factors of production nonessential. Notice that in all four figures the efficient subsets are bounded, but the isoquants need not be bounded. For each figure the production input set L(0) is bounded by the positive axes with the null input vector the single efficient point. The production input sets are convex and closed, and for $u_1 < u_2 < u_3$ these sets are nonincreasing with each contained in its predecessor. Also if an input vector x belongs to one of the input sets any input vector at least as large as x also belongs to that input set. Further if $x \geq 0$ and the ray $\{\lambda x \mid \lambda \geqq 0\}$ intersects an input set for positive output, the ray inter-

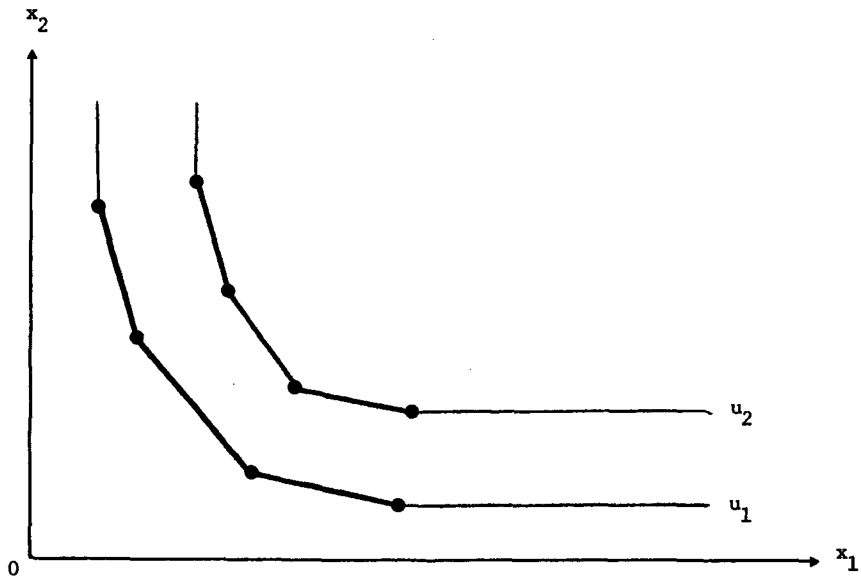


FIGURE 4: SEVERAL FACTOR MIX ALTERNATIVES

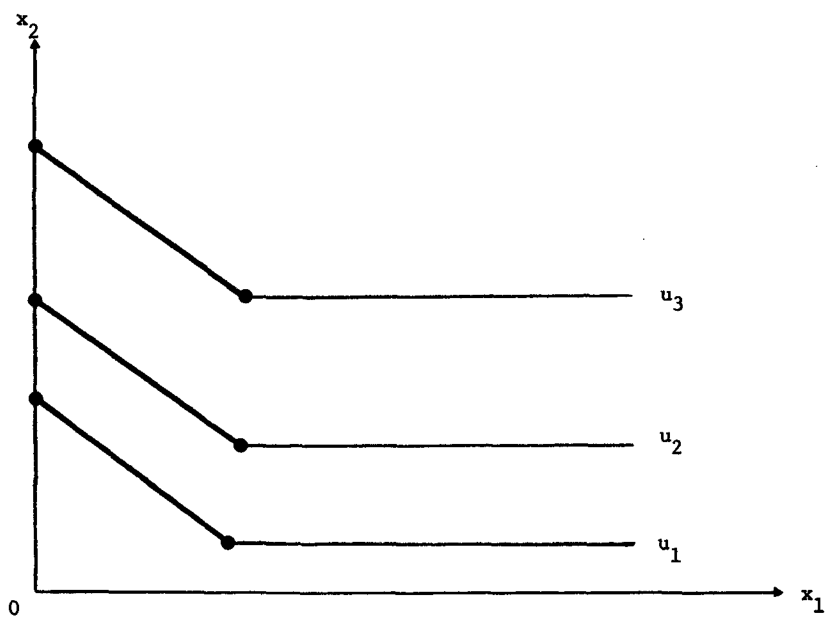


FIGURE 5: LINEAR SUBSTITUTION WITH ONE FACTOR NONESSENTIAL

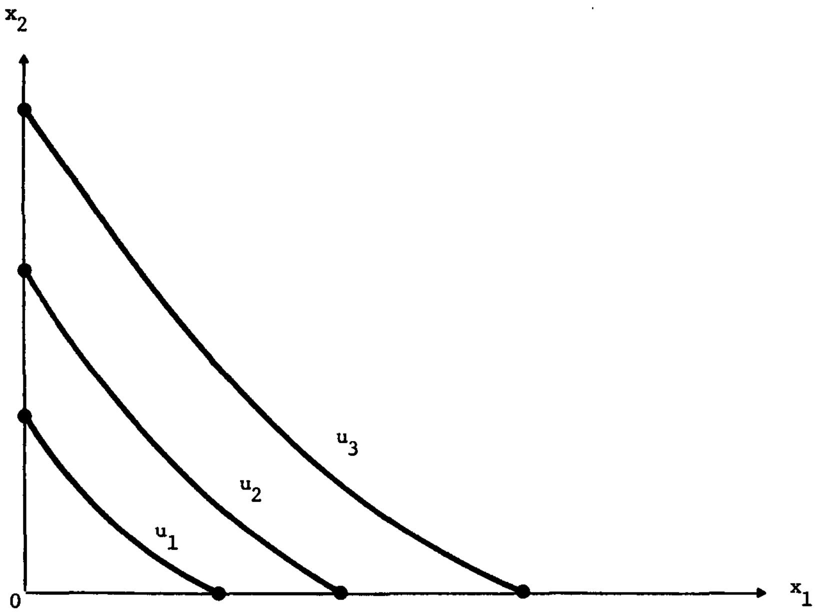


FIGURE 6: SMOOTH SUBSTITUTION WITH BOTH FACTORS NONESSENTIAL

sects all input sets. Properties P.5 and P.6 cannot be illustrated in such figures. Observe that for each figure L(u) = E(u) + D, and a similar decomposition holds in terms of the production isoquant.

## 2.2 Definition and Properties of the Production Function

The production function is a mathematical form defined on the production input sets of a technology, with properties following from those of the family of sets L(u), $u \varepsilon [0,+\infty)$ which can be best understood this way instead of making assumptions ab initio on a mathematical function.

For any input vector $x \varepsilon D$, consider a function $\Phi(x)$ defined on the sets L(u) by

$$\Phi(x) = \text{Max}\{u \mid x \varepsilon L(u), u \varepsilon [0,+\infty)\},\ x \varepsilon D \qquad (2)$$

giving to the production function $\Phi(x)$ the traditional meaning as the largest output rate obtainable with x. It is not obvious that Max exists for all $x \varepsilon D$, and this fact needs to be proved first.

Let $x \varepsilon D$ be chosen arbitrarily. The input vector x belongs to L(0), see Property P.1, and there exists a finite value $\bar{u} > 0$ such that $x \notin L(\bar{u})$, due to Properties P.4 and P.6. Hence, $\text{Sup}\{u \mid x \varepsilon L(u), u \varepsilon [0,+\infty)\} = u_0$ is finite. Then, it follows that $x \varepsilon L(u)$ for $u \varepsilon [0,u_0)$ and Property P.5