EXTENSIONS OF CONTOUR ANALYSIS IN ECONOMIC THEORY

By

J. A. Galbraith

Submitted in partial requirement for the Master of Commerce degree (in Economics).

McGill University August 1950

PREFACE

To a footnote on closed indifference curves in an article by Professor F.H. Knight, the writer first owes the direction of his interest to the general theme of this thesis. This interest led to further reading and, subsequently, to the setting down, in a class paper, of some preliminary ideas on the extension of existing indifference curve analysis. The favorable reception of this paper led to the consideration of expanding it into a thesis. Once this step was decided upon it was natural that the scope of the prospective thesis be broadened by including isoquant analysis. So arose the idea of extending contour analysis. At first it appeared to be an original idea, but as the investigation of the subject progressed it revealed that others had, to some extent, preceded the writer in this field. However, the literature does not offer a complete or full development of the theme.

Most of the thesis represents much laborious and independent thinking. How much represents original work is another matter on which the writer hesitates to express an opinion. Some originality may exist in the discussion of ridge lines in Chapter IV, in the analysis contained in Chapter V, in the treatment of thresholds in Chapter VI, and in a few miscellaneous ideas presented in Chapter VII. Possibly parts of Chapters XII and XIII are also original. However, aside from any question of originality, it may be claimed that the thesis, by gathering together and presenting as a unified whole all available knowledge on the subject, passes on to economic analysis a clearer and more refined conception of contour analysis than has hitherto existed.

In developing the theme of the thesis the writer was aided considerably by certain printed sources to which mere reference in a footnote, or bibliography, does not do full justice. These sources are as follows: F.H. Knight, "Realism and Relevance in the Theory of Demand", Journal of Political Economy, 1944; J.M. Clark, "Realism and Relevance in the Theory of Demand", Journal of Political Economy, 1946; H.Schultz, The Theory and Measurement of Demand, esp. pp607-628; and J.M. Cassels, "On the Law of Variable Proportions", Readings in the Theory of Income Distribution.

Outside the normal supervision and advice given by his research director, and apart from the literary sources noted in the text, no other assistance has been received by the writer.

McGill University
August, 1950 J. A. Galbraith

CONTENTS

CONTENTS

INTRODUCTION

This study is on the extension of contour analysis. The extension is an elaboration of existing contour analysis and not an application of that analysis to new fields. The object is to carry the analysis of indifference curves and isoquants to a logical conclusion. This entails a detailed consideration of areas of the contour maps not required for ordinary discussions of demand and production problems. Three preliminary arguments may be advanced for extending the analysis.

First, the extension of the analysis serves to sharpen powerful tools of economic analysis by presenting a clearer conception of these tools. Economists should be familiar with the properties of all areas of contour maps even though the greater part of such areas are never used. Such knowledge minimizes the danger of faulty analysis. Secondly, a discussion of unused areas permits a more enlightened comparison of the two main types of contour map. Lastly, it can be shown that variations in these maps arise from different assumptions that are made about the redundant areas. These variations are a source of confusion. The extension of the analysis eliminates this source of trouble.

The thesis is divided into two parts. Part I deals with the extensions of existing indifference analysis. The opening chapter sets out in tabular form basic postulates and useful definitions for the description of two experiments upon which indifference analysis can be founded, and from which it can be extended. This chapter is laid out in outline form so as to set off all postulates, definitions and secondary assumptions in a manner convenient for ready reference. Chapter II is a preliminary discussion dealing with the extension of analysis for the neutral case of independence and for the two extreme cases of perfect complements and perfect substitutes. To deal with more general situations a definition of commodity relations is required. Various definitions are examined in Chapter III, and the most appropriate of these is considered in more detail in Chapter IV. With its aid the analysis of the more general cases of partially related goods is fully extended. In Chapter V, certain restrictions imposed on the analysis in Chapter I are removed for the purpose of considering types of indifference maps not treated, even in part, by existing analysis. Chapter VI is a digression on absolute necessities. These only arise when the fundamental experiments are formulated on terms which differ from those set out in Chapter I. The longest chapter of Part I is Chapter VII, in which are presented some applications of the extended analysis. The findings and conclusions of Part I are briefly summarized in Chapter VIII.

Part II deals with extensions of iso-product analysis. Its plan is similar to that of Part I. The preliminary chapter, Chapter IX, sets out in tabulated form postulates, definitions and secondary assumptions, along with a fundamental proposition on elasticities, required, and useful, for existing analysis and for its extension. Chapter X is a preliminary discussion dealing with the extension of the analysis for the neutral case of independence and for the two extreme cases of perfect complements and perfect substitutes. In Chapter XI a suitable definition of input relations is developed, and with its aid the analysis of more general isoquant maps is extended. Certain restrictions imposed on the analysis in Chapter IX are removed in Chapter XII. The consequence is the development of special isoquant maps not accounted for by existing analysis. Chapter XIII is a digressionary chapter which treats of the contour analysis of linear homogeneous production functions. Chapter XIV returns to the main lines of the analysis by presenting some applications of the extended analysis. This chapter is briefer than its counterpart in Part I since many of the ideas presented there can be applied here. The findings and conclusions of Part II are briefly summarized in Chapter XV.

At the end of the thesis appear the General Conclusions. Here an attempt is made to knit together the two separate parts of the thesis. The main conclusion agrees with Professor Norris's statement: "Logical extremes are often illuminating in the field of Economic Theory".

PART I

EXTENSIONS OF INDIFFERENCE ANALYSIS

CHAPTER I

ASSUMPTIONS AND DEFINITIONS BASIC TO PART I

Existing indifference analysis can be extended with much profit. The purpose of this preliminary chapter is to set out the following: first, postulates basic to two experiments on which classical existing indifference analysis[1] rests;[2] secondly, definitions useful in the description of the experiments, and in the extension of existing analysis; thirdly, description of the first of the experiments, which abstracts from economic considerations of price and income; and fourthly, description of the second experiment, which permits such economic considerations to operate. Under the last two heads certain secondary assumptions are set down. Some of these are useful for carrying out the described experiments and for tabulating the results; others are taken from existing analysis; still others are helpful for the extension of the analysis. In addition to these, under the third head, note is made of the geometrical method for tabulating the results of the experiment and for deriving from the results

---

1 The classical existing indifference analysis can be found in the following sources which constitute a partial list of useful references on the subject: J.R. Hicks, Value and Capital (2nd ed), Pt I, and pp55-7; R.G.D. Allen, Mathematical Analysis for Economists, pp124-6, 290, 438; K. Boulding, Economic Analysis (1941), ch 30; H. Schultz, The Theory and Measurement of Demand, Pt I; R.G.D. Allen and A.L. Bowley, Family Expenditure, ch 3.

2 The analysis need not rest on these experiments. It may be derived by the method of revealed preferences. Cf. P.A. Samuelson, "Consumption Theory in Terms of Revealed Preference", Economica, 1948, pp243-253.

the indifference maps associated with existing analysis. Under the fourth head, some preliminary remarks on prices and income are laid out along with a discussion of the price situations under which the analysis can be extended. Again, under the last head, note is made of the conditions under which the experiment described there conforms most closely to the real world. The scheme described here is given below in tabular form.

I - Postulates Basic to the Experiments

A. An individual has a given system of wants, at a given moment of time, for different amounts of a wide variety of goods.

B. An individual selects that collection of goods from among those alternatives available to him which he expects will most fully satisfy his wants.

> Comment: This selection is an observable fact, and provides the empirical evidence on which the case for indifference analysis rests.

II - Definitions

A. The term good, or commodity, applies qualitatively to amounts of any object which helps to satisfy any of the individual's given wants.

> Comment: When the individual prefers the zero amount of an object to some other positive quantity of it, that other quantity does not help to satisfy any want. Therefore, all quantities of an object inferior to the zero quantity can not be referred to as quantities of a good.

B. The saturation quantity of a good is that quantity which completely satisfies the want for the good when all other commodity amounts are fixed at given values.

Comment: The saturation quantity may vary for different fixed values of other commodity quantities.

C. A satiable good is one for which saturation occurs for some finite amount of the good.

Comment: It is with satiable goods that the analysis is concerned.

D. A non-satiable good is one for which saturation does not occur for any finite amount of the good.

E. Quantities of a satiable good greater than the saturation quantity are said to be a nuisance if such quantities are inferior in preference to the saturation quantity.

Comment: For a nuisance good all excess units have a negative marginal utility. This must mean either that it causes the individual some trouble and effort to get rid of the excess units (speculation and recourse to the market for resale purposes being ruled out), or that the physical nature of the units is such that it is not possible to overlook or ignore them.

F. Quantities of a satiable good greater than the saturation quantity are said not to be a nuisance if such quantities are on a level of indifference with the saturation quantity.

Comment: Such indifference means that excess units of the good all have a zero marginal utility. Abstracting from economic considerations such as speculation and resale such a possibility indicates either that the redundant units can be overlooked and ignored or that they can be disposed of without any effort on the part of the individual, disposal referring to riddance without profit.

III - Description of the First Experiment

A. An experiment based on the above postulates and definitions is now described. The significant elements in the experiment are as follows:

1. A hypothetical subject is asked, at a given moment of time, to state which of two alternative collections of goods presented to him he prefers if he can only have one of the collections.

2. Each collection of goods contains only quantities of objects to which the term good applies with reference to the given amounts in the collection.

   Comment: Zero amounts are not excluded from any collection; only those positive non-zero amounts of an object to which the qualitative definition of a good does not apply, when all other commodity amounts are given, are excluded. Also excluded, of course, are all amounts of all objects for which the subject has no wants.

3. All possible combinations of pairs of collections containing all possible permutations of the goods are presented to the subject.

   Comment: Every conceivable collection of goods is used in the experiment.

4. Out of each pair of collections the subject either selects one or, if he can not indicate a preference, indicates his indifference as between the collections.

5. The experiment is considered to take place at a given moment of time. It abstracts from such economic considerations as prices, income, speculation, resale, *etc.*, so that wants alone determine choice.

B. Four secondary assumptions, taken from existing analysis, are now set down. The first two are useful in conducting the experiment just described while the last two are useful in tabulating and analyzing the results of the experiment.

1. The order of presentation of the pairs of collections has no effect on the subject's choice.

   Comment: This assumption ensures that, under the appropriate conditions, the indifference curves close instead of spiraling. It is similar to the Paretian assumption that the order of consumption is indifferent and does not influence the pleasure which it occasions.[3]

2. The subject's choice behavior is consistent so that he does not indicate he prefers one collection to a second, the second to a third, and the third to the first.

3. The measurement of preference behavior is ordinal.

   Comment: The scale against which the subject's preferences are recorded is an arbitrary one so that there exist an infinite number of ways of tabulating the results of the experiment.

4. The variation from one collection of goods to more preferred collections is continuous as is also the variation from one batch of goods to other batches on the same level of indifference with it.

   Comment: The continuity assumption is adopted here for theoretical convenience.

---

3 Cf. H.T. Davis, The Theory of Econometrics, pp78-9.

C. The results of the experiment can be plotted in hyper space on an arbitrary scale.

1. An infinite number of hyper figures may result from the plotting as a consequence of the arbitrariness of the utility scale.

   Comment: Each hyper figure has one more dimension than the number of goods used in the experiment.

2. For purposes of analysis only one of the possible hyper figures need be considered, provided that only those properties are dealt with which are invariant for a monotonic transformation.

   Comment: The analysis of Part I is worked out using invariant properties only. It is left to the reader to check the invariance for himself.

3. The analysis of the selected hyper figure is simplified by considering various plane cross-sections of it. This reduces the analysis of preference behavior from a multi-dimensional to a two dimensional frame of reference. It is accomplished by taking a plane through any point on the hyper figure parallel to any plane containing two commodity axes, referred to here as X and Y. The resulting cross-section yields an indifference map for X and Y which is the subject matter of indifference analysis.

   Comment: The indifference map of existing analysis is similar to this map derived from the hyper figure if the axes used in fixing the cross-sectional plane are suitably restricted. The restriction required is that neither the X nor Y good have a perfect substitute or a perfect complement unless that perfect substitute or complement is X or Y.[4]

D. Existing analysis is only concerned with a certain so-called "relevant" section of the indifference map. The following assumptions are associated with this area:

1. At all points within this area the margional utilities of both X and Y are of the same sign so that each indifference curve has a negative slope[5] and is continuous.

2. The second derivative of each indifference curve is positive.[6]

3. All indifference curves appear in a "continuously variable order of ascending preference" so that both marginal utilities are positive.

E. Existing analysis can be extended by considering the utility surface[7] for which the indifference curves are the utility contours, and by making certain assumptions about that surface.

---

4 This restriction is relaxed in Chapter V below. It is not a severe restriction since commodities with perfect relations are not numerous.

5 The slope of an indifference curve is given by the ratio of the marginal utilities of the two goods, preceeded by a minus sign.

6 This assumption is deduced from the stability conditions required for equilibrium.

1. The utility surface rises monotonically to a maximum height for increases in X alone, or Y alone, thus implying that larger amounts of a good are always preferred to smaller amounts when both amounts are less than the saturation quantity.

> Comment: These maximum heights of the surface form ridge lines on the surface whose projections on to the indifference map trace out the loci of the saturation amounts for X or Y, when Y or X are fixed at different values. These loci are called saturation curves, or more briefly (but less correctly) ridge lines. Since along these saturation curves the marginal utility of one good is zero, the slopes of the indifference loci at their point of intersection with these curves become parallel to the axis of the saturated good.[8]
>
> These ridge lines play an important part in the analysis to follow.

2. For satiable goods the utility surface has only one set of ridge lines[9] thus implying that all amounts of a good greater than the saturation amount are either on the same indifference level with, or are inferior to, the saturation quantity.

---

7 The indifference curves derived from the given hyper figure represent the contours of just one utility surface since all other arbitrary measures of utility were eliminated when the given hyper figure was selected.

8 This is so because the slope of the indifference curves is given by the ratio of two marginal utilities. When one of these is zero the slope will either be zero or infinite. Thus the X (Y) saturation curve is also the locus of all X (Y) values for which the indifference curves turn parallel to the X (Y) axis.

9 This assumption eliminates Fisher's special case of two distinct maxima of satisfaction since two ridge lines, with a single point of intersection, determine only one locus of maximum satisfaction. More than one maxima can not be accounted for by the analysis. For an attempt to account for it cf. Irving Fisher, "Mathematical Investigations in the Theory of Value and Prices", Transactions of the Connecticut Academy of Arts and Sciences, v. IX (1892-95), p70.

Comment: If the latter condition exists the above assumption also implies that the smaller of any two amounts, both greater than the saturation amount, is always preferred to the larger amount.

3. If excess units of X and Y are not a nuisance, the utility surface flattens out beyond the ridge lines in the X and Y directions, and remains horizontal for all excess amounts of X or Y up to infinity.

Comment: The hyper preference figure extends indefinitely in the positive X and Y directions since no positive amount of the X and Y goods are excluded from the experiment.

4. In lieu of 3 above, it might be assumed that excess units of X and Y are nuisances. The utility surface then decreases monotonically beyond the ridge lines as X and Y continue to increase.

Comment: Under this assumption the hyper figure does not extend indefinitely in the positive X and Y directions. It is terminated at those X (Y) values which are so great that they are on the same indifference level as zero X (Y). Amounts greater than these are excluded from the experiment.[10]

The Utility surface associated with the XY cross-section of the hyper figure also terminates at the X or Y values where X or Y cease to be a good. Beyond these points the XY plane is "empty".[11]

---

10 This is an artificial exclusion, but it is a meaningful one. All those amounts of an object which are inferior to the zero amount do not meet a meaningful definition of an economic good. If the subject possesses none of the object and is asked to choose between two amounts of it, both amounts being inferior to the zero amount, he will not make that choice but rather will indicate that he prefers the zero amount. Thus, with respect to amounts of only one object he will never freely choose an amount which is inferior to the zero amount. He may, however, choose such an amount if paid to do so. However, the experiments described here are only concerned with choice behavior when all objects are objects of desire.

IV - Description of the Second Experiment

A. In this experiment certain economic considerations which were excluded from the first experiment are permitted to operate. Some preliminary remarks on these are set out here.

---

11 If total negative utility is not ruled out for each good individually but is ruled out for a collection of goods, then two possibilities arise.

First, if the origin of the indifference map denotes a point in hyper space associated with a collection of goods yielding some total utility, then the utility surface of the resulting indifference map falls below the XY plane until the utility yielded by the fixed quantities is cancelled. The surface then becomes parallel to the XY plane. Of course, if absolute negative total utility is not ruled out the surface would continue to fall indefinitely. Thus the surface lies both above and below the XY plane so that the plane is not "empty". If the total utility yielded by X and Y is measured as a deviation from the utility of the fixed quantities so that the zero X and zero Y point on the indifference map is associated with zero total utility, then some of the indifference curves will be designated by negatively numbered parameters. Such indifference curves indicate that the amounts of X and Y represented by points on them substract from the total utility of the fixed quantities.

Second, if the origin of the map denotes a point in hyper space associated with a collection of good yielding no utility, then the utility surface of the indifference map only falls until it meets the XY plane everywhere. For all greater X and Y quantities the surface coincides with the XY plane, and that part of the plane is then "empty". In this case one of the goods yields negative total utility but both together must yield positive total utility. If this latter condition is relaxed the surface falls below the XY plane. The definition of a good given here rules out negative total utility for any one good as well as for any collection of goods.

This restriction must be what Fisher had in mind when he stated: The utility surface "need not extend indefinitely over the plane.... [It] may approach vertical plane or cyclindrical asymptotes so that for some points in the plane there may be no surface vertically over or under....Economically it is impossible that the individual should consume quantities...indicated by the coordinates of such points. Those parts of the plane where such points are may be called 'empty'." ("Mathematical Investigations in the Theory of Value and Prices", <u>op. cit.</u>, p69) Schultz, <u>op. cit.</u>, p13, endorses this statement.

1. The price attached to any good has no effect on choice behavior.

   Comment: This is in general true, although there are a few goods, e.g., orchids and diamonds, where the price plays a part in shaping the choice of the individual for the commodity.

2. All units of the same good have the same price.

3. With real income constant,[12] the lower the price of a good the more units of it the individual will buy.

   Comment: As long as the marginal utility of a good is positive, its price will be positive. Only when its marginal utility is negative will its price be negative. The individual must be paid to take a unit of a good when that unit is a nuisance.[13]

B. All the significant elements in the description of the first experiment, except the fifth (section III - A.), apply to the description of the second experiment. In addition to these, the following elements are required for the latter description.

1. The experimentalist is some large central organization having complete control of all available supplies of every good.

2. This agency sets a fixed, but not necessarily identical, price on each good.

3. It gives to the subject a sum of money on the terms that he spend all of it on buying a collection of goods.

---

12 i.e., constant price level and money income

13 Cf. F.H. Knight, "Realism and Relevance in the Theory of Demand", Journal of Political Economy, 1944, p292; K. Boulding, Economic Analysis (rev. ed), p762.

<u>Comment:</u> By varying prices and the sum of money given to the subject, the experimentalist can confront the individual with an infinitely large number of choice situations.

From among all the collection of goods which the subject can buy at a given set of prices and with a given sum of money, there will be only one collection which is preferred to all others. This collection will be the one chosen if choice is based on the same considerations of wants as ruled in the first experiment.

C. In addition to the four assumptions set down under section III - B the following assumptions are required to render similar the results of the two experiments.

1. The experimental agency deals with all other individuals and institutions with which the subject may be in contact on the same terms as with the subject. Under this, all buyers must be indifferent as from where they buy.

   <u>Comment:</u> If the fact that the agency will deal with others on the same terms as with the subject is known, the subject can not gain anything by buying from the experimentalist and selling to others on different terms.

2. when the agency announces a set of prices for its goods, the subject expects these prices to remain so fixed for as long in the future as he makes plans.

   <u>Comment:</u> The individual is not induced by a low price to buy more goods than he needs because he expects prices to rise later.

3. Production, as well as speculation, is ruled out.[14]

   <u>Comment:</u> The subject does not buy a collection

---

14 Production is not dealt with until Part II of the thesis.

of goods for the purpose of transforming them into different goods.

D. The results of the second experiment plotted in hyper space yield the same type of hyper figures as obtained from the first experiment. Cross-sections of any one of these hyper figures, taken in the appropriate manner, produce indifference maps similar to those derived by means of the first experiment.

Comment: To obtain an indifference map reflecting preference behavior of an individual in a realistic situation, the point on the hyper figure, selected for the cutting plane to pass through, must represent the equilibrium purchases of the individual, at a given moment, for a relevant price and income situation. That is, the point in the hyper figure must represent the equilibrium purchases of the individual in the real world at the given moment. The plane passed through that point parallel to the XY plane then yields the relevant indifference map for studying the preference behavior of the individual for X and Y, under varying income and price situations, when the prices and quantities of all other goods are fixed at values given by the equilibrium point.

E. The analysis can be extended under the assumptions of section III - E, but the price situations under which such extensions can proceed must be considered. To do so fix the sum of money spent on X and Y, all commodity quantities but X and Y, and all prices but the price of X or Y. Obviously, granted that the subject exercises free choice, the more X or Y the experimentalist wishes him to take, the lower he must set the price of X or Y.

1. If the variable price is set equal to zero, and if excess units are not a nuisance, the experimentalist can get the subject to take an infinite amount of the good.

2. If excess units are a nuisance, the agency can not get the subject to take more than the saturation amount by setting a zero price. Once saturated with X or Y, the individual will not choose larger amounts unless the experimentalist pays him to take more by setting a negative price.[15]

> Comment: The negative price situation may appear more nearly relevant if an example is considered. Suppose the experimental agency has a glut of X which it seeks to dispose of in a socially desireable manner. It offers X free to all individuals, who all find excess units a nuisance. After they have taken all the X they want when it is free, assume that the agency still has left on hand a large amount of the good. It it wants individuals to take more of the good it must set a negative price on X. This negative price must be high enough, absolutely, to induce sufficient takers to come forward and clear away the desired amount. These takers, it should be noted, can not sell or even give away their excess amounts of X because everyone has as much of it as he wants and would therefore have to be paid to take more.
>
> Thus the free choice of redundant quantities is possible if negative prices are possible when required. No compulsion is necessary. The individual need only respond to the incentive of negative prices when it is in his interest to do so. The situation may be improbable,[16] but its occurence does not rule out, or is not incompatible with, free choice.[17]

---

15 "The subject would then judge whether he would be willing to accept so many items of a disadvantage [the excess units] in order to possess so many items that are desirable [e.g., money]. The supposition is that as the amount of advantage is increased he would be willing to accept a greater amount of corresponding disadvantage." (L.L. Thurstone, "The Indifference Function", Journal of Social Psychology, v.2, 1931, p145)

16 However, "undesired surpluses do accumulate, even of commodities of which a moderate amount is of some use, and people pay to have them removed. But the cost of removal is normally so low that the outer reaches of the supersatiation curves would seem to lie beyond the range of probably experience." (J.M. Clark, "Realism and Relevance in the Theory of Demand", Journal of Political Economy, 1946, p353).

F. The conditions under which the second experiment conforms most closely to the real world are noted below. Some of these require that certain conditions of the experiment be relaxed; others are simply repetitions of conditions already laid down.

1. The subject is confronted with an infinite number of distributing agencies instead of just one. In his environment there are now a large number of sellers as well as buyers.
2. The individual, as buyer, is indifferent as between the various suppliers of the goods he purchases.
3. The price of each good, whatever it may be, is the same for all buyers and sellers, and is independent of the amounts bought by any one individual.

---

17 That is, the choosing of excess amounts, in the experiment, does not involve any notion of force or compulsion. All that is necessary for free choice to operate is sufficient compensation and the absence of critical points at the ridge lines. This differs from the conclusion of some other writers. In his discussion of this point, Professor Knight states: "We assume consumption compulsory, if necessary, ignoring the possibility that later increments of A [the excess good] might be simply ignored or might be disposed of in some other way." ("Realism and Relevance in the Theory of Demand", op.cit., p292) He further states, "The ascending part of the [indifference] curve is unrealistic under ordinary conditions where the subject has freedom of choice." (loc.cit.)

Professor R.L. Bishop also argues that "under the more natural assumption that no consumption is *forced* beyond satiation, the indifference curves would simply become parallel to the appropriate axis, instead of bending away from it." ("Professor Knight and the Theory of Demand", Journal of Political Economy, 1946, p144n8. Italics mine.)

Comment: The condition permitting these prices to range in value over the entire scale of real numbers is retained.

4. The individual provides his own stock of money, i.e., his income, for trading purposes.

On the basis of the second of the two experiments described, and under the conditions noted under section IV - F, a hyper preference figure can be constructed for each individual showing his preference for all goods at any given moment of time. This figure, it will be recalled, is restricted to only those commodity quantities which can be classed under the term good, as that is defined here. On this figure any point may be selected and a plane taken through it parallel to a plane containing any two commodity axes. Any two axes may be chosen for this purpose, the only restriction being that the commodities represented by these axes must not be perfectly related to any other commodities except each other.[18] The cross-sectional plane, taken as described, yields an indifference pattern with which two dimensional indifference analysis is concerned. By changing the axes of reference various cross-sectional planes yielding different preference maps are obtained. The chapters of Part I which follow are taken up with the extension of the analysis of these various maps.

---

18 Cf. comment on section III - C, p9.

CHAPTER II

EXTENSION OF ANALYSIS FOR INDEPENDENT GOODS, PERFECT COMPLEMENTS AND PERFECT SUBSTITUTES

Three special types of cross-sectional planes, discussed in the last chapter, are considered in this chapter. These correspond to three critical commodity relationships. The first type of cross-section yields an indifference map for independent goods, the second, an indifference map for perfect complements, and the third, an indifference map for perfect substitutes. The determination of each of these types imposes a restriction[1] on the selection of the (two) commodity axes, the nature of the restriction depending on the particular definition of commodity relationships adopted. Precise definitions are def~~in~~erred until Chapter III, but working definitions are presented below.

I - Indifference Maps for Independent Goods.

Consider a selection of the X and Y axes such that any XY plane cross-section of the hyper-figure yields indifference curves with slopes which can be expressed in the form $f(x)/g(y)$.[2] This implies that the equation

---

1 This is in addition to the restriction imposed in the last chapter preventing the two commodities from being perfectly related to other commodities but each other.

2 J.R. Hicks and R.G.D. Allen, "A Reconsideration of the Theory of Value", Pt I, Economica, 1934, pp74-5, take the presence of this property to define independence.

of the utility surface, associated with the given curves, is of the form $u = f(x) + g(y)$,[3] where u represents total utility. Since X and Y make separate and independent contributions to total utility in this equation, the two commodities may be classed as independent goods. Any plane cross-section determined by such a selection of the commodity axes then gives an indifference map for independent goods.

On such a map, a saturation point for X occurs for that value of X for which $\partial u/\partial x = f'(x) = 0$. Since this expression is independent of Y, the saturation curve for X is a line parallel to the Y axis. In other words, all saturation points for X have the same X value.[4] Similarly, all Y saturation points have the same Y value.

Under the assumption that excess amounts of X and Y are not a nuisance, the indifference map is of the type

---

3 Cf. R.G.D. Allen, "The Nature of Indifference Curves", Review of Economic Studies, v.I, No.2, Feb. 1934, pp110-112. Professor Samuelson, Foundations of Economic Analysis, pp174-176, proves that, in general, there will be no utility function that can be written in this form. However, it is assumed here that one does exist only for the purpose of getting a tentative definition of independence. To this there can be no objection.

4."Under the conditions of a two commodity system...[the saturation points for X]...should ordinarily have the same X value...if X and Y are 'independent', i.e., noncomplementary but different goods." (Knight, "Realism and Relevance in the Theory of Demand", op.cit., p292) illustrated in Figure 1. The lines AN and BL are, respectively, the Y and X saturation curves. The area AOBM is the effective region because it is the area of

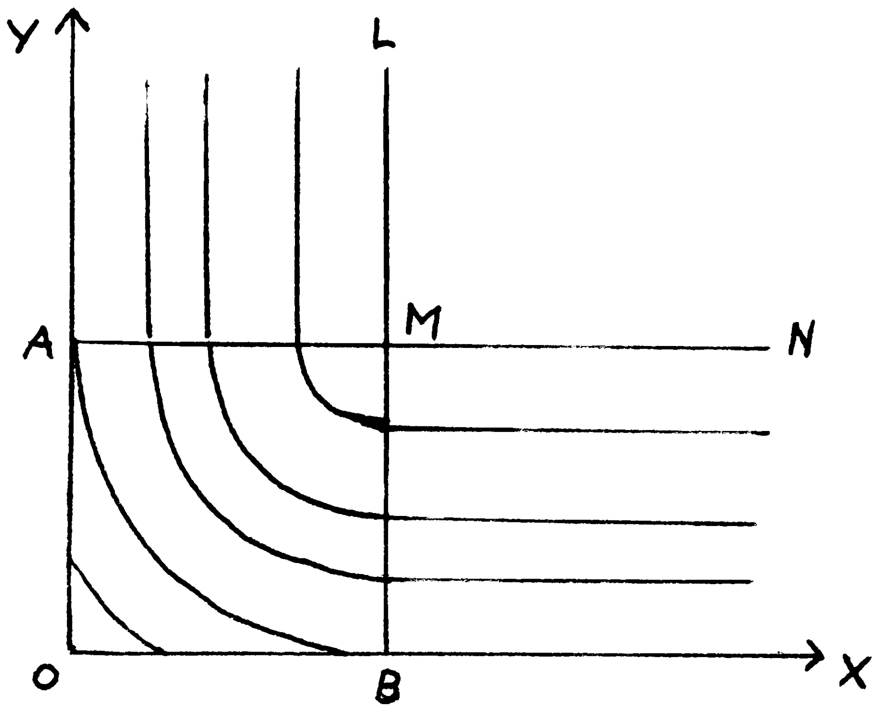


Fig. 1. - Independent Goods. Excess units no nuisance.

significance in realistic demand theory, and for this reason is usually the only section of the map given in the textbooks. The point M where the ridge lines intersect is a point of maximum satisfaction or of "bliss". But it is not the only point of "bliss" because any point lying within the area bounded by LM and MN is on the same indifference level as M. This is because the marginal utility of X and Y within the area is zero. Therefore, the locus of maximum satisfaction, or maximum indifference, is a plane bounded by LMN and stretching towards infinity in north and east directions.

Suppose now that excess X is a nuisance but that excess Y is not. The indifference map is then of the type illustrated in Figure 2. The indifference curves have a positive slope in the area BCNM because the marginal utility of X is negative while the marginal utility of Y is positive. The point M is again a point of maximum

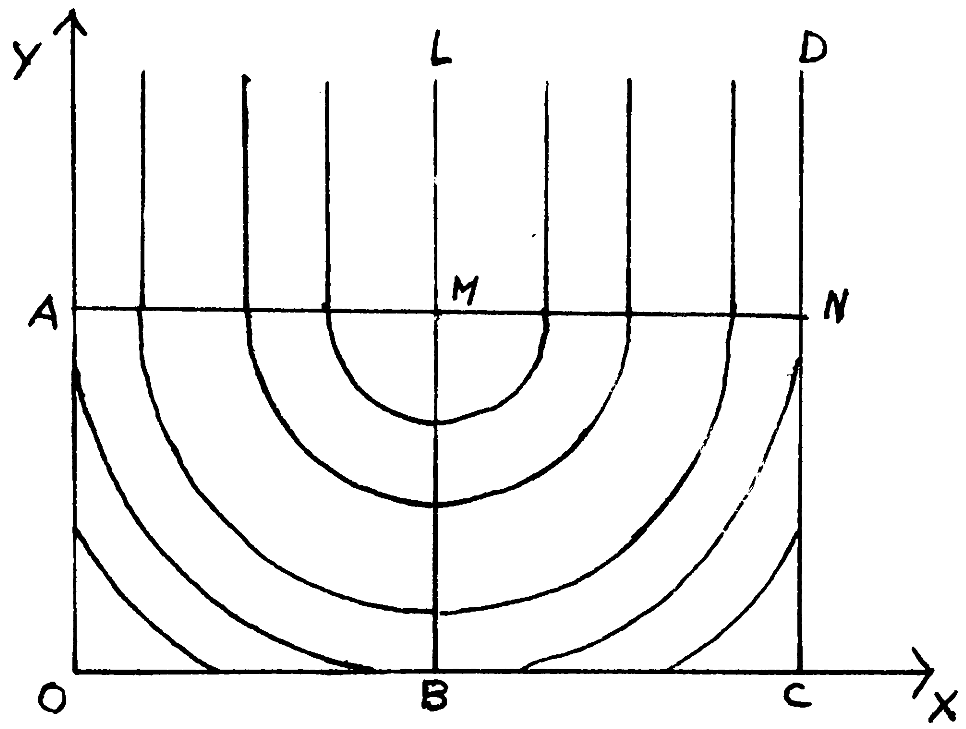


Fig. 2. - Independent Goods. Excess X a nuisance.

satisfaction. Any addition of X at this point results in a loss of satisfaction, but additions of Y have no effect on utility. Therefore, the locus of maximum indifference is the line parallel to the Y axis lying between M and infinity.

The map is bounded on the right by the straight line CND drawn perpendicular to the X axis. It marks the boundary of the map because it is the locus of all X values which individually are on the same level of indifference as zero X.[5] As a result of the independent relationship there is only one such X value, OC. The

5 Beyond this boundary the XY cross-sectional plane is "empty" since the hyper-figure does not exist for X values lying in the "empty" area when all other commodity quantities are fixed at the values associated with the origin of the map in hyper space.

points O and C may therefore be considered as points of zero total utility,[6] both being on the same level of indifference. In general, for any given Y value there are two X values which form two distinct combinations both on the same level of indifference.[7]

The possibility that both excess X and excess Y are nuisances is represented by a family of closed curves suchas those in Figure 3. The point M is <u>the</u>

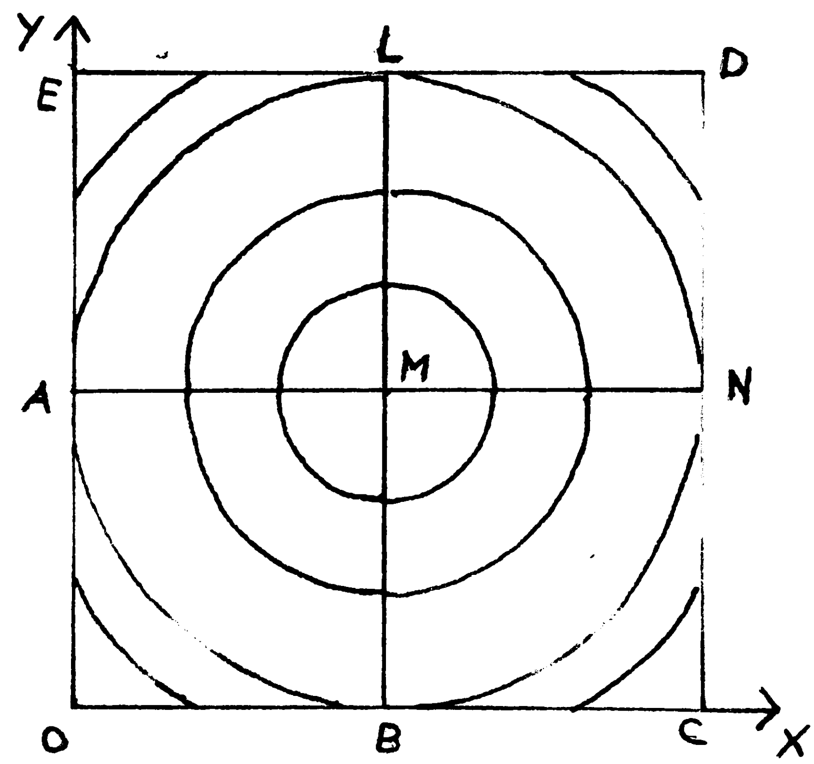


Fig. 3. - Independent Goods. Excess units nuisances.

point of "bliss" since all other points are inferior to it. It is the point at which the subject would arrange his purchases if X and Y cost nothing, or if his income

---

6 i.e., only if the total utility associated with the origin is taken to be zero for purposes of calibrating the indifference curves.

7 "There will be two equivalent combinations of each quantity of Y with different quantities of X, one corresponding to a deficiency, the other to a superfluity, of X". (Knight, "Realism and Relevance in the Theory of Demand", <u>op</u>.<u>cit</u>.,p292)

permitted.

In the area BMNC the individual must be paid in terms of Y to take more of X (real income kept constant), while in the area AELM he must be paid in terms of X to take more Y. In the region MLDN, he must be compensated for suffering more Y by being made to suffer less X. Or, to induce him to take more X he must be relieved of some Y. So in this area, movement along an indifference curve involves the substitution of one discommodity for another. Further, both marginal utilities are negative here while in the effective region, OAMB, both are positive. In the other two areas of the map the marginal utilities have opposite signs. It is the signs of the two marginal utilities which determine the sign of the indifference slope in the various regions.[8]

The line CD is the locus of all points where the total utility of X is zero, the amount of X being so great that it is on the same level of indifference as zero X. The line ED is the similar locus for Y. Therefore, these lines, along with the axes, mark the limits of the map. At the points of intersection of the boundaries, the four points O, E, D and C are all on the same level of indifference and may be considered to be points of zero total utility. Under the conditions of

8 For a very clear and vivid description of closed indifference curves cf. A.P. Lerner, "The Diagramatical Representation of Demand Conditions in International Trade", Economica,1934, pp319-334.

the experiment, these boundaries enclose all possible X and Y amounts at which the hyper-figure exists, for the given fixed quantities of all other commodities.

II - Indifference Maps for Perfect Complements

Consider now a selection of the X and Y axes such that the effective region of an XY indifference map is reduced to a straight line. The two commodities measured along these axes can then be defined as perfect complements since the effective range of substitution between them is zero. Unless quantities of the two goods form a given fixed ratio some amount of one of the commodities will be redundant. Thus, in Figure 4, the slope of the straight line OM represents the fixed ratio

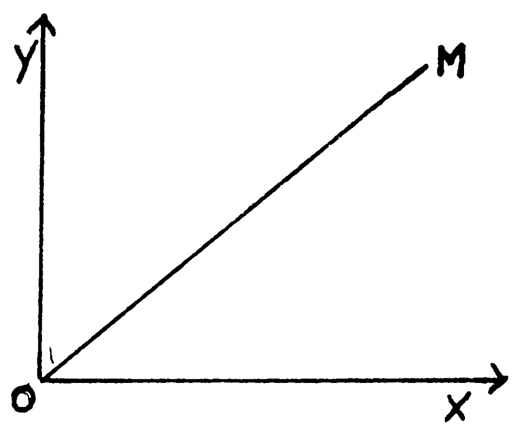


Fig.4

in which quantities of Y may be combined with quantities of X without amounts of either being in excess. However, anywhere to the right of OM X is in excess and anywhere to the left of it Y is excessive. Therefore, OM is the saturation curve for X when variations in X alone are being considered, but it is the saturation curve for Y when attention is centered on variations for it. In increasing amounts of X and Y, in a fixed proportion, along the line OM, more and more preferred positions are reached until the point M is attained. Since the goods are satiable ones, the point M is a position of maximum preference.

Existing analysis usually assumes that excess amounts of perfect complements are not a nuisance, the assumption being implicit in the form of the indifference curves. Under this assumption the indifference curves are parallel to the axes outside of the effective region (in this case a straight line) for all excess amounts so that the curve system may be drawn as in Figure 5, ~~overleaf~~. The locus of maximum indifference is a plane bounded by LM and MN on the west and south and by infinity on the north and east.[9] The only finite boundaries of the map are the axes which

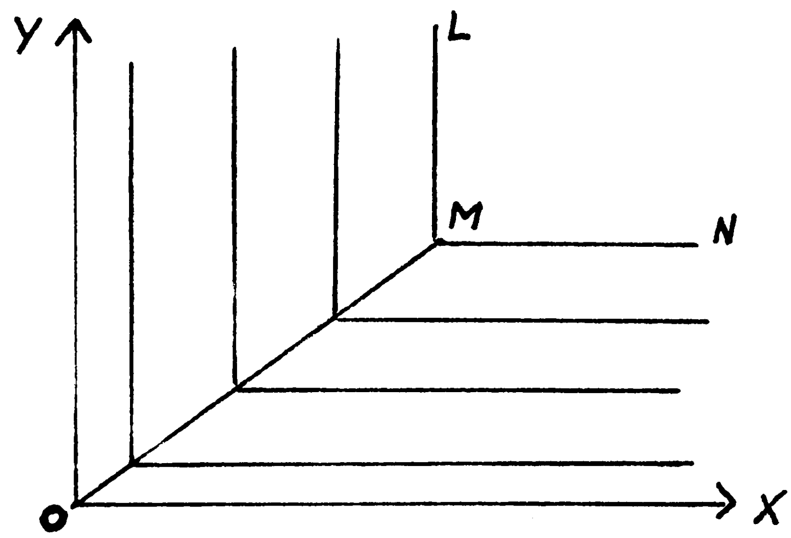


Fig. 5 - Perfect Complements. Excess Units no nuisance.

here constitute an indifference curve of zero total utility since all points on it are on the same level of indifference as the origin.

Now if excess units of one of a pair of perfect complements are a nuisance then excess units of the other commodity must necessarily also be a nuisance. Perfect complements, such as a pair of shoes, can be considered as

---

9 As a consequence of this locus being a plane, note what happens to the ridge lines. They become separated at the point M with LM being a branch of the X ridge line and MN a branch of the Y ridge line. This behavior of the ridge lines under the above assumptions has some bearing on later

a single commodity and if excess quantities of a part of that commodity can become a nuisance redundant amounts of all other parts of it would in all probability also be undesirable. This then rules out the possibility of excess supplies of one complement being a nuisance while excess supplies of the other are not. The only other situation to consider is the possibility that excess supplies of both X and Y are undesirable.

In this case the indifference curves outside of the effective region turn away from the axes and close around some point M - the point preferred to all others. This

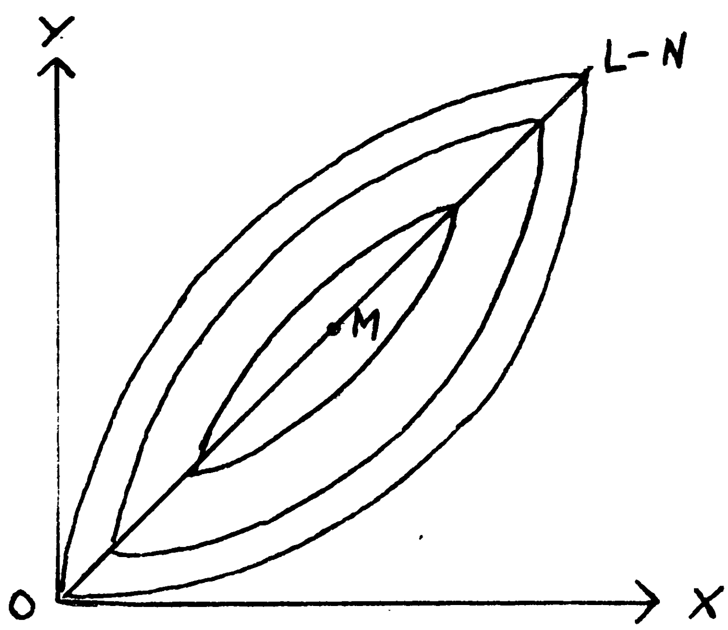


Fig. 6 - Perfect Complements. Excess units nuisances.

indifference pattern is illustrated in Figure 6. The slope of each curve to the left of the line OML-N varies from infinity where it cuts OM to zero where it cuts ML-N. To the right of OML-N the slope varies from zero on OM to infinity on ML-N. Thus, the line OML-N is the locus of all points where the indifference curves are parallel to both axes. As a result the slopes of the indifference curves are everywhere positively sloping.

Along OM and up to the point M larger amounts of X and Y are preferred to smaller amounts. Past M, along ML-N, smaller amounts are preferred to larger amounts. The outer curve in Figure 6 passing through the origin and the point L-N is an indifference curve marking the boundary of the map since it is the locus of all points which are on the same level of indifference as the origin. The hyper figure does not exist for quantities of X and Y lying outside this boundary when all other quantities are fixed at the values associated with the origin of the map in hyper space. Therefore, the XY cross-section yields isoquants only within the area indicated in Figure 6. The rest of the plane is "empty".[10]

III - Indifference Maps for Perfect Substitutes

The next selection of the X and Y axes considered is a very special one indeed. The two commodities measured along these axes must be two distinctly different goods yet so nearly identical that a unit of one is a perfect

---

10 The writer is not familiar with any source where the indifference map for perfect complements is drawn as in Figure 6. Yet, as will be argued later, this may be the most general indifference map for this class of goods. If this is so, Figure 6 should then replace Figure 5 in the economic texts as the standard illustration for perfect complements.

substitute for a unit of the other. If two goods can be found, a plane cross-section of the hyper figure parallel to the plane containing their axes yields a family of straight lines such as those in Figure 7. If neither good becomes a nuisance, then all points beyond some indifference line AB yield the same total utility because the goods are satiable ones. Beyond AB quantities of X and Y are in excess, and the locus of maximum indifference is a plane bounded by AB, the axes and infinity on the north and east.

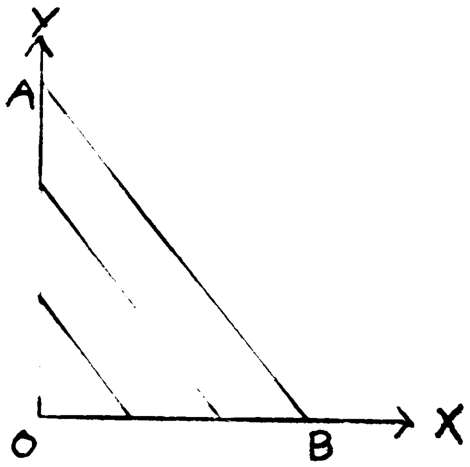


Fig. 7

Now, as in the case of perfect complements, it is impossible for excess units of one perfect substitute to be a nuisance while excess units of the other are not. This follows from the fact that perfect substitutes can be considered as one and the same commodity.[11] Thus several units of X and Y could be put into a large bag, mixed up, and then withdrawn, one by one, from the bag. After the saturation point is reached, if a unit of X (all units having previously been marked for identification purposes) is withdrawn and found to be a nuisance, then when the next Y unit is withdrawn it too will be found to be a nuisance.[12]

---

11 "....if two commodities are perfect substitutes for one another, they are for all practical purposes the same commodity, and there is no point in distinguishing between them." G. Stigler, The Theory of Prices, (New York 1947), p72.

The only other instance in which perfect substitutes are possible is where excess units of both goods are nuisances. The cross-sectional planes for such goods then

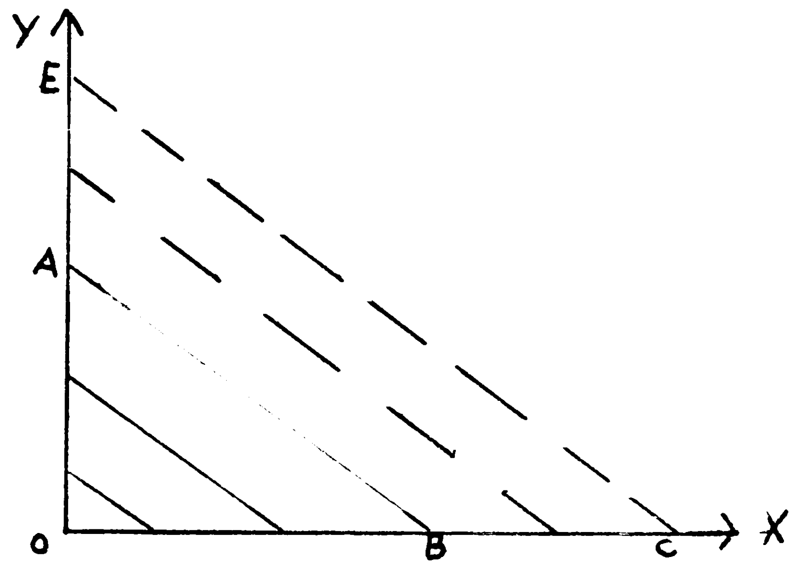


Fig. 8. - Perfect Substitutes. Excess units nuisances.

produce the type of preference pattern illustrated in Figure 8.[13] The line AB is now the highest indifference level, while EC is a boundary of the map since it is the locus of all points onthe same level of indifference as the origin.

---

12 Professor Clark, "Realism and Relevance in the Theory of Demand", op.cit., pp252-3, apparently errs when he draws an indifference pattern for perfect substitutes under the assumption that excess units of one perfect substitute only is a nuisance.

13 This figure differs from the one drawn by Professor Clark to show the same thing. Beyond AB his indifference curves cease to be linear. See his figure 2, ibid., pp252-3.

CHAPTER III

DEFINITIONS OF COMMODITY RELATIONS

Most pairs of commodities are not perfect substitutes or perfect complements; nor are they independent in the sense in which this term is used in Chapter II. Most pairs of goods, on a common-sense basis, are partial substitutes or partial complements. This means that to obtain a typical cross-section of the hyper-figure, axes of reference must be selected to represent goods which are partial substitutes or complements. But such a selection cannot be made until the significance of the relationship for the derived preference pattern has been determined. This is a matter of definition. Therefore, in order to obtain a typical cross-section it is necessary to frame a definition of commodity relations in terms of some property of the indifference map, or associated utility surface, which alters in a calculable manner for different selections of the commodity axes. In this connection some existing definitions are considered in this chapter.

The classical definition of commodity relations hinges on the sign of the cross derivative of the utility function.[1] But this is completely dependent on the notion of utility as a cardinal measurement. Two goods will be complements or substitutes according to the particular arbitrary measure of utility chosen.[2] Unless the measurability of utility is

1 Cf. F.Y. Edgeworth, Papers Relating to Political Economy, Vol. I, p117n1.

2 Hicks and Allen, "A Reconsideration of the Theory of Value", op. cit., pp59-60 and p196.

assumed, neither the form nor sign of this cross derivative is determinate.[3] Since merely ordinal measurement is assumed here, the classical definition can not apply.

It was for this very reason that the neo-classical definition was developed. However, this definition involves more than a formal change from the classical definition, required by independence from measurable utility.[4] This definition involves a radical change which altogether removes the problem of related goods from two dimensional analysis. On the basis of the definition "if a consumer is dividing his income between purchases of two goods only ... there can not be anything else but a substitution relation between the two goods."[5] Complementarity does not appear until at least three goods are considered.[6]

---

3 Cf. Schultz, The Theory and Measurement of Demand, pp22-3; O. Lange, "Determinateness of the Utility Function", Review of Economic Studies, Vol. 1 (1934), p221; and H. Bernardelli, "The End of the Marginal Utility Theory?", Economica,1938,p194.
The one exception is the case of perfect complements. This is because the effective region is a straight line representing both ridge lines, and along a ridge line the sign of the cross derivative is determinate.

4 Cf. Schultz, op. cit., pp619-20

5 J.R. Hicks, Value and Capital, p46. All references to this work are to the 2nd ed.

6 "Complementarity, in the strict sense in which we shall define it, is not a possible property of two goods; it only has sense when the goods in question are at least three." Hicks and Allen, "A Reconsideration of the Theory of Value", op. cit.,p69. See also p202 .
This is because complementarity is defined here with respect to a constant level of preference. Thus, on the indifference map for two commodities, the relationship between the goods, defined in terms of what happens along a given indifference curve, can only be a substitution one within the effective region. If the indifference curves are closed

For analytical purposes a definition of related goods that holds when there are only two goods is required, but it is exactly in this case that the neo-classical definition breaks down. A definition which treats two goods as substitutes when they appear on a common sense basis to be complements is not very helpful.[7] Therefore, the sign of the slope of the individual indifference curves is not a suitable property for use in defining commodity relations. Neither is the curvature of the curves a suitable property.[8] However, there are available for this purpose two

---

curves such as those in Figure 3, p24, then the goods will be complements, on the same definition, in those areas where the curves have a positive slope.

In travelling completely around one of the closed curves in Figure 3, the relationship between the goods varies from substitution, in the effective region, first to complementarity and then to substitution, in the area where both are discommodities. Then it changes to complementarity again and finally to substitution back in the effective region. In other words, the relationship varies with the sign of the slope of the indifference curves, being one of substitution when this is negative and one of complementarity when it is positive.

This follows from the fundamental equation: $\partial y/\partial P_x$ equals income effect plus substitution effect, where P_x is the price of X. Since the income effect is zero for movement along an indifference curve, the left hand term of the equation is negative when the indifference slope is negative and positive when it is positive. Therefore, the sign of the substitution term varies accordingly, and the sign of this term defines the relationship.

7 "The elementary fact of complementarity can...be studied in the case of two commodities only, and entirely without the interference of money.... It seems reasonable to demand that a realistic theory of value should also fit this elementary relationship between two goods." (H. Bernadelli, "A Note on the Determinateness of the Utility Function", Review of Economic Studies, 1934-35, II, 72)

For criticism along the same lines cf. R.L. Bishop, "Consumer's Surplus and Cardinal Utility", Quarterly Journal of Economics, May 1943, p438n6.

8 Cf. Hicks, Value and Capital, p42.

properties of the indifference map, independent of measurable utility, which are suitable.

One of these properties is the slope of the ridge lines. Milton Friedman uses the sign of these slopes as a criterion for complementarity. He suggests that "two commodities be defined as completing, competing or independent according as the saturation quantity of each commodity increases, decreases, or remains the same when the quantity possessed of the other commodity increases."[9] Thus, where the ridge lines have positive slopes the goods are complements, where they have negative slopes, substitutes. The intermediate case of zero and infinite slopes defines independence.

There is a simple connection between this ridge line definition and the classical definition. The equation of the X saturation curve is given by $u_x/u_y = f(x,y) = 0$, where u_x and u_y are the marginal utilities of X and Y. The slope of this curve is equal to the expression $\frac{u_y u_{xx} - u_x u_{xy}}{u_x u_{yy} - u_y u_{xy}}$.[10]

This reduces to $-u_{xx}/u_{xy}$ since $u_x = 0$ along the X saturation curve. Now u_{xx} is not, in general, invariant under a monotonic transformation of the utility function, except on

9 From an unpublished paper by Friedman. Quoted by Henry Schultz, op. cit., pp615-6.

10 Cf. Schultz, op. cit., p616.

the ridge line of X where it must always be negative for a maximum value of X.[11] Therefore, if the slope of the ridge line is positive, u_{xy} must be positive for points along the ridge line, and if the slope is negative, then u_{xy} must be negative.[12] If the slope is infinite, then u_{xy} must be zero. So if the goods are complements or substitutes by the classical definition at the ridge lines then they will also be complements or substitutes by the Friedman criterion.

Before evaluating the usefulness and validity of the Friedman definition we will consider another definition based upon the second property of the indifference curves referred to above. This property depends on the variation of the marginal rate of substitution, or indifference slope, when quantities of one good only are varied, and quantities of the other good are held at a fixed value. Johnson has suggested that commodity relations be defined in terms of this variation. The normal form of this variation is illustrated in Figure 9A,[13] where Q and R are two points

11 $U_{xx} = F'(u)u_{xx} + F''(u)(u_x)^2 = F'(u)u_{xx}$ when $u_x = 0$. Now since $F'(u) > 0$, and $u_{xx} < 0$, U_{xx} is negative definite for any monotonic transformation.

12 "The only case in which $\partial u/\partial x \partial y$ is definite in sign is when ϕ_x or $\phi_y = 0$. [u_x or $u_y = 0$]...This possibility does not concern us in our considerations of the individual under market conditions." (Hicks and Allen, op. cit., p197n2)

"Only on these satiation loci...can the unique or invariant independence, complementarity, or substitution of commodities — in the utility sense — be established." (Bishop, "Professor Knight and the Theory of Demand", op. cit., p150n16).

13 Figure 9 is reproduced from W. E. Johnson, "The Pure Theory of Utility Curves", Economic Journal, 1913, p493. For his treatment of commodity relations see pp493-99.

on the same indifference curve while P is a point on an adjacent and less preferred curve. The sloping lines through Q, R and P are the tangents to the indifference curves at those points. In the normal case the slope at

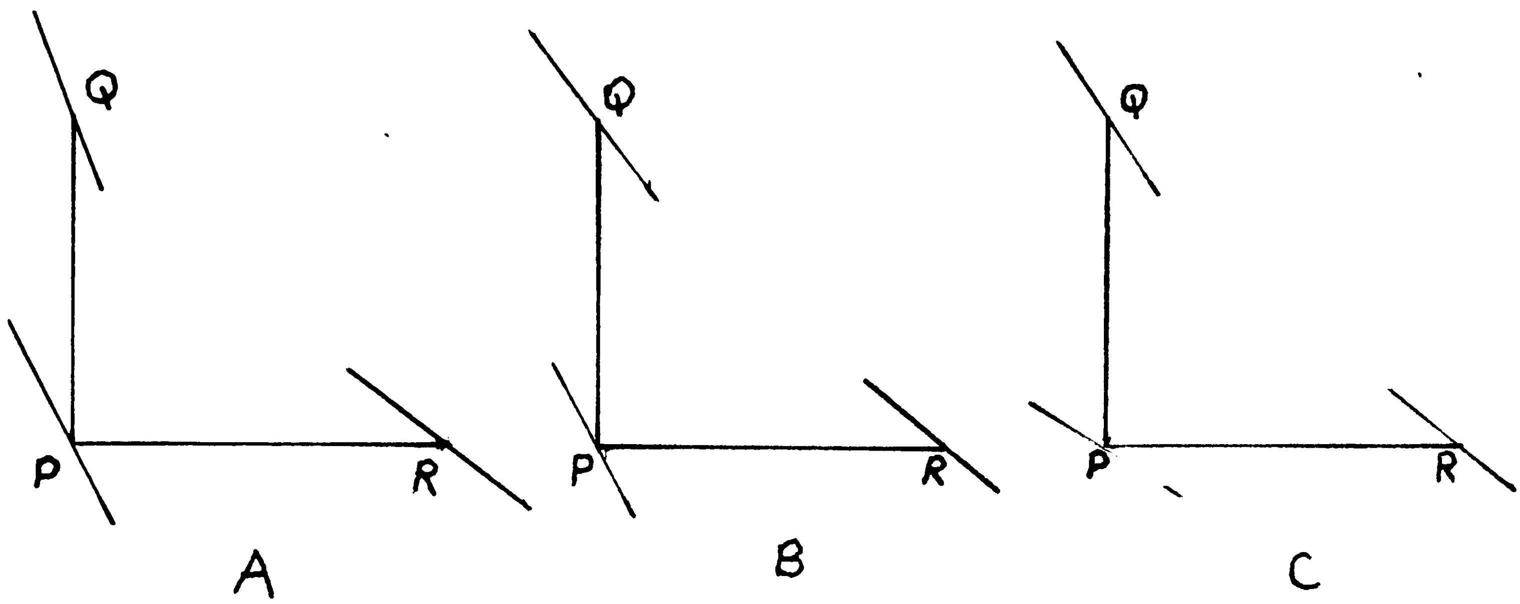


Fig. 9.- Variations in indifference slopes illustrated.

P lies in between the slopes at Q and R. So at P, for increase in Y, X fixed, the slopes of the indifference curves become more negative, while for Y constant and increases in X the slopes become less negative.[14] Johnson defines this case as complementarity.

In some cases this property of the indifference curves does not hold but breaks down, and the situation illustrated in Figure 9B or C arises. In Figure 9B the slope at P is steeper than the slopes at Q and R, while in 9C it is flatter than at Q and R. These are the only two possible exceptions

---

14 The mathematical statement of this property, given by Schultz, op. cit., p19, is $\delta X\ dy/dx > 0$, $\delta Y\ dy/dx < 0$, where dy/dx is the indifference slope, δX the variation in that slope along a line parallel to the X axis, and δY the variation in slope along a line parallel to the Y axis.

because it is impossible for the slope at P to be both steeper than the slope at Q and flatter than the one at R at the same time, because of the convexity of the indifference curves. When either of these exceptions occur Johnson defines the goods as substitutes.

However, Johnson did not indicate whether to define two goods as competing everywhere when the situation in Figure 9B or C occurs, or only within the range for which that situation holds. But Allen,[15] in refining Johnson's treatment, defines two goods as complementary if they are "normally related", as in Figure 9A, at all points, and competitive if they are "non-normal", as in Figure 9B or C, at some (but not necessarily at all) points. On this basis, the two goods are considered to have the same relationship at all points within the effective area.

The connection between the Johnson-Allen and Friedman definitions of commodity relations can be shown graphically.[16] Consider first the case where the ridge lines are parallel to the axes. Any straight line drawn anywhere within the effective region parallel to the X axis cuts the X ridge line at a point where it is a tangent to some indifference curve. At this point the indifference slope is zero. But to the left of this point the straight line cuts indifference curves at

---

15 R.G.D. Allen, "A Comparison between Different Definitions of Complementary and Competitive Goods", Econometrica, Vol. II, (1934), pp168-75.

16 The connection is shown mathematically by Schultz, op. cit., p617.

points where their slopes are less than zero, and the further to the left such points of intersection lie the steeper the negative indifference slopes.[17] So with parallel ridge lines, along any line parallel to the X axis the indifference slope at any point is flatter the nearer the point is to the X ridge line. While along any line parallel to the Y axis, the indifference slope at any point is steeper the nearer the point is to the Y ridge line. Therefore, although the goods are classed as independent on the Friedman definition, they are defined as complementary by the Johnson-Allen test.

Where the ridge lines are positively sloping at all points, the goods are still "normally related" at every point within the effective region so that they are classed as

---

17 This seems to be intuitively obvious from the convexity of the curves. The gradient at each point on the indifference map must be directed towards the point of intersection of the ridge lines and it must also be normal to the indifference curve passing through that point. Therefore, with respect to two points lying on a horizontal line, the one on the right must have a more steeply, or more northerly, directed gradient than the other. So the indifference slope at that point must be greater (less negative) than the indifference slope at the other point. This applies to all points lying on the same horizontal line. The further the point is away from the Y axis the flatter the indifference slope passing through it.

This is a very sketchy proof but it may be supported by quoting Schultz. "...the Johnson-Allen test gives an unequivocal determination of the relationship existing between the two commodities only in the neighborhood of the boundaries of the effective region....In regions removed from the boundaries, we should expect the...[normal case]...to be satisfied, no matter whether the goods are completing or competing...." (*op. cit.*, p610) Also, "it is almost certain,that if no 'non-normal' positions are found on the boundaries of the effective region, none will exist elsewhere...." (*ibid.*, p617) complements on both the Friedman and Johnson-Allen definitions.[18]

The "non-normal" situation described by Johnson and Allen only arises when the ridge lines are negatively sloping.[19] Thus, in Figure 10, the goods are "normally related" along any line parallel to the X axis lying within the area ORMB and along any line parallel to the Y axis lying in the area OAMS. But in the area SMB any such vertical line cuts indifference curves whose slopes become flatter as Y is increased. Likewise, in the area AMR any horizontal line cuts indifference curves whose slopes become steeper as X is increased. Therefore, it is evident that the regions AMR and SMB contain all the "non-normal" points.[20]

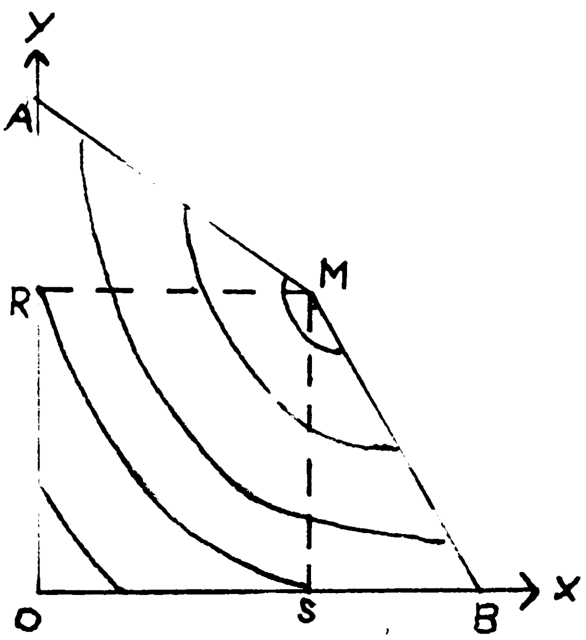


Fig. 10

In this particular case, the two commodities are classed as substitutes on both the Johnson-Allen and Friedman definitions.

---

18 "If two commodities are everywhere completing [complements] according to the former definition [Friedman], they are likewise completing according to the latter [Johnson-Allen]...." (Schultz, op. cit., p619)

19 This follows by deduction. If it can not occur where the ridge lines are positively sloping, and if it occurs at all, then it must occur where the ridge lines are negatively sloping.

In support of this deduction, Allen says that the consequence of negative sloped ridge lines is as follows: "at points in the effective region near the ridge lines... the negative sloped tangent to an indifference curve gets flatter, not only as X increases but also as Y increases. Similarly, the slope becomes steeper in both directions at points of the effective region near the other ridge line." ("The Nature of Indifference Curve", op. cit., p116) In other words, "non-normal" situations occur near the ridge lines when these are negatively sloped.

The connection between the Johnson-Allen and Friedman definitions can be summarized in the statement that only where it is possible for a horizontal line, drawn parallel to the X axis within the effective region, to cut the Y ridge line before it cuts the X ridge line, or where a vertical line, drawn parallel to the Y axis within the effective region, cuts the X ridge line before it cuts the Y ridge line, can non-normal positions occur. In all other cases where the ridge lines have positive slopes, the goods are "normally related" at all points within the effective region.

In evaluating the usefulness of these two definitions, it can be said of both that they are free of any dependence on measurable utility.[21] Only the Friedman definition

---

20 "It is not necessary, in the case of competing goods, that the non-normal positions...should be located near the boundaries of the effective region. All that is necessary for competing goods is that the effective region should contain, somewhere within it, at least one non-normal position." (Allen "A Comparison between Different Definitions of Complementary and Competitive Goods", op. cit., p173).

The above argument disagrees with this statement in maintaining the contrary. In further support of this position consider the isoclines (cf Chapter VII below) of the map. In Chapter VII it is illustrated that all isoclines converge on the point M and that the ridge lines are two special isoclines. Moving anti-clockwise from AM to MB, the indifference slopes defining the various isoclines vary from the smallest slope at AM to the greatest slope at MB. In the area ARM, because the isoclines are downward sloping, increasing X alone means moving from isoclines representing greater indifference slopes to isoclines representing smaller slopes. This is only possible within ARM. Again, within the area SMB, the isoclines are backward sloping so that increasing Y alone means moving from isoclines representing smaller indifference slopes to isoclines representing greater slopes. This also is only possible within SMB. Hence, all "non-normal" positions are located within SMB and ARM.

21 Cf. Schultz, op. cit., p164.

provides a criterion for the borderline relationship of independence. When goods are independent by the ridge line criterion they are complementary by the Johnson-Allen test. Because the ridge line definition provides for independence and also includes the Johnson-Allen test in the case where the ridge line slopes do not vary in sign within the effective region, it is the better definition of the two for the purpose at hand .

In considering four major definitions of related goods,[22] it was necessary to reject the classical definition because of its dependence on measurable utility and the neo-classical definition because of its uselessness for two dimensional analysis. Of the two remaining definitions, the Friedman one was found the better. It may be argued that it is a highly unrealistic and artificial definition because it is based on properties of areas of the indifference map seldom covered by events in the real world. But a definition based on the saturation curves is not as remote from reality or common sense as would at first appear. One would expect , if goods are partly complementary, that the more one had of Y, for instance, the more of X one could take before being saturated with it. So common sense considerations here call for a positively sloping X saturation curve. If X and Y are mild substitutes, it would be expected that the more Y chosen the sooner saturation with X would be reached, and that therefore the X saturation curve would be negatively sloping.[23] There-

22 These definitions are neatly summarized in Schultz, op. cit., pp607-628.

fore, although ridge lines may lie in unrealistic areas they are still relevant for a definition of commodity relations as they are a property of the indifference map and "an individual's introspective, spontaneous, intuitive description of two goods as complementary or competitive (sugar and coffee, beef or pork, *etc.*) corresponds...to behavioristic properties of the preference field and demand functions."[24] The slopes of these ridge lines are a characteristic of the indifference map which will alter in a definite manner for different selections of the commodity axes. They therefore make a suitable indifference property for formulating a definition of commodity relations.

---

23 Cf. Clark, "Realism and Relevance in the Theory of Demand", *op. cit.*, p352. In other words, a good that satisfies one want may partially satisfy the desire for another good, and so the goods are substitutes. The ridge lines must slope towards the axes. Or quantities of a good may increase the want for another good; so the goods are complements and the ridge lines slope away from the axes.

24 P.A. Samuelson, *Foundations of Economic Analysis*, p184. "We see that, as the position of saturation points puts a certain limit to the possible slopes of the indifference elements, the former are more intricately connected with human behavior than the term 'saturation point' suggests at first." (N. Georgesue-Roegen, "The Pure Theory of Consumer's Behavior", *Quarterly Journal of Economics*, Aug. 1936, p557n7)

CHAPTER IV

EXTENSION OF ANALYSIS FOR PARTIALLY RELATED GOODS

A suitable criterion for classifying commodity relations is the slope of the ridge lines. It is necessary now to investigate further properties of this criterion before proceeding to the analysis of the most general plane cross-sections of the hyper figure.

Consider first two ridge lines which are straight lines with positive slopes, such as the lines AM and BM in Figure 11. The two commodities of this map are partial complements on the Friedman definition. If they were perfect complements, AM and BM would coincide and would pass through the origin; whereas, if they were completely independent, AM and BM, while still intersecting at M, would be parallel to the X and Y axis respectively. Since the ridge lines of Figure 11 lie between these two extreme positions, which correspond to the extremes of the complementary scale, the goods must be imperfect complements.

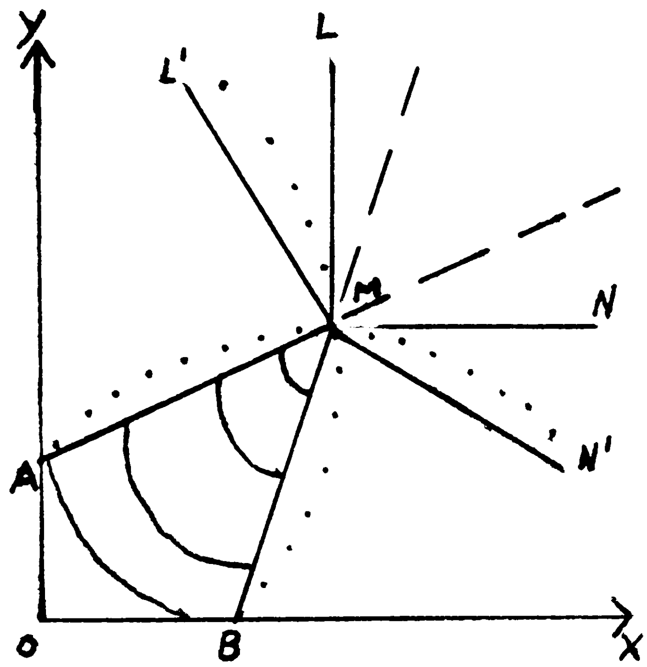


Fig. 11

Consider now what happens at the point M, a point of "bliss", when excess units of X and Y are not nuisances. Since all the indifference curves are parallel to one of the axes outside of the effective region, the ridge lines bend abruptly at M to form the right angle LMN containing the plane of "bliss". This change in slope signifies that at M the two goods become independent and remain independent for all saturated quantities.

A priori, there is no reason to suppose that two goods will abruptly change from mild complements, to complete independents.[1] If it is assumed that the change in the relationship is gradual, so that as the quantities of X and Y increase up to saturation the complementary relation weakens until the goods become independent at M, the linear form of the ridge lines must be replaced by a curvilinear form , such as the dotted curves AM and BM have in Figure 11. The manner in which these curves are drawn indicates that the complementary relation holds all over the effective region but that it weakens as the point M is approached.

What happens now at M if excess units of X and Y are nuisances? Beyond M the marginal utilities of X and Y are negative. The change at M from positive marginal utilities through zero to negative marginal utilities should have some effect on the shape of the ridge lines. It was noted above that the change from some positive marginal utilities to zero marginal utilities had the effect, figuratively speaking, of pulling the linear extensions of AM and BM (the dashed lines in Figure 11) over to lie on LM and MN. So if the marginal utilities become negative beyond M it is quite

---

1 i.e., as between two different batches of X and Y, one just inferior to M and the other indifferent with M, there is no reason to expect that the subject's preference behavior will disclose strong complementarity in the inferior batch and complete independence in the superior one. It might be expected if there was very little complementarity disclosed in the inferior batch,

reasonable to expect the effect of this to be reflected by a negative slope for MNand LM. For, if within the effective region increases in Y, X constant, lead to more preferred positions and at the same time increase the desire for X, increases in Y which lead to less preferred positions should reduce the desire for X. That is, as the excessiveness, or deficiency, of Y grows, the saturation quantity for X is reduced. Stated in another way, if a positive marginal utility for Y acts to postpone saturation for X, a negative Y marginal utility should serve to hasten it.[2] Thus, past the point of "bliss", the X ridge line slopes towards the Y axis and the Y ridge line, towards the X axis, when the excess units are nuisances.

Therefore, under the assumption that excess amounts of both goods are nuisances, and postulating gradual change, the ridge lines appear as the dotted curves BML' and AMN' in Figure 11. The complementary relation gradually weakens as M is approached. At M the relation is one of independence, and past M it is one of substitution.

The gist of the above argument is that the commodity relations, as reflected in the ridge line slopes, change at the intersection point of the saturation curves, so that, in general, the same relationship does not hold at all points

---

2 "It may well be, for example, that, until a certain point, the more butter an individual possesses, the greater the amount of bread he can use; but that, after that point has been reached, an increase in the amount of butter will decrease the maximum amount of bread he can use." (Schultz, op. cit., p617)

on the indifference map.[3] If the intersection of the ridge lines marks the only critical point at which their slopes change in sign, then the same commodity relation holds for all points within the effective region. But if the ridge line slopes change in sign before they intersect then the same relation does not hold all over the effective area. This latter situation is represented in Figure 12.[4] At the

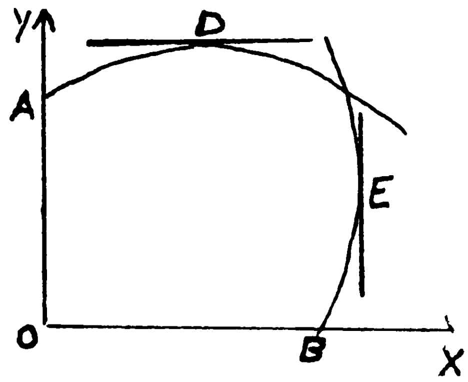


Fig. 12

points D and E the slopes of the ridge lines change from positive to negative values signifying a change in the commodity relationship from a complementary to a substitution one. However, it is not evident why there should be an arbitrary change at these points. It is more likely that the slope of the ridge lines only changes at some critical point. Such a point exists where both marginal utilities become zero. Since there appear to be no critical points on the ridge lines within the effective region, it is argued here that, in general,

---

3 "The relations between competing articles and completing articles are not always so simple, for articles may be competing at some combinations and completing at others." (Fisher, "Mathematical Investigations in the Theory of Value and Prices", op. cit., p76)

Allen does not agree with this but maintains that the same relationship holds all over the indifference map. He draws the saturation curves as straight lines everywhere. cf. "A Comparison between Different Definitions of Complementary and Competitive Goods", op. cit., p169; and "The Nature of Indifference Curves", op. cit., pp116-7.

4 This Figure is a copy of one of Friedman's reproduced in Schultz, op. cit., pp617-19. Both Friedman and Schultz hold, without very good reasons, that the slope of the ridge lines vary in sign, as shown in Figure 12.

the same relationship holds all over the effective region but not over the entire map.

There are a few exceptions to this general conclusion. First, if two commodities are completely independent over the entire effective region, it is most probable that they are independent everywhere on the indifference map. For if positive marginal utilities do not shift the ridge lines, negative marginal utilities should not shift them. If the assumption that excess supplies are not a nuisance is dropped, then perfect complements and perfect substitutes are also relations which hold over the entire map.[5]

For the analysis of more general plane cross-sections of the hyper figure one final adjustment is made to the terms of the fundamental experiment. It is now assumed that in the experiment only those commodities are presented to the subject which are of the type for which excess units can not readily be disposed nor ignored. This restriction, if it can be termed one, has no appreciable effect on the number of goods entering into the experiment since almost all commodities are of such a physical nature that their excessiveness can not be simply overlooked. Therefore, under these conditions, all goods used in the experiment become nuisances when their quantities become excessive.

---

5 Only in these exceptional cases are the ridge lines linear, that is, when the relation is perfect or non-existent.

This is in closer accord with reality than the assumption that excess units are not nuisances.[6]

The most general type of indifference map is obtained when the X and Y axes are so selected that the two commodities to which they refer are only partially related. Consider first a selection of the axes for which X and Y are imperfectly related in a complementary relationship. The complete indifference map for this case is of the type illustrated by Figure 13. The ridge lines AMN and BML are

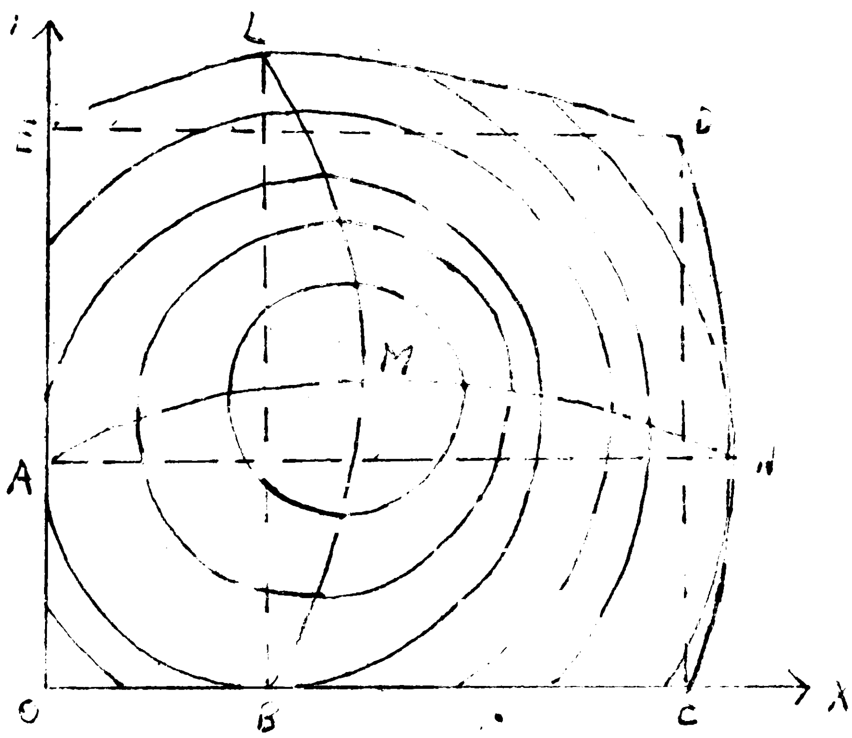


Fig. 13. - Indifference map for partial complements.

6 Ruling out the non-nuisance assumption may make the analysis more realistic, but is it necessary to do so? The answer is yes. The assumption has proved a useful building block so far but it is no longer necessary or relevant. The indifference curves can always be made to take on a positive slope by paying the subject in terms of one good for taking excess amounts of the other. Further, a determinate equilibrium solution for the case of free goods, which surely exists in the real world, requires closed indifference curves, and so the absence of the non-nuisance assumption. Finally, the presence of this assumption leads to the highly unlikely result that, on the ridge line definition, perfect complements become completely independent at the saturation point. Such a drastic change in a relationship seems highly improbable.

upward sloping over the entire range of the effective region, OAMB. The simplifying assumption is made that these curves vary in a continuous and smooth manner. The point E on the Y axis marks that quantity of Y which is so great that it is on a level of indifference with zero Y. Since the amount of X at E is zero, it is also a point of zero total utility, or a point indifferent with the origin. The point C on the X axis is a similar point of zero total utility. The northern boundary of the map is marked by the curve ELD which is the locus of all those Y values which are on a level of indifference with zero Y. This locus has the same general shape as the Y saturation curve since any shift in the point of maximum total utility for Y causes a similar shift in the point of zero utility. Similarly, the curve DNC forms the east boundary of the map representing the locus of all points where the total utility of X is zero. The intersection of these two boundaries at D forms another point of zero total utility which is on the same level of indifference as the origin and the points E and C. The coordinates of the point D are given by OC and OE.[7] The end points, L and N, of the ridge lines are determined by the fact that for two values of one good which are on the same level of indifference to each other, the same fixed quantity of the other good, when

7 The argument is as follows: For OC of X there must be two different Y values which, with the given X, form two batches of X and Y indifferent to each other. Since the batch at C is indifferent with the origin, the other equivalent batch must also be indifferent with the origin. The only other point which could be indifferent with the origin and contain OC of X is D. So OC is one of the coordinates of D. By the same line of reasoning the other coordinate of D must be OE.

combined with these two values, forms two indifference combinations. Thus, at B and L the total utility of Y is zero so that the same amount of X at both those points yields the same amount of total utility. But OB is the saturation quantity for X at B. Therefore OB of X must also be the saturation amount for X at L. Thus both B and L are points on the X ridge line. The same reasoning applies to the positioning of the points A and N.

If the X and Y axes are so selected that the commodities are partially related in a substitution relation, the complete indifference map resembles the one drawn in Figure 14. The main difference between this map and the one represented in Figure 13 lies in the slope of the ridge lines. Here the ridge lines vary in slope from negative

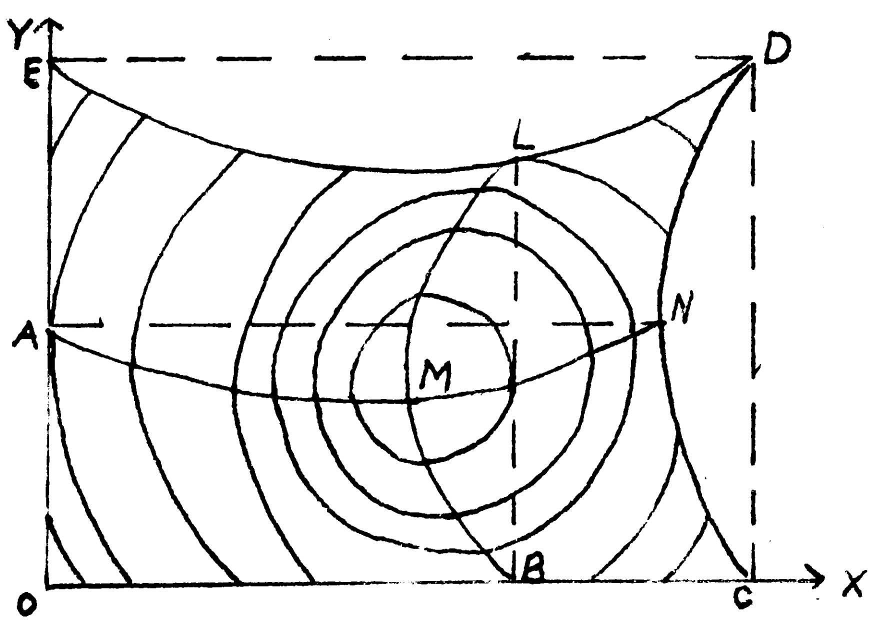


Fig. 14 - Indifference map for partial substitutes.

values to zero at M, and then retain positive values for slope in the saturated regions. All the lettered points and all the boundaries of this map are located in the same

---

The above proof is based on the proposition on dual values laid out on p24 above.

manner as the similar lettered points and boundaries of the map in Figure 13.

Figures 13 and 14[8] represent the most general preference patterns which emerge from the two-dimension analysis of the hyper-figure. Briefly, the assumptions under which they are derived are as follows: first, all commodities are satiable goods, excess units of which are nuisances; secondly, quantities of X and Y inferior to their zero amounts (for the given fixed quantities of all other goods), are excluded from the hyper-figure as not being relevant for a choice experiment; thirdly, the relationship between X and Y is unchanged as to kind over the entire effective region, but not over the entire map,[9] so that the ridge line slopes change in sign only at the point of "bliss"; and fourthly neither X nor Y is perfectly related to any other good.

---

8 These figures differ widely in form, outside of the effective region, from the rather full maps drawn by Allen and Fisher. Cf. Allen's "The Nature of Indifference Curves", op. cit., figures 3 and 4, p116; and Fisher's "Mathematical Investigations in the Theory of Value and Prices", op. cit., figures 20, 22, p70, 71. Only within the effective regions do all three sets of maps look the same.

9 The relationship between any two goods may alter with the location of the origin of their indifference map in hyper space. That is, for one location of the origin the goods may be partial complements, while for a different location they may be partial substitutes. Thus, when the origin is fixed in hyper space, and the frame of reference is altered, the preference pattern alters. Also when the frame of reference is fixed, and the location of the origin in hyper space is altered, the preference pattern may alter.

CHAPTER V

ANALYSIS OF SOME SPECIAL INDIFFERENCE MAPS

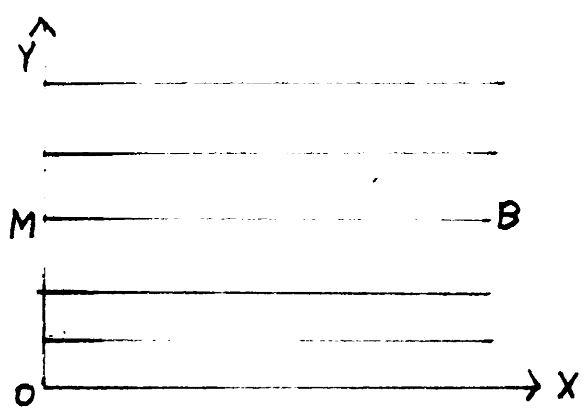
This chapter is an analysis of special indifference maps which can be derived from the hyper figure when no restriction is imposed on the selection of the axes. Only one such restriction has been imposed, and it can now be removed. In the last chapter axes representing goods perfectly related to other goods not appearing on the indifference map were not selected. It is now proposed to consider the indifference patterns which arise when such axes are selected.

First, consider problems connected with complementarity. Let Y and Z be perfect complements and X and Y independent goods, where X is any good having neither a perfect substitute nor a perfect complement. Take Y and Z so that they are not perfectly related to any other commodity except (possibly) X. As in the last chapter all excess units of any good are assumed to be nuisances. Now in hyper space there are a whole series of points associated with zero Y and varying amounts of Z. All planes taken through these points parallel to the XY plane give the same type of cross-section of the hyper figure, except the plane through the point associated with zero Y and zero Z. Consider first this latter possibility.

Since Z and Y are perfect complements any amount of Y is redundant. If the hyper figure were not restricted to exclude all quantities inferior to the origin the indifference map would be like the one in Figure 15A, p55. Any increase in X values, between O and M, lead to higher levels of indifference, while any increase in Y leads to lower levels. The curve passing through the points A,B and O can be considered to be the zero indifference curve. MB is the X ridge line while the X axis is the Y ridge line. Therefore, the point of "bliss" is M, since this is the point of intersection of the ridge lines. However, the restriction of the hyper figure which excludes those amounts of an object, inferior to its zero amount wipes out these indifference curves leaving only the line segment OA.[1] The rest of the XY plane is "empty".

---

1 If Figure 15A is turned around so that the X axis represents the commodity with the perfect complement, and if excess X is not a nuisance, then a linear indifference map is obtained,



identical in form to a preference map appearing in W.Fellner, Competition Among the Few, (New York, 1949), p256. This map is reproduced here and is, in the above book, associated with a labor union, with Y representing wage rates and X, employment. It expresses "the assumption that the satisfaction of the union increases with rising money wage rates, regardless of the amount of employment". (p255) MB is the Y ridge line so that excess Y is a nuisance while excess X is not. This type of preference pattern is ruled out here by the assumption that all excess amounts are nuisances.

If it is not ruled out three analytical interpretations can be given to it. First, the appearance of this type of map map imply that something which is a perfect complement to employment is missing from among "all other things held constant". Secondly, it may imply that all other things held constant include the saturation amount of a perfect substitute for employment. Finally, it may mean that amounts of employment measured along the X axis measure only those amounts in excess of the saturation amount of employment. In other words, it might show that, given a certain amount of employment for all union members, all additional employment adds nothing to union satisfaction.

Professor Fellner also suggests that in the above figure, accompanying this note, the wage ridge line, MB, may lie at infinity. This implies that the desire for increased wage rates

If X and Y are partial substitutes instead of independent, the X ridge line MB slopes towards the Y axis in the unrestricted indifference map and the Y ridge line falls below the X axis. For the restricted map only a line segment exists and the rest of the plane is "empty". Where X and Y are perfect complements or perfect substitutes the cross-section yields the single point O. However, if X and Y are

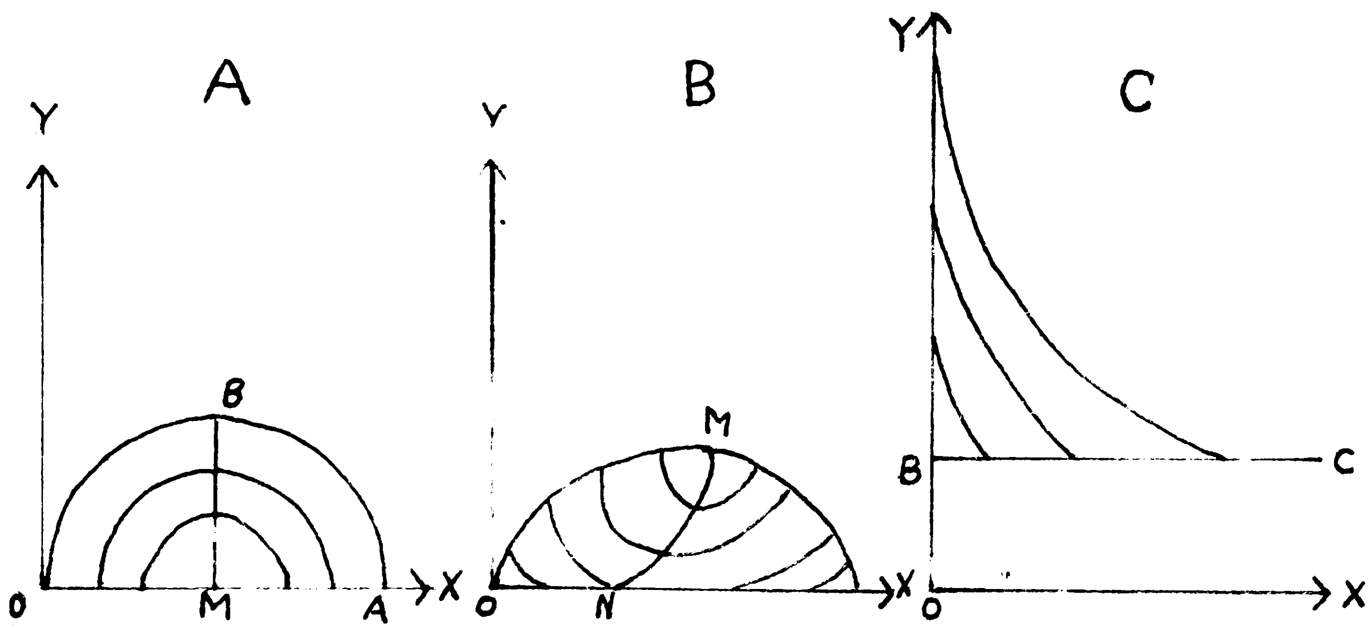


Fig. 15. - Preference maps when Y has perfect complement.

partial complements a partial indifference map exists which has the appearance of the map illustrated by Figure 15B. The Y ridge line, OMA, passes through the origin since for zero X (and zero Z), zero Y is the saturation amount for Y. However, as X is increased, the saturation amount of Y is increased until the saturation amount of X is reached. Beyond this, the

---

is insatiable. The above three interpretations also apply to this case, but here a fourth interpretation is possible. The indifference curves may be circles with their common center given by a finite X value and an infinite Y value - i.e., by the coordinates of the point of "bliss". Segments of these circles would then appear as straight lines parallel to the X axis. The implication of this is that the desire for wage rates is so great relatively to the desire for employment that it completely swamps any consideration the union may want to give to employment.

the saturation amount of Y is reduced to zero again. The Y ridge line is also the locus of all Y values which are on the same level of indifference as the zero Y amount. It is, therefore, along with the X axis, a boundary of the map. All other Y amounts lying within the area bounded by OMA and the X axis are superior to zero Y. Therefore Y is a commodity in this area and so the indifference map of 15B can be derived from the hyper figure.

Suppose now that the cross-sectional plane is taken through a point in hyper space associated with zero Y and M of Z, instead of zero Z. Of course, for zero Y, M of Z is not a commodity.[2] All those points in hyper space associated with zero Y and positive amounts of Z do not lie on the hyper preference figure. Because of this the above plane gives an XY cross-section similar to the one in Figure 15C when X and Y are independent. In this figure OB of Y is just sufficient when combined with M of Z to put that combination of Y and Z on the same indifference level as zero Y and zero Z. The hyper figure does not exist for those quantities of Y less than OB and for M of Z, the space between the X axis and CB being "empty". This indifference map can be made to appear normal merely by shifting the origin to the point B.[3]

---

2 i.e., M of Z is inferior to zero Z since excess units are nuisances. Thus M of Z does not meet the test for a commodity.

3 Above CB the indifference map is normal. The saturation amount of Y is that amount which when combined with M of Z yields a combination of Y and Z superior to any other YZ combination, under the given conditions.

If X and Y are partially related instead of independent, then the line BC in 15C either slopes away from, or towards, the axis, depending on the nature of the partial relationship. If X and Y are perfect substitutes, then the "empty" area is a triangle formed by the lowest indifference curve appearing on the map and the two axes. If they are perfect complements, the lowest indifference curve does not pass through the origin as it does in Figure 6 of Chapter II. Instead it passes through some point representing positive non-zero amounts of X and Y.

What is the situation like if not only Y and Z, but also if X and W, are perfect complements? If a plane parallel to the XY axes is passed through one of the points in hyper space associated with positive non-zero amounts of Z and W the cross-section it yields of the hyper figure differs little from the situation where there is only one perfect complement. Where X and Y are perfect substitutes or perfect complements there is no difference. Where they are independent, partial substitutes or partial complements there are two "empty" spaces, instead of one, on the preference map. There is an "empty" space next to each axis in the same manner as there is an "empty" space next to the X axis in 15C. An appropriate shift in the origin would eliminate these "empty" spaces.

However, if the plane is passed through one of the points on the hyper figure associated with zero amounts of Z and W, it produces a cross-section which is a single point. This is true regardless of the relationship between X and Y. If the hyper figure includes all amounts of X and Y, and if X and Y are independent, then the type of cross-section which appears in this case is illustrated in Figure 16A. The ridge lines are the axes so that the point of "bliss" falls at the origin.[4] The further any one indifference curve is from the origin the lower the level of indifference it represents. Thus if the origin represents the zero indifference level,

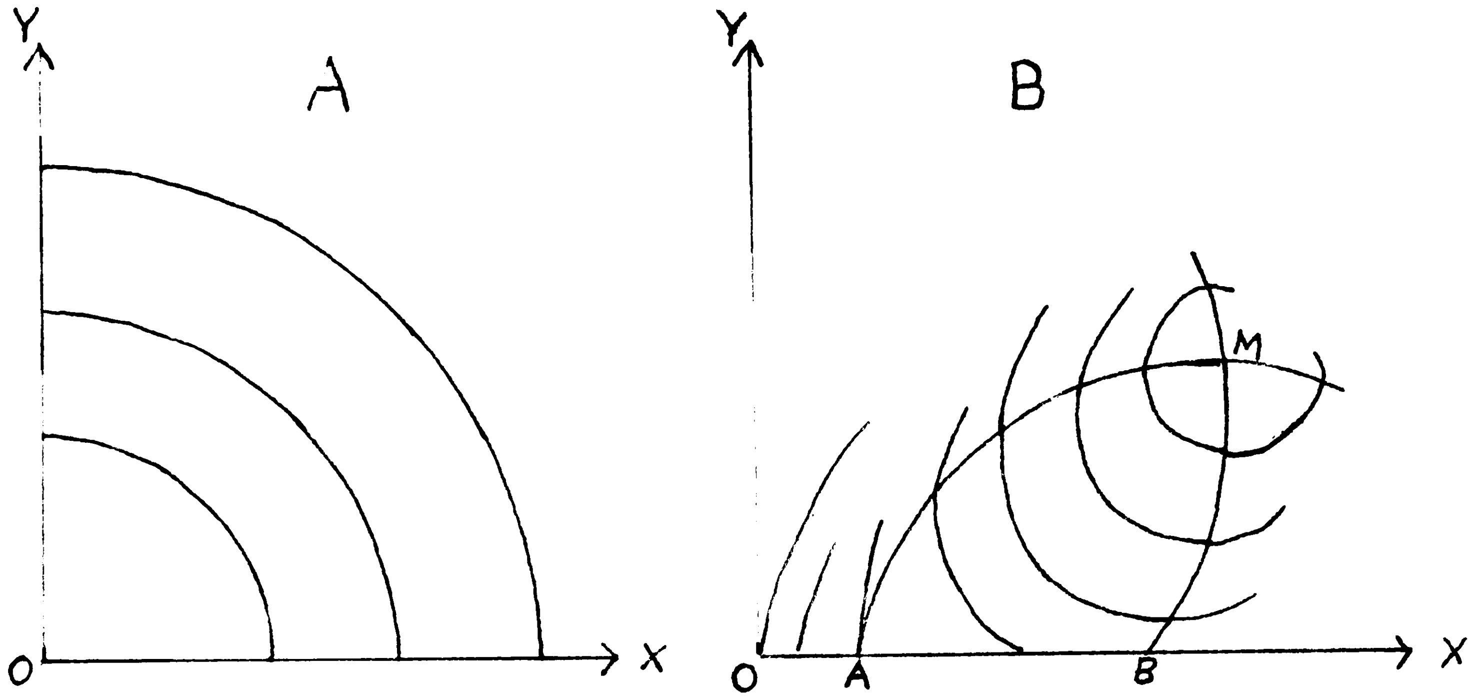


Fig. 16. - Two special indifference maps.

the indifference curves represent different negative indifference levels.[5] The restrictions imposed on the hyper figure eliminates preference patterns similar to this.

To summarize, special cases associated with complementarity fall into three classes. For one of these the amount of a perfect complement for either X, or Y, is held constant at its

---

4 In the Chapter on applications below it is pointed out that all the isoclines of the indifference map converge on the point of "bliss" but not on the origin. In this case where the origin and point of "bliss" coincide all isoclines converge on the origin. This is the only situation in which this could occur.

5 i.e., the curves are negatively calibrated.

zero value. The XY plane is then "empty" except when X and Y are partial complements. This exception gives rise to the second class, illustrated by 15B. For the third class the amount of a perfect complement for X, or Y, is held fixed at some non-zero value. In this third class, the preference pattern only appears normal when the origin is appropriately shifted. If both X and Y have a perfect complement then the resulting situations fall into the first class for zero amounts of the perfect complements, and into the third class for non-zero amounts. For the zero amount of one complement and a non-zero amount of the other complement the existing situations fall into either the first or second class.

Consider next the problems connected with substitution. Take Y and Z to be perfect substitutes neither being perfectly related to any other good unless possibly X. Let X be any other good having neither a perfect substitute nor perfect complement, unless Y. Obviously those XY cross-sections for which Z is zero are normal in appearance. However, those XY cross-sections for which the amount of Z is held constant at its saturation amount are similar to those special cross-sections which arise when a perfect complement for Y is missing. The XY plane is either "empty", or the peculiar map of 15B arises. If not only Y and Z, but also W and X, are perfect substitutes, the XY cross-sections for the saturation amounts of Z and W are similar to those XY cross-sections for the zero amounts of perfect complements for X and for Y. The results are of course "empty" XY planes. The result is also the same if Z is a perfect substitute for

Y whose quantity is fixed at the saturation amount, and if W is a perfect complemement for X whose quantity is fixed at the zero amount. In brief, special cases which arise when a perfect complement is missing from the quantity of goods which are fixed for two dimensional analysis also arise when the saturation amount of a perfect substitute is present in the fixed collection of goods.

Two special possibilities exist for perfect subsitution which are not possibilities for perfect complementarity. One occurs for those cross-sections where the amount of the perfect substitute for Y, Z, is held constant at a value in excess of its saturation amount, and where X and Y are partial complements. This possibility is illustrated by Figure 16B. For any amount of X less than OA all amounts of Y are excessive. However, since X is complementary with Y, additions of X reduce the excessiveness of Y (or of Z) until OA of X eliminates the excessiveness. For greater amounts of X than OA, small amounts of Y are no longer redundant. Thus the consequence of the presence of an excess amount of a perfect substitute for Y is to cause the Y ridge line to cut the X axis. AM is the Y ridge line and BM is the X ridge line. The indifference curve passing through the origin (not completely drawn in the figure) forms one of the boundaries of the map. In this respect the map is somewhat similar to the one in 15B.

The other special possibility occurs for cross-sections where the amount of the perfect substitute for Y is held constant at a value slightly less than the saturation amount, and where X and Y are partial substitutes. In this case the saturation amount for Y is very small when X is zero. As X is increased the saturation points for Y approach the X axis until possibly the Y ridge line cuts the X axis for a value less than the saturation amount of X. The situation where the Y ridge line cuts the X axis for the saturation point of X is illustrated in Figure 22, p89. There the point of "bliss" lies on the X axis.

All XY plane cross-sections of the hyper preference figure have been discussed. In general, the typical cross-sections appear as those illustrated in Figures 13 and 14 of Chapter IV. Less typical are the cross-sections represented by Figures 3, 5 and 6 in Chapter II. Less general are the cross-sections having "empty" spaces in the neighborhood of the axes. Cross-sections which are completely "empty" (except for a segment or a point) are rare, and those similar to Figures 15B, 16B, and 22, and variations thereof, are very special cases. Which particular preference pattern is yielded by any cross-section depends not only on the selection of the axes but also on the point chosen in the hyper figure to represent the origin of the indifference map. For any two given goods the indifference map may be altered greatly with changes in the location of its origin in the hyper figure. Thus, for one location of the origin, the preference pattern may be of a nature which reveals the goods to be partial complements, while for a different location of the origin the indifference pattern may reveal the goods to be independent or partial substitutes.

Therefore, to repeat, the preference pattern depends upon the location of the indifference map origin and upon the two axes selected as X and Y.

CHAPTER VI

A DIGRESSION ON INDISPENSABLE GOODS

This chapter is a digression on indispensable goods.[1] An indispensable good is defined as one for which there is a lower positive limit, not zero, to the amount the subject will, or can, do without.[2] Such a lower limit is called a threshold, and its value is usually determined by the amount of the indispensable good necessary to sustain life. Therefore, the value of the threshold denotes the minimum quantity of the object to which the term good can be applied. Quantities less than this can not be considered to be quantities of a good because such quantities do not serve to satisfy any want. The individual has no wants for amounts of objects less than the absolute minima needed for subsistence. In other words, the desire of the individual for an indispensable good begins with its threshold amount.[3]

---

1 This digression owes its inspiration to a note by Professor A.G. Hart ("Pecularities of Indifference Maps Involving Money", Review of Economic Studies, v. VIII, 1941, pp 126-28) on the same subject. Although Professor Hart has set up the problem in a suitable fashion, his analysis is confused and in error on points of detail. His solutions appear as very special cases of the solutions presented here.

2 "There is a lower limit in the amount of the commodity below which the owner will not or cannot barter. If, for example, you are accustomed to possess two new hats every year, this amount of the commodity might be so commonplace for you that it is taken for granted. You may regard it as a minimum necessity. Such a rate of consumption causes you neither noticeable pleasure or pain." (Thurstone, "The Indifference Function", op. cit., p141)

3 In this sense the threshold amount for indispensable goods is the zero amount.

Absolute necessities, which are the type of goods that have thresholds, are hard to find. For a period of time as short as a given moment, nothing, not even air, is indispensable.[4] However, for a period long enough to permit the individual to starve or die of privation, there are some goods which become indispensable for survival.[5] Since the experiments described in Chapter I exclude the time element it is necessary to describe a third experiment on which can be based the analysis of indispensability.

The description of the third experiment is similar to the description of the second one in every detail except for the introduction of the time element into the former. The introduction of time means that the results of the third experiment hold for some period of time, in contrast to a moment of time. The subject now indicates which collection he prefers from among the alternatives, not for a given moment, but for some fixed period in the future. That is, wants and choices are invariant for changes in time over the

---

4 i.e., it takes longer than a moment of time to kill off an individual.

5 Professor Hart suggests that leisure, as a minimum time to sleep and eat, and that drugs, such as insulin for a diabetic, may be classed as strict necessities when the period of time to which their quantities are referred is of substantial length. Cf. "Pecularities of Indifference Maps Involving Money", op. cit., p126.

Professor Hart also suggests that Hicksian "money", defined as a representative commodity, is an indispensable good. However, this concept of "money" is a tenuous one, the use of which often leads to confusion and difficulties. Therefore, it is best to avoid the use of any such concept of money. Even the introduction of money proper causes trouble because then price changes shift the indifference curves. Cf. Samuelson, Foundations of Economic Analysis, p119 period under consideration. The period taken must be sufficiently long so as to permit at least two of the goods, from among all those considered in the experiment, to become indispensable for certain relevant collections.[6] (By relevant collections are meant those collections which are relevant with reference to the real world.) For those collections which contain zero amounts of almost every commodity, minimum amounts of some of the other commodities will be indispensable when the period of time is of sufficient length. But such collections are irrelevant as they do not represent realistic situations. Within the relevant collections containing the indispensable goods the quantities of all the dispensable goods, and especially of those which are close substitutes for the indispensable ones, are such that they alone can not satisfy the survival needs of the individual.

---

6 i.e., at least two of the goods must be so unique and distinct that minimum amounts of them are indispensable in enabling relevant collections of goods to meet the survival needs of the subject. Of course, these minimum, or threshold, amounts vary from collection to collection, depending on the quantities of partial substitutes for the indispensable goods contained in each collection. But for any given collection containing indispensable goods the threshold amounts are fixed and irreducible. The subject has neither the desire nor incentive to lower the threshold amounts. In the experiment, the individual does not choose amounts less than the thresholds, supposing they were offered to him which is not the case, with the idea of later in the period acquiring more of the threshold goods through a substitution operation. Under the market conditions of Chapter I the operation is possible but the economic incentive for it is missing.

One of the hyper figures obtained by plotting the results of the third experiment is singled out for consideration. Only those plane cross-sections of it which reflect indispensability are selected for analysis. These cross-sections are obtained by taking a suitable point in the hyper figure and by making an appropriate selection of the X and Y axes. A suitable point is one which represents a relevant collection of goods containing a least two indispensable goods. This point, of course, fixes the quantities of all other goods, but X and Y, for two dimensional analysis. An appropriate selection of the axes reimposes the restriction on the X and Y axes which was dropped in Chapter V. That is, X and Y can not be perfectly related to any other goods except each other. The further necessary restriction is added that at least one of the axes must represent one of the indispensable goods. For expository convenience all indispensable goods are considered to be non-satiable ones while all dispensable goods are considered to be satiable ones. The cross-sections which form the subject matter of the analysis of this chapter are then derived by taking a plane through some "suitable" point on the hyper figure parallel to two "appropriately" selected commodity axes. The various possibilities are now considered.

Consider first a cross-section on which the X axis represents some dispensable good not related to Y.[7] Let

---

7 The ridge line definition of commodity relations and independence breaks down when one of the goods is a non-satiable one. However, independence may be defined to

the Y axis represent an indispensable good. The indifference map yielded by the cross-section is illustrated by Figure 17. Ignore for the time being the indispensable nature of Y. Since Y is a non-satiable good, the Y ridge line does not appear because it lies at infinity. BM is the X ridge line

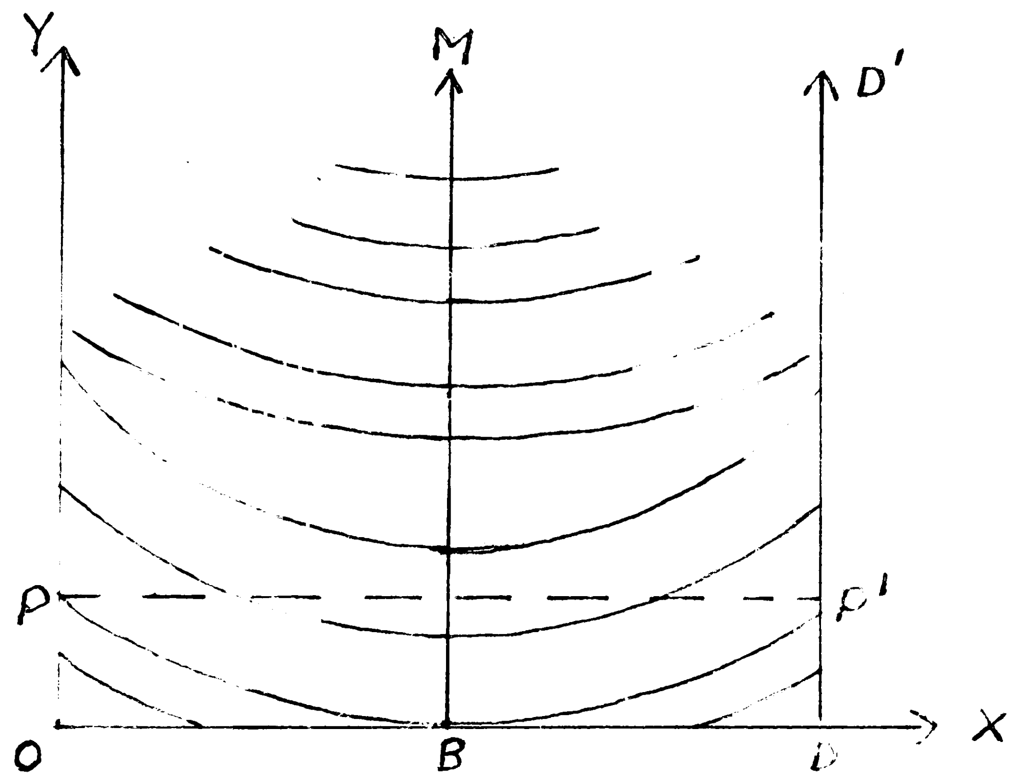


Fig. 17 - Indifference Map for an Indispensable Good.

intersecting the Y saturation curve where the value of Y is infinite. Therefore, the effective region of the map lies within the area bounded by the Y axis, OB, BM and infinity on the north.[8]

Taking account now of the indispensable nature of Y, let OP of Y be the absolute minimum amount the subject requires for survival for the relevant collection of goods selected.[9]

---

exist here if the saturation amount of X for zero Y is the same as it is for infinite Y.

8 A peculiar property of this map is that the indifference curves are parallel in a vertical direction. At any fixed X value the slope of the indifference curves are all equal. That is, $\delta Y \; dy/dx = $ a constant, where dy/dx is the indifference slope and δY the variation in that slope for variations in Y along a line parallel to the Y axis.

Life with less than OP of Y is impossible no matter what amount of X is possessed, since the latter good is not a partial substitute for the former. Therefore, the line PP' drawn parallel to the X axis in Figure 17 is the locus of all those combinations of X and Y necessary for survival. Thus, the hyper figure does not exist for those amounts of X and Y represented by points lying within the rectangular area OPP'D because all amounts of Y less than OP are not amounts of a good and so are excluded from the experiment. Therefore, below PP' the cross-sectional plane is "empty". As a result the indifference curves do not cut the X axis but terminate on cutting PP'. This indicates the fact that substitution between X and Y (along any one curve) breaks down at PP' because the subject can not do with less than OP of Y.[10] The conclusion to be drawn from this is that the only restriction placed on the indifference map by the threshold locus PP' is one preventing the curves from cutting the X axis,[11] and so excluding quantities of Y less than OP.

---

9 The length of OP varies with the initial collection selected. Cf. p65n6 above.

10 This fact does not necessitate redrawing the indifference curves in Figure 17 as Professor Hart suggests. In his analysis he treats such loci as PP' as indifference curves representing zero total utility. From such representation it follows that all higher indifference curves can not cut PP' so that the curves must be redrawn. The error lies in identifying PP' as an indifference curve, and then deriving the whole family of preference curves from it.

11 As will be noted below in Part II, the effect of indispensability on the isoquants is the same as it is here for the indifference curves.
The results of indispensability for the contour curves is similar to the results of perfect complementarity when the

If the X commodity is complementary[12] to Y instead of independent, there is no change in the form of the indifference map yielded by the XY cross-sectional plane. Figure 17 can still represent the derived indifference pattern. The X ridge line is parallel to the Y axis because the X and Y ridge lines intersect where the value of Y is infinite. Even if their point of intersection fell on the Y axis, so that the X ridge line cut the Y axis at an infinite Y value, the X ridge line would still be parallel to the Y axis for all finite Y values.[13] Not only the ridge line, but also the threshold locus, is unaffected by the complementary relation. The minimum amount of Y needed for survival is still OP since no amount of X, which is complementary to Y, can replace one unit of Y in meeting the survival needs of the individual.

---

quantity of a perfect complement for either X or Y is fixed at some non-zero value. Cf. Figure 15C above. Although the results are similar a distinction can be drawn between them. The "empty" area due to complementarity arises from an arbitrary definition of a good. The "empty" area due to indispensability does not depend on any arbitrary definition. The area is truly "empty" because it is impossible for any individual to select any quantities lying within the "empty" area since such quantities would not permit him to survive.

12 By this is meant that the saturation amount of X for zero Y is less than it is for infinite Y. Cf. p66n7.

13 For this reason it is impossible to define commodity relations in terms of the ridge line slopes when either good is a non-satiable one. In the above case only the X ridge line appears on the map and it is always parallel to the Y axis. Professor Hart does not seem to realize this because after defining "money" as a non-satiable good, he states: "since X and 'money' must always be substitute rather than complementary goods, saturation points for X must lie further to the left on higher indifference curves." (op. cit., p127).

However, if the X commodity is a partial substitute for Y, then although the X ridge line is still not affected, the threshold locus is. For if X is a partial substitute for Y, some X units will serve as well as Y units

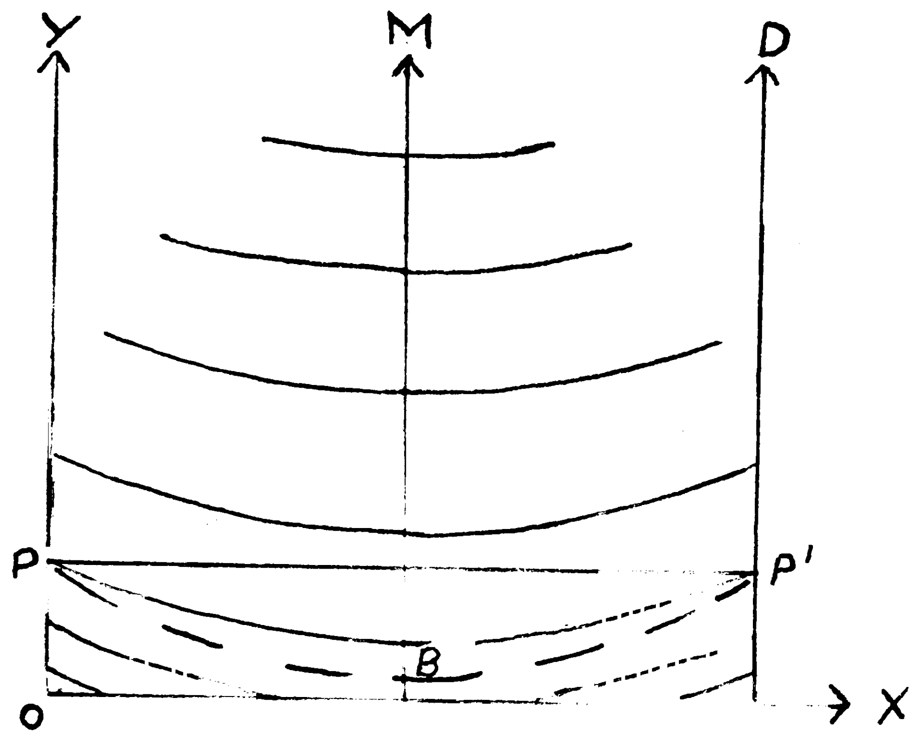


Fig. 18. - Indispensability for substitute goods.

to meet the need for survival. Thus, as the quantity of X possessed increases from zero, the threshold amount of Y is decreased. The map derived in the situation where X can never completely replace Y in meeting the survival need of the individual is depicted in Figure 18. The threshold locus, PBP', lies below the straight line PP' but does not cut the X axis. It would only cut this axis if some amount of X could completely replace Y in the survival combination. Along the section of the threshold locus, PB, increased X reduces the minimum requirement for Y. Beyond B, along BP', increased X increases the amount of Y required for survival since the ability of X to substitute for Y is weakened by its excessiveness.

The special case where the curve PBP' coincides over its entire length with an indifference curve[14] implies that the individual, under the given conditions, requires a given level of utility to survive. The marginal rate of indifference substitution is then identical to, what may be called, the marginal rate of survival substitution. But there is no need for these two rates to be equal. Why should two combinations of X and Y both yield the same amount of total utility merely because they represent a minima necessary for life? For example, two different batches of food may both contain the same minimum number of calories necessary for survival, but the individual may still prefer one combination to the other, depending on his taste. He is then not just indifferent as to which batch he survives on but prefers survival with one to survival with the other. Therefore, in general, the subject will not be indifferent to various survival combinations so that the threshold locus cuts across the indifference curves. It will lie entirely below the indifference curve passing through P in Figure 18 if the individual prefers survival with less X and more Y to survival with less Y and more X. If the reverse is true, then the locus will lie entirely above the indifference

---

14 This situation corresponds to the "general" case for one indispensable good presented by Professor Hart, op. cit., p127 and Fig. 1. Even in this case there is no need to redraw the indifference curves as Professor Hart argues.

curve through P. Its exact path depends entirely on the individual's taste. All that can be said about the path is that it must lie below the straight line PP', because of the substitution relation.[15]

All amounts of X and Y which are commodities are contained by the Y axis, the threshold locus PBP', the line DP' on the east, and infinity on the north. Below PBP' the plane is "empty". To get a less complicated and more normal looking map, it is only necessary to shift the X axis upwards until it passes through the point P. This can be done by redefining the origin to be that point representing the combination of X and Y necessary for survival, under the given conditions, when no X is held.[16]

Now consider the indifference patterns which can be derived when both the X and Y axes are chosen to represent indispensable goods. Since both commodities

---

15 In this case a substitution relation could be defined, or identified, by the deviation of the threshold locus from the horizontal.

16 In the situation illustrated by Figure 18 all the indifference curves are debarred from cutting the X axis. However, if the threshold locus cuts the X axis for some value less than saturation, indicating that some amount of X can replace Y in the survival combination, then the indifference curve passing through that point of intersection, and all higher curves which do not cut the X ridge line, cut the X axis.

Thus, the substitution case differs from the complementary and independent cases in this respect since in the latter cases under no circumstances can the indifference curves cut the X axis.

are non-satiable ones, the saturation curves do not appear on this type of indifference maps, one of which is illustrated by Figure 19. Here OP is the irreducible amount of Y required for survival, under the given con-

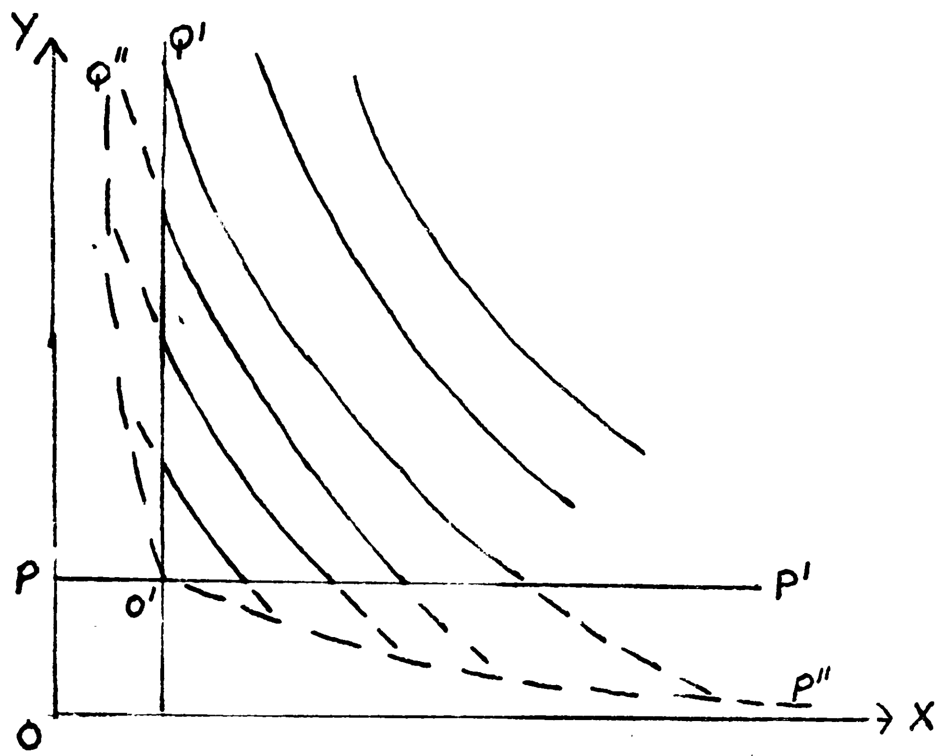


Fig. 19. - Indifference map for two indispensable goods.

ditions, when only the threshold amount of X is present. OQ is the threshold quantity for X, determined in the same manner as OP. Now if X and Y are not partial substitutes[17] for each other, then the initial threshold amounts of OP and OQ remain constant for all quantities of X and Y. The threshold loci, O'Q' and O'P', are parallel to the axes, as shown in Figure 19. The indifference curves do not cut the axes but stop at the threshold

17 i.e., no amount of Y can replace any amount of X, and vice versa. The two goods may then be either independent or complementary, but there is no criterion available by means of which these relations can be classified.

boundaries. Quantities of X below O'P' and quantities of Y to the left of O'Q' are excluded from the hyper figure. The map can be made to appear normal by shifting the origin to O'.

If X and Y are partial substitutes for each other, then the threshold amounts do not remain constant for all quantities of X and Y. As a consequence, the threshold loci deviate from the horizontal in the manner illustrated by the curves O'Q" andO'P" in Figure 19.[18] If X and Y can never replace each other, under the given circumstances, then O'Q" will have either the Y axis, or some vertical line, drawn for a positive fixed X value, as an asymptote. The latter possibility suggests that even with infinite Y some minimum amount of X is required. The same considerations apply to the O'P" locus. When the loci behave in this manner they cannot coincide with any one indifference curve.[19] As a result none of the indifference curves can cut the axes. If, however, X and Y can completely replace each

---

18 The same considerations discussed in connection with the deviation of the threshold locus of Figure 18 apply here as well.

19 i.e., all survival combinations are not on the same level of indifference. This is readily apparent from Figure 19. If all the indifference curves were extended beyond the threshold loci, they would cut the axes. This follows because the goods are non-satiable. On the other hand, the threshold loci can never cut the axes under the above assumption of irreplaceability. Therefore, the loci must cut across the preference curves.

other for certain finite values, then the threshold loci will cut the axes at these values. The indifference curve, or curves, passing through these two points of intersection cut the axes, as do all higher indifference curves. Only when this is possible can the two threshold loci lie entirely on the indifference curve passing through O'.[20] However, the loci are just as likely to cut across the preference curves. It depends on the individual's taste. Consistent choice behavior requires that if O'P" lies above the indifference curve through O', O'Q" must lie below it, and *vice versa*; if O'P" lies below the curve, O'Q" must lie above it. In all cases, by simply redefining the origin of the indifference map, i.e., shifting the origin up to the point O', the indifference map can be made to appear normal.

Therefore, the conclusion which emerges from this chapter is that no change is required in the form of the indifference curves when one or both goods are indispensable. The only peculiarity of indispensability is that it gives rise to "empty" spaces near the axes on the cross-

---

20 If one locus, say O'P", coincides with the indifference curve through O', then the other locus, O'Q", must also coincide with it. If it did not this would indicate inconsistency in behavior. For along O'P" the individual indicates indifference to surviving with less X and more Y, but along O'Q", if it did not coincide with the given indifference curve, it would indicate that the individual was not indifferent to survival with less X and more Y.

The analysis of this special case of coincidence is given by Professor Hart as the general analysis of the preference situation when both goods are indispensable. Cf. *op. cit.*, p127 and Fig. 2.

sectional planes. The only change this could call for would merely be a shift of the origin. The writer must therefore diagree most vigorously with Professor Hart when he claims that the "direct consequences" of a good being a strict necessity are to "restrict us considerably in our drawing of indifference curves."[21]

---

21 _Op_. _cit_., p127.

CHAPTER VII

APPLICATIONS OF EXTENSIONS

In general, only the most typical indifference maps, analyized in Chapters II and IV, are considered in this chapter. With one or two exceptions, all special cases are ignored. The purpose of the chapter is to examine the uses, or applications, that can be made of the extended analysis. Since indifference curves are primarily associated with demand theory, all applications are confined to this field, the idea being to indicate how extending existing analysis contributes to the analytical clarification of neo-classical demand theory. In this field, the extended analysis makes for a clearer statement of the theory than does the existing curtailed analysis. The following applications serve to illustrate this point.

Consider first the usefulness of the extended analysis for the treatment of the isoclines and income-consumption curves of the indifference maps. For illustration take the indifference map in Figure 20, overleaf, for two independent goods. The point M is a single point of maximum indifference. The dashed extensions of the indifference curves beyond the axes are merely geometrical extensions having absolutely no significance for preference behavior. Only if negative values were admissable would indifference curves such as ROS and ABCD exist. The only relevant parts of these curves consist of the point O and of the segment BC. The slope of any one of the indifference curves in Figure 20 (including the irrelevant extensions of all curves cutting the axes) varies continuously from minus infinity at the Y ridge line to zero at the X ridge line. Of course, the higher the curve the

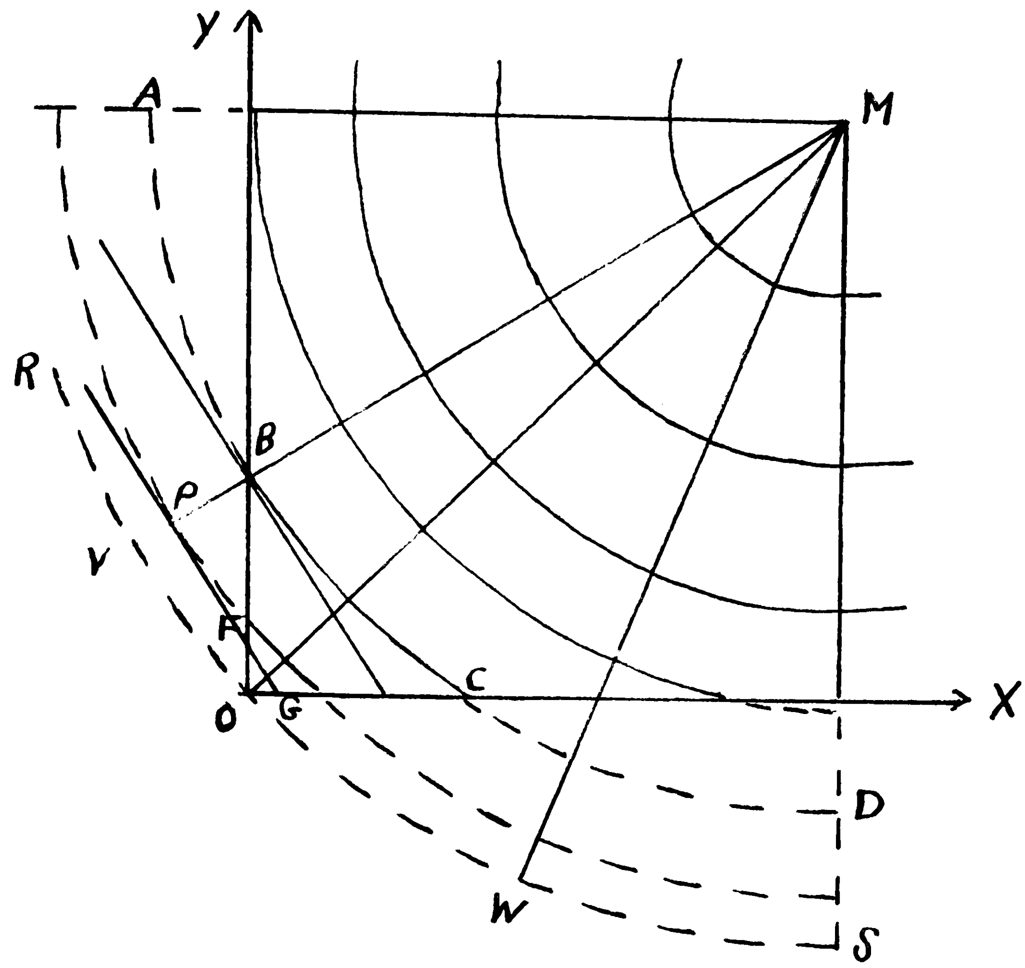


Fig. 20. - Indifference map for isoclines.

greater its curvature. It follows from this variation of the slopes that there must be an infinite number of isoclines — loci of equal indifference slopes — passing through the curves. The only restrictions placed on these isoclines prevent them from intersecting each other or from cutting the same indifference curve twice within the effective region. The ridge lines are two special, or particular, isoclines, and because of this, and the restrictions imposed upon the isoclines, all isoclines converge on the point M. This point, since it represents a unique indifference level, represents an infinite number of indifference slopes. In contrast to M, the origin, O, is only one point on the geometrical indifference curve RS, and so represents just one indifference slope. Therefore, there can be only one isocline passing through the origin, and all other isoclines must pass through different points on the axes.

To determine the precise connection between the isoclines and the income-consumption curves, consider the path traced out by the point representing the equilibrium position of the individual when the income he has to spend on X and Y varies from zero[1] by a very small amount, expenditure on all other goods being held constant. If the ratio of the prices of X and Y is such that the slope of the price line is equal to the slope of the indifference curve RS at O, Figure 20, then clearly the equilibrium point moves up the isocline OM as income is increased, relative prices remaining fixed. For this particular price ratio then the individual purchases both X and Y at every level of income so that the income

---

1 i.e., the initial equilibrium position is at the origin. Income now refers to the money available for expenditure on the two goods X and Y when the expenditure on all other goods is held constant. Any variation in this income must be the result of exogeneous gains or losses in this sum of money.

curve[2] is the isocline OM.

Now suppose that the price ratio is such that for a very small level of income the point of tangency between the price line and an indifference curve lies in an area of negative X or Y values. One such possible situation is illustrated in Figure 20 where the price line FG is not tangent to any indifference curve within the effective region. If the indifference curves could be extended into the negative X area, the point of tangency would be at P. Obviously, the equilibrium point cannot move to P, but can only proceed up the Y axis from O as far as F. As income expands, relative prices constant, the equilibrium position continues moving up the Y axis until the point B is reached. Here the price line is tangent to an indifference curve at a point on the Y axis. From B, as income is further increased, the equilibrium point leaves the Y axis and moves up the isocline VBM. So in this case the income curve consists of two parts. It corresponds to the Y axis for the income range OB, and it coincides with an isocline for the income range BM. For any point on the income curve between O and B the marginal rate of substitution does not equal the given price ratio, yet such points denote equilibrium

---

2 Defined as the curve which "shows the way in which consumption varies when income increases and prices remain unchanged." (Hicks, <u>Value and Capital</u>, p27)

positions for appropriate income levels.[3]

In general, there is only one price ratio for which the income curve coincides with an isocline at all income levels, and that is the price ratio for which the price line has the same slope as the indifference slope determining the path of the isocline passing through the origin.[4] For this price ratio expenditure is divided between the two goods at every level of income. If the prevailing price ratio differs from this unique ratio, then, for low income levels, only one good is bought, and not until income has risen to a high enough level is the other good also purchased.

The good which is bought at all income levels is determined by the price and preference situation. If, when prices are identical, both goods are not purchased at all income levels, then that good which is most

---

3 Cf. Bishop, "Professor Knight and the Theory of Demand", _op. cit._, p149.
This is "the case of items which do not enter into a consumer's budget until income increases or their relative prices decrease to some critical levels." (Samuelson, _Foundations of Economic Analysis_, p70n7)

4 The obvious exceptions of perfect complements and substitutes are excluded. For perfect complements all the income curves coincide and pass through the origin. For perfect substitutes there is also an infinite number of isoclines passing through the origin as well as through every point on the axes. This is because the locus of maximum indifference is here a line instead of a point. Between any given point on the locus and the origin there is only one isocline.

preferred[5] is the one bought, regardless of income restrictions. Of course, if there is no "most preferred" good, then for identical prices both goods are purchased, regardless of income. When the prices differ the situation is more complicated. If the price differential is not too great, the "most preferred" good is still the only one taken (providing the given price ratio does not permit both to be chosen) even if it is the dearer of the two. But as the price differential widens the more likely is it that cheapness, regardless of preference as defined here, decides which is the chosen good. Thus, for every price ratio but one[6] the income curve coincides with one of the axes for a given income range before it coincides with the isocline appropriate for the given price ratio. A typical and normal shaped income curve has the general appearance of the curve OBM in Figure 20.[7]

---

5 By most preferred good is meant here that good whose ridge line is further removed from its axis.

6 Strictly speaking, there are three exceptions, but two of them are trivial. They occur when one of the prices is zero, the other not zero. The income curve then coincides with one of the ridge lines for all income levels and does not lie anywhere on either axis.

7 There has been a great deal of confusion over the exact path of the income curves, possibly because the developers of the new demand theory did not stop to work out all the properties of their curves.

The confusion stems from the obvious fact that all income curves must start at the origin, if income is reduced to zero, and from the complete identification of income curves with isoclines. For if the curves must start at the origin and must, at the same time, be isoclines, then all income curves must converge on the origin without first intersecting the axes. This is erroneous, yet repre-

The point to consider now is the range of slopes, and paths, to which the isoclines[8] are restricted. Obviously, in the case of independent goods, all isoclines slope upwards towards the point of "bliss". Their slopes must always be positive but may vary between the two extremes set by the ridge lines.[9] The effect of a complementary relation is to restrict the number of isoclines while retaining the requirement of positive slopes. The stronger the complementary relation the less the variability in the path of the curves.[10]

---

sentative of the type of income curve analysis presented in economic texts. For examples, cf. Boulding, Economic Analysis, (1st ed) pp677-8 and Fig. 103; and R.T. Norris, The Theory of Consumer's Demand, pp25-6.

Schultz also appears to fail into the same analytical error, for he maintains "that when the commodities are independent $[U_{xy} = 0]$ in consumption an increase in income always brings about an increase in the quantity purchased of each commodity" and that "when the commodities are completing $[U_{xy} > 0]$ ...the conclusions are the same as for independent goods". (The Theory and Measurement of Demand, p49. My italics.) This statement cannot hold for low income levels.

8 i.e., those parts of the income curves which do not coincide with the axes. For the time being, those segments of the income curves which lie on the axes are ignored, and only the isoclines are considered. This simplifies the discussion.

9 Because she identifies income curves completely with isoclines, Professor Norris (op. cit., pp25-6) concludes that it is the axes which set the limits to the variability of the income (isocline) curves.

10 Hicks and Allen, "A Reconsideration of the Theory of Value", op. cit., p62, because they reject considerations of "unrealistic" indifference regions, can only conclude that the variety of possible slopes the income curve can show "is a little more restricted when the elasticity of substitution is low than when it is high". A realization of the complete appearance of the indifference map adds clarity to this point.

If abnormal or negative sloping isoclines (and so negative sloping income curves) exist, they must be associated with substitute goods.[11] Since all isoclines pass through the point of "bliss" they can only have positive slopes when the indifference curves, within the effective region, extend above or to the right of this point. This is only possible when the goods are substitutes, as can readily be seen from Figure 21. Only within the areas

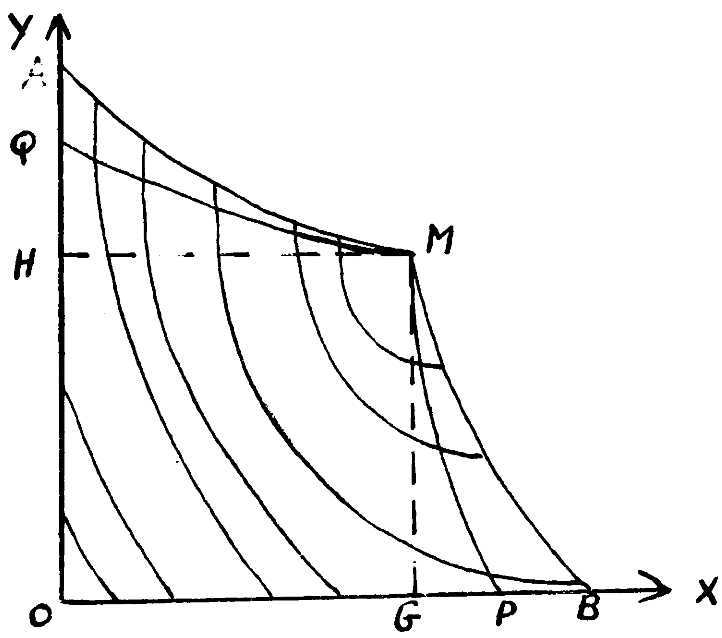


Fig. 21. - Negative sloping income curves.

AMH and GMB can the isoclines have negative slopes. Figure 21 shows at a glance that the greater the substitution

---

11 "Complementary goods have income-consumption and price-consumption curves of an orthodox character.... Both price-consumption and income consumption curves have the same appearance and limits as do the curves of unrelated goods. There is no case corresponding to the backward sloping curve of 'inferior goods' which occurs for less-than-perfect substitutes <u>only</u>." (Norris, <u>op</u>. <u>cit</u>., p42. My italics)

Although Professor Norris does not use the same definition of commodity relations as employed here, it is gratifying to see that she arrives at the same conclusions.

relation (short of perfection) the greater the possibility for a negative income effect. It indicates too that it is only a matter of relative prices and income which determines the inferior good. For appropriate income levels, if the price of X is very low compared to that of Y, then X is the inferior good. If the reverse holds, Y is the inferior commodity. If income, or the price ratio, is such that the point of tangency between the price line and an indifference curve lies within the area OHMG in Figure 21, both goods are normal[12] so that neither is inferior or superior. Thus there is nothing intrinsic in a good which makes it per se an inferior or superior one.[13] Therefore, one should not talk of inferior and superior goods but only of goods which can, and which do, become inferior or superior.

The negative income curve is a relatively unimportant exception since it only arises within the shadows of the saturation areas.[14] A typical abnormal income curve has

---

12 A good is normal if more of it is bought as income rises, and inferior if less is bought. Cf. G. Stigler, The Theory of Prices, p78.

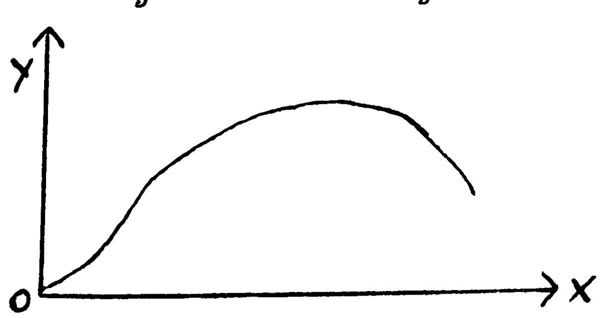
13 "It is possible for a commodity to be 'normal' at some income levels and 'inferior' at others." (Ibid.,p78) For an exceptional case where only one of two goods can become inferior see p89 below.

14 "In the case of complementary goods, it can never happen that one good becomes superior to the other when the individual is nearly saturated with the latter. On the other hand...one of a pair of competitive goods is to be expected to become superior to the other just at these 'saturation positions'." (Allen, "A Comparison between Different Definitions of Complementary and Competitive

some such shape as the curves OPM or OQM in Figure 21.[15] Along these curves income elasticity is plus 1 for the range OP or OQ, but it is negative between P and M or between Q and M. Thus, all abnormal income curves start out by being normal, meaning that for very low incomes goods do not become inferior. Only when income rises to a certain critical level can one good become inferior and the other superior. That critical income level is the level for which both goods enter into the individual's budget since both must be bought before the superior-inferior relation can arise. It is not surprising therefore that a good which is destined to become inferior should be bought exclusively as income expands from a very low level and that the good which becomes superior only enters the budget as the inferior one is dropped. Along income curves such as OQM and OPM each new increment of income is either spent entirely on the commodity destined to become inferior or entirely on the one which has become superior,[16] depending on the income level.

Goods", op. cit., p172)

15 Contrast this with Johnson's income curve ("The Pure Theory of Utility Curves", op. cit., pp491-2 and Fig. 5) reproduced here. It shows increasing income involving at first an increase in the purchase of both goods, and later an increase of X with a decrease of Y.



This mistaken conception of the shape of the abnormal income curve is also held by Schultz, op. cit., pp31-2; by Boulding, op. cit., pp679-80; and by R.D.G. Allen and A.L. Bowley, Family Expenditure, p106.

A special case arises when the income curve is like OHM or OGM, Figure 21. If the income curve is like OHM, then, in the range OH, the income elasticity of demand for Y is unity, while in the range HM, it is zero. Y is on the borderline of "normality". A slight change in relative prices either way makes all the difference here in determining whether Y will be normal or abnormal.

In connection with inferior and superior goods,[17] one special case remains to be dealt with. In this special case one good always becomes inferior when both

---

16 Not only is every increment of new income spent on the "superior " good but some money formerly spent on the "inferior" good is transferred to expenditure on the "superior" good.

17 The discussion of Figure 21 is now complete. It presents nothing new but it makes understandable and concise what hitherto has only been conveyed in vague or implicit statements of theory — statements kept vague because of a lack of a complete conception of the entire indifference map. For example, Hicks and Allen state that "very abnormal expenditure curves [income curves] (downward or backward sloping) are undoubtedly most likely at the extremes of the indifference curves...." ("A Reconsideration of the Theory of Value", op. cit., p62n) The above discussion shows that without a doubt it is the only place they can arise. Further, these authors state that such abnormal curves "may arise whether or not the goods are easily capable of substitution, but they are distinctly less likely when the elasticity of substitution is low." (Ibid., p63) Here it is argued that they only arise when the goods are easily capable of substitution.

commodities are purchased.[18] This is only possible when one of the ridge lines cuts both axes.[19] An illustration of one such situation is given in Figure 22, overleaf.[20]

---

18 "This class of competing goods is characterized by an unsymetrical relation, one of the pair being definitely superior and the other definitely inferior." (Allen, "A Comparison between Different Definitions of Complementary and Competitive Goods", op. cit., p174)

An unsymetrical relationship requires one ridge line to have a negative slope and the other a positive slope. This is possible for a special case not discussed in Chapter V. This special case only arises when a plane cross-section of the hyper figure is taken through the origin of the figure. One of the axes contained by this plane may be any one of the commodity axes, say Y. The other axis must be a straight line in a positive quadrant of hyper space not coinciding with a commodity axis, and it must be perpendicular to the Y axis at the origin. Such a line represents a homogeneous combination of commodities. This unique combination may be treated as a single composite good, X. The cross-sectional plane thus determined by these X and Y axes may yield an indifference pattern reflecting an unsymetrical relationship. For if the composite good, X, contains many close substitutes for Y, then increases in X will reduce the saturation amount of Y. That is, the Y ridge line will have a negative slope. However, additions of Y need not have any effect on the saturation quantity of X, especially if X also contains large amounts of complements for Y. It is even conceivable that the X ridge line will have a positive slope.

If such an unsymetrical relationship exists between X and Y then the composite good, X, can never become inferior, although Y can. However, Y need not become inferior when both goods are bought. This differs from the situation portrayed in Figure 22 below. There Y always becomes inferior when both goods are purchased.

19 Fisher considers a situation such as this but his analysis of it differs widely from the analysis presented here. Cf. "Mathematical Investigations in the Theory of Value and Prices", op. cit., pp73-4.

20 For the derivation of this type of indifference map see pp60-61 above.

From this diagram it is clear that whenever both goods are bought Y must always be the inferior,[21] and X the superior, good. But note that Y need not become infer-

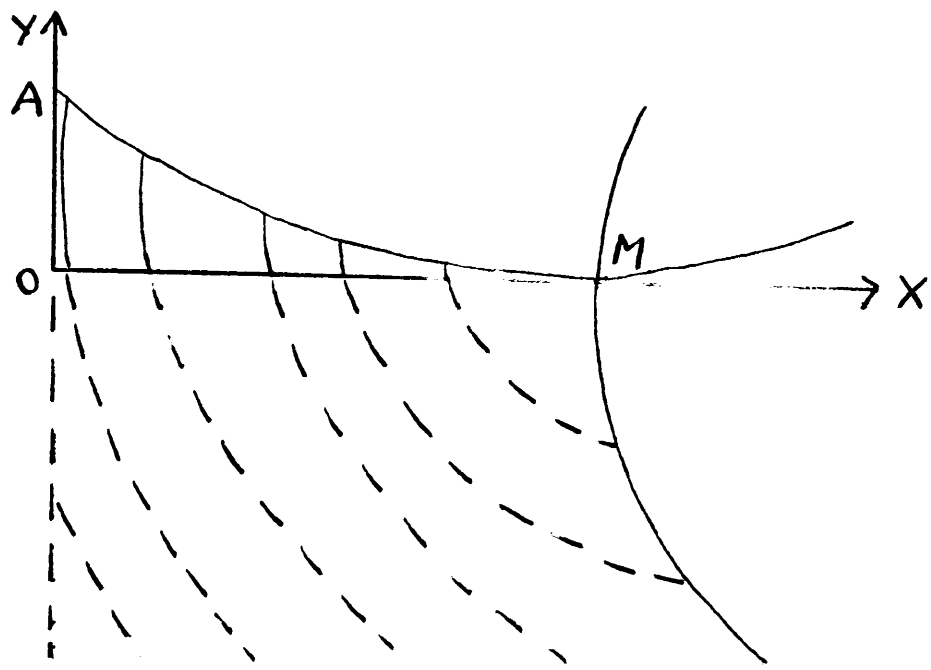


Fig. 22. - Y commodity a "generalized inferior" good.

ior or X superior, and they will not take on these characteristics if only one of the goods ~~alone~~ is bought. Y only enters into the budget if its price is very low compared to the price of X, but if the differential between the two prices is not great then X alone is bought even if it is the dearer good. However, for the appropriate price differential, and for all "significant" income levels, Y becomes inferior.

The income curve has now been dealt with in detail by making use of the extended indifference analysis. Con-

---

21 Y may be referred to as a "generalized inferior" good, to use an undefined term coined by Professor Samuelson, Foundations of Economic Analysis, p169.

sider next the utility of the extended analysis for the examination of the price-consumption or demand curve.[22] Such a demand curve for, say, X is drawn with income and the price of Y fixed, the price of X being a variable. If X and Y are completely independent, the demand curve always shows an increase in the quantity of X bought as its price falls (once X enters into the budget). Its slope varies from high negative values, through zero, to positive values in the fashion illustrated by the curve PQ in Figure 23, overleaf. The lower the price of Y is fixed relative to income (i.e., the larger the Y intercept OP), the greater the variability in the slope of

---

22 This is not a demand curve in the usual sense of the term. Although this curve indicates the amounts of a good demanded at various prices, its context differs from the context of the usual demand curve.

It is a simple matter to move from the demand curve discussed here to the Marshallian one. For an illustration of how this may be done cf. Boulding, "The Concept of Economic Surplus", <u>Readings in the Theory of Income Distribution</u>, esp. Fig. 3A and B, p648.

Although Professor Boulding's method of making the transition is correct, his figure is not. In passing from one demand curve to the other he makes use of what he refers to as marginal rate of substitutioh curves, which are derived from an indifference map. These curves are drawn as a family of non-intersecting (and apparently parallel) straight lines. Actually, in the general case, these curves will be curvilinear, and for substitute goods they all intersect each other at a given positive unit price in Professor Boulding's Figure 3B. In the case of independent goods, the point of intersection of these curves occurs for zero price. Only for complements would these curves be non-intersecting, but they would still be curvilinear.

the demand curve. And of course, the smaller the Y intercept the less the variability in slope. As the Y intercept approaches zero, the demand curve approaches a straight line.[23] The variability of the slope of the

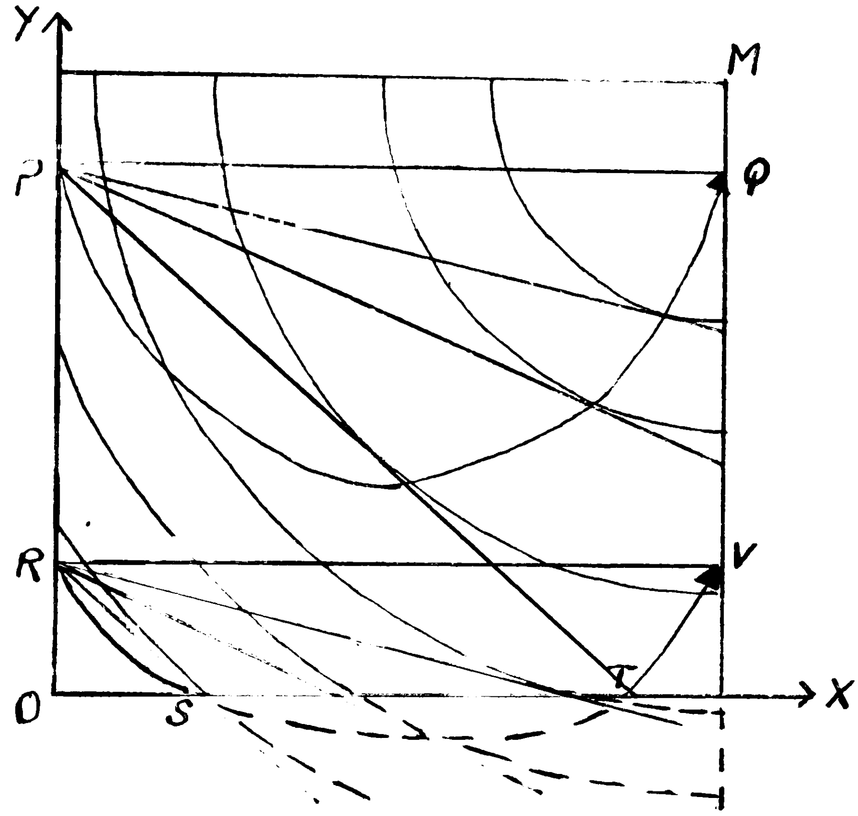


Fig. 23. - Two demand curves for X.

demand curve is also affected by the distance between the axis and the ridge line. Thus, the more easily saturated the individual is with X the less the slope of the demand curve can vary.[24]

A special case arises when the price of Y is fixed high relative to income such that the Y intercept of the demand curve is sufficiently small, and when the X ridge line is far enough removed from the Y axis, to permit some

---

23 i.e., tends to coincide with the X axis.

24 The validity of these propositions can easily be verified by drawing a few diagrams.

of the points of tangency between the various price lines and indifference curves, extended, to lie below the X axis. The demand curve for X then coincides with the X axis for that range of prices within which the tangency points fall below the X axis. The demand curve might appear as the curve RSTV in Figure 23. In the range ST it coincides with the X axis, and at all such points it has a price elasticity of minus 1.

The above discussion on demand curves is valid for independent and complementary goods but a substitution relation calls for further analysis. For substitute goods the demand curve may be re-entrant, i.e., may turn back towards the Y axis. This can only happen when the income curve has a negative slope, since the demand curve through any given point lies to the right of the income curve passing through that point.[25] However, a backward sloping income curve is not a sufficient condition for a re-entrant demand curve. Sufficient conditions require that the price of Y be fixed low enough relative to income, and that the X ridge line be set far enough away from the Y axis, to permit the demand curve to turn back towards the Y axis. These conditions are fulfilled in Figure 24, overleaf, when the demand curve PQ starts from

---

25 Cf. Hicks and Allen, "A Reconsideration of the Theory of Value", *op. cit.*, p66.

the point P on the Y axis. This curve turns back towards the Y axis in the vi~~c~~cinity of the point Q. However, had the demand curve started at P" (the price of

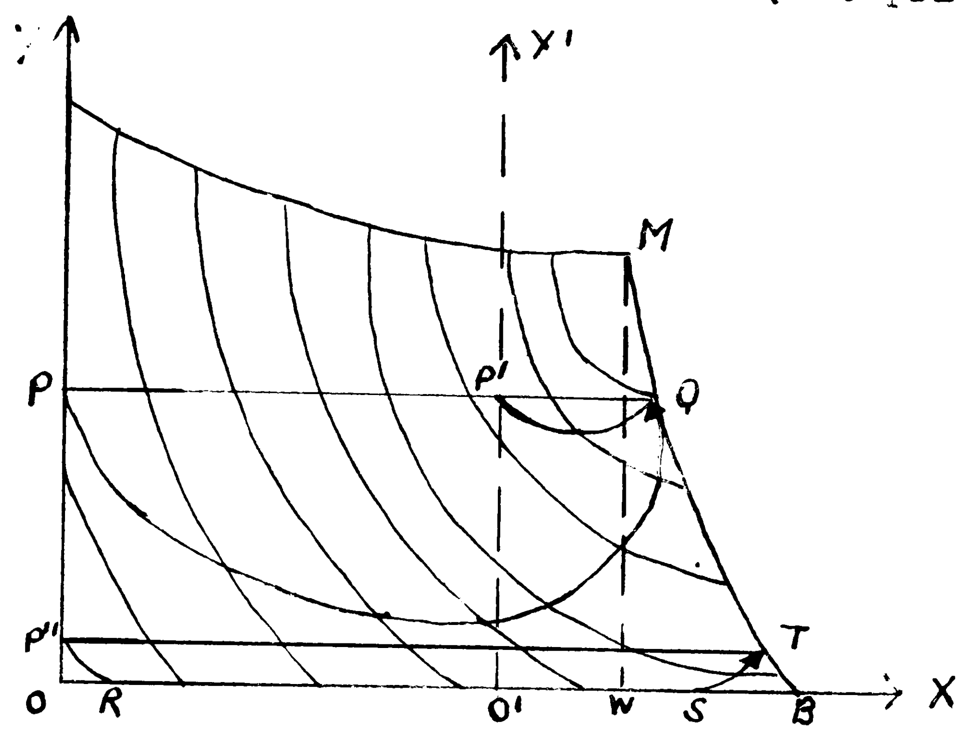


Fig. 24. - Re-entrant demand curve for X.

Y fixed high in relation to income) it w~~h~~ould have traced out the path P"RST, It would then probably not be re-entrant. Likewise, if the origin is at O', instead of at O, so that the saturation amount of X is lessened, and if the demand curve starts at P', it appears from Figure 24 that it would not be re-entrant. So the necessary and sufficient conditions for an abnormal demand curve are much more severe than those for an abnormal income curve.

The Giffen paradox arises when the demand curve is re-entrant, and Figure 24 indicates why it is so unusual. The amounts bought of X and Y must be represented by a point near the X ridge line in an abnormal area[26] such as MWB. The saturation quantities of X must be large and the commodities must not only be very strong substitutes — stronger substitutes than necessary for a negative income curve — but the price of the alternative substitute Y must be low relative to income.[27]

The extension of the analysis has made it possible to formulate more precisely the concepts of income and demand curves associated with neo-classical demand theory. With the aid of these more fully developed concepts an examination can now be made of the possible geometrical connections between the new demand theory and the Marshallian demand theory. Such a connection depends upon the interpretation taken of Marshall's constancy of the marginal utility of income.[28] This interpretation de-

---

26 i.e., an area where the income curves have negative slopes.

27 The lower the Y price the smaller the substitution effect, for there will be less substitution between X and Y if the price of Y compares favorably with that of X. And one condition for the Giffen paradox is that the substitution effect be small.

Also, if the saturation quantities of X are large, then when the equilibrium point resides in the abnormal region the quantity of X purchased will be large. This enables the negative income effect to outweigh the substitution effect.

28 Marginal utility of income is used instead of marginal utility of money since the former concept is less ambiguous. Here we follow Samuelson who holds that "Marshall, when he spoke of money, ordinarily meant nothing more than income". (Foundations of Economic Analysis, p187)

termines geometrically how Marshall's theory is to fit into the indifference map. This, in turn, depends on the variables with respect to which the marginal utility of income is considered to be constant.

Now it can be proved that the marginal utility of income is a homogeneous function of the order minus 1 in each and both prices and in income.[29] Therefore, the marginal utility of income cannot be constant with respect to changes in all prices and income. At most, it can be independent of all but one of these three magnitudes. That is, two first partial derivatives can be set equal to zero, but not three. The two to be chosen can be selected in three alternative ways. One involves constant marginal utility with respect to changes in both prices but not income. The other two involve constant marginal utility with respect to changes in income and one price only.

The last two alternatives imply that the marginal utility of income is constant for changes in the price of X, say, and for changes in income (both real and money). It can be shown that the empirical implications of this interpretation is that any increase in income is spent completely on one good.[30] The geometrical implication

---

29 Cf. P.A. Samuelson, "Constancy of the Marginal Utility of Income", Studies in Mathematical Economics and Econometrics, in Memory of Henry Schultz, ed. O.

of this is that the income curves either coincide with the axes or are parallel to them. The "coincident" condition is met by segments of all income curves, except one,[31] for all low income levels at which the curves coincide with the axes.[32] If X and Y are substitutes, the "parallel" condition is also met by segments of two income curves.[33]

Thus, in the case of satiable goods, Marshall's theory,

Lange, F. McIntyre, and T.O. Yntema, pp78-9.

30 Ibid., pp84-5.

31 i.e., the income curve which coincides with the isocline passing through the origin. The trivial exceptions of the income curves coinciding with the ridge lines are ignored.

32 This condition is broader than the condition that the marginal utility of income be constant. That is, the marginal utility of income need not, and in general will not, be constant along that part of the income curve which coincides with an axis, but for points on that segment the results are the same as if the marginal utility of income were constant.

33 When this condition holds the marginal utility of income cannot be constant, but the results are the same as if it were. The reason that it cannot be constant is that, the goods being substitutes, marginal utilities are not independent. Along any income curve the ratio of the marginal utilities is constant. When the income curve is parallel to an axis, the quantity of one good is fixed, and for independent utilities the marginal utility of the fixed quantity does not change so that the marginal utility of the variable quantity must also be constant. If the variable quantity is called income, then the marginal utility of income is constant. But when utilities are not independent, changes in the variable quantity affect the marginal utility of the fixed amount which in turn reacts on the marginal utility of the variable quantity. Thus, marginal utilities cannot remain constant, although a ratio of marginal utilities can.

when interpreted as meaning that the marginal utility of income is constant with respect to changes in income and one price, is identical to the new theory for the situations in the latter where segments of the income curves coincide with, or are parallel to, one of the axes. On this interpretation then it applies to all price ratios (but one) when income is in the vicinity of zero and to two price ratios when income is not in the vicinity of zero. It is only in such circumstances that the income curves, or segments of them, behave in the manner required by the assumption of constancy of marginal utility of income.

However, where one of the goods is a non-satiable one (let it be the numeraire or "money" good) Marshall's theory holds, on the above interpretation, for as many price ratios and income levels as does the newer theory.[34] For when the numeraire good is a non-satiable one the ridge line for it lies at infinitely great values. Therefore, the other ridge line must always be parallel to the numeraire axis, regardless of the relationship between util-

---

34 Those who support the assumption that Marshall meant by constant marginal utility of income constancy with respect to changes in income take this to imply that the numeraire good is a non-satiable one, so that the indifference curves are parallel in the vertical direction. Cf. Hicks and Allen, "A Reconsideration of the Theory of Value", op. cit., pp64-5; Boulding, Economic Analysis (rev. ed), p762; Johnson, "The Pure Theory of Utility Curves", op. cit., p499; Samuelson, Foundations of Economic Analysis, p191.

ities.[35] What is true for this ridge line is true for all isoclines. They must all be parallel to the *numeraire* axis, which is also an isocline. Thus, all parts of the income curves are either parallel to the *numeraire* axis, or coincide with the satiable good axis, indicating that any increase in income is spent either on one or the other good.[36]

If both commodities are non-satiable ones then all the isoclines are parallel to each other but are not parallel to either axis.[37] Therefore, the theory would apply to all price ratios, except one, only for incomes in the vicinity of zero. This preference situation is the most restrictive of the three considered to Marshall's theory, on the given interpretation.

---

35 See the discussion on this point on p69 above.

36 Once again this condition of parallel isoclines is broader than the condition that the marginal utility of the *numeraire* be constant. Only for independent utilities would this marginal utility be constant. If utilities are not independent, the marginal utilities will not be constant. Only the ratio of marginal utilities is constant. So the final result is the same as if the marginal utility of the *numeraire* were constant. Cf. p96n32,33 above.

37 In this case both ridge lines intersect at the point (∞, ∞). Since all isoclines converge on this point they must be parallel straight lines because only parallel lines intersect at infinity.

Professor Allen, "The Nature of Indifference Curves", *op. cit.*, p114, asserts that the form usually assumed for the indifference curve system rests on the assumption that the individual's purchases on the market are so small compared with the saturation values of X and Y that as a first approximation these can be taken as infinite. Because of the behavior of the isoclines under such conditions, this form is useless for the analysis of negative income effects and the Giffen paradox.

The alternative method of considering Marshall is to assume that he meant to hold the marginal utility of income constant with respect to changes in both prices but not with respect to changes in income, real or money. It can be proved that the implication of this hypothesis is unity income elasticities.[38] The geometrical implications of this is that the income curves must either coincide with, or be parallel to, an axis, or be a straight line passing through the origin. On such an interpretation Marshallian theory can only hold for one (at most three) special price ratio and all income levels, or for all price ratios and those income levels in the vicinity of zero.[39]

Professors Knight[40] and Keirstead[41] implicitly accept this latter interpretation of Marshall by arguing that he could not have meant to hold the marginal utility of income constant with respect to changes in income. However, for them the literal implication is different. They take it

---

38 Samuelson, "Constancy of the Marginal Utility of Income", op. cit., pp80-1.

39 If all income curves converged on the origin in accordance with the analysis of Professors Boulding and Norris, then Marshallian theory would hold for all price ratios when these curves were linear, on the above interpretation.

40 "Realism and Relevance in the Theory of Demand", op. cit..

41 The Theory of Economic Change, p188. Also cf. Milton Friedman, "The Marshallian Demand Curve", Journal of Political Economy, v.lvii (Dec. 1949), 463-495.

to imply not only that the marginal utility of income is constant with respect to changes in both prices but also that changes in income, both real and money, do not take place. Therefore, if the price of one good falls, there must be a compensating price change in the other price so as to keep the price level and real income constant.[42] The equilibrium position of the individual is then represented by points lying on the same indifference curve, and the only effect of a price change is the Hicksian substitution effect.[43] There is no income effect for it has been assumed away. From this viewpoint, Marshall's theory is only one part of the Hicksian theory.

This attempt to find a link between Marshall and Hicks reveals that according to the viewpoint taken as to the meaning of constant marginal utility of income, Marshall's theory can be represented in indifference analysis as the case where the income curve is parallel to, or coincides with, one of the axes; where the income curve is also a straight line passing through the origin in addition to where

---

42 Professor Keirstead argues that unity elasticity of demand for a good is an exception. But this cannot be so. There must always be a compensating price change, no matter what the elasticity of demand, if real income is to remain constant. For once permit the price level to change and real income changes.

43 By this is meant the substitution effect illustrated in the text of <u>Value and Capital</u>, and not the one laid out in the mathematical appendix, which is the same as Slutsky's substitution effect. For a discussion of the difference between the two cf. J.L. Mosak, "On the Interpretation of the Fundamental Equation on Value Theory", <u>Studies in Mathematical</u>

it is parallel to, or coincides with, one of the axes; where only the substitution effect is present. Since it is not clear just what Marshall's assumptions mean, it appears desirable to replace his demand theory with the neo-classical theory in which it is plain enough what the assumptions are.

Economics and Econometrics, in Memory of Henry Schultz, pp69-74.

APPENDIX TO CHAPTER VII

A NOTE ON OFFER CURVES

Since an offer curve is derived in the same way as a demand curve it is appropriate that it too should be discussed here. However, before proceeding with the analysis of the offer curve, the indifference map must be redefined so that it applies not to an individual but to a whole community of individuals.

Consider a hypothetical community made up of a large number of individuals. Suppose that in this open economy only the two goods, X and Y, enter into use.[1] For every individual there is an indifference map for these two goods, with probably no two of these maps being identical.[2] Suppose also that for each level of total real income for the community (as measured by the aggregate amounts of X and Y possessed by members of the community) it is possible to determine exactly what each person's share is in that real income.

Given an initial real income which the community enjoys at a given moment, it is possible to conceive of a large scale experiment in which each member of the community is asked to vote on a proposed change from this initial real income. Everyone is asked to vote whether he wants the community to keep the present aggregate combination of X and Y it now holds or to exchange for some

---

1 It is simpler here to deal with just two goods instead of with a large number of goods.

2 i.e., tastes will not be identical.

other proposed aggregate combination. To determine how he should vote each citizen computes the amount of X and Y that will fall to him at the new national income level and refers to his indifference map, as it exists at that moment, to compare his prospective new real income with his present one.[3] He votes for the proposed change if he will be better off for it; against it, if worse off; and abstains, if indifferent.

It is now necessary to draw up rules for counting the votes. The situation where a simple majority vote decides the fate of a proposal is ruled out by assigning to every voter the power to veto any change. One "no" vote is then sufficient to prevent any change. Now suppose that those who stand to gain from any change attempt to bribe all "no" voters. If the loss of the "no" voters has to be made good to prevent them exercising their veto, and if the bribers (all the "yes" voters) stand ready to pay away all their gains in bribes, then on the supposition that no one subject objects to bribing, the proposed real income can be defined as inferior to the present one if, when all gains are paid out in bribes, there are still some negative votes cast; superior, if all "no" votes are eliminated before all gains are paid out in bribes.[4] The two alternative

3 i.e., to compare whether his new share of X and Y lies on, above, or below the indifference curve on which his present share falls.

4 In the first situation some people suffer a loss of real income while all others are no better off. In the second situation some people enjoy a gain of real income while all others are no worse off.

aggregate incomes can be defined to be on the same level of indifference in the limiting case where the total amount of all gains paid out in bribes exactly matches the total amount of all losses. In such a case no votes will be cast either for or against the change. As between the two alternatives, everyone's real income will be constant. By considering a sufficient number of changes from the given initial position enough points can be obtained to trace out an indifference curve. Along this community curve the distribution of real income among all the citizens is the same as it is at the initial position so that every point on the curve is indifferent to the initial point from everybody's point of view.

Now since this community indifference curve represents a given distribution of real income,[5] it will have the same geometric property as an individual's indifference curve.[6] However, there will be an infinite number of such curves passing through the initial point, call it P, each curve corresponding to a different distribution of real income. This is because the distribution of incomes at P enters into the calculation of the total amount of loss (measured in terms of, say, X) which must be compensated for; and the amount of compensation, measured in terms of Y, depends on the average rate of substitution between X and Y at which

5 i.e., constant absolute amounts of real income for everyone at all points on the curve.

6 Cf. T. de Scitovszky, "A Reconsideration of the Theory of Tariffs", Readings in the Theory of International Trade, p366. The method of defining community indifference curves laid out here is based on this article.

each person has to be compensated for his loss, which in turn depends on distribution. For small changes from P, the slope of an indifference curve through P depends on the ratio of the amount of compensation paid in terms of Y to the amount of loss in terms of X. Change the distribution of incomes at P, which has to be maintained, and the amount of compensation and loss changes, thus altering the slope of the indifference curve. Once given the distribution at P the indifference curve through that point is uniquely determined.[7] It is an indifference curve in the sense that all points on it are on the same level of indifference only if the distribution of real income is similar at all points.

To proceed from one community indifference curve to a family of such curves which obey the rules of an individual's indifference map, the assumption must be made that the distribution of welfare (real income) is the same at each indifference level. That is, as between any two adjacent curves, the higher curve indicates that everyone has more income than he has on the lower curve and that the percentage difference in incomes at the two levels is the same for everyone.[8] The community indifference

---

7 Ibid., pp364-66.

8 i.e., if one individual's real income is 10% more on the higher curve, every individual's income is 10% more on the higher curve. This implies cardinal measurement. (Cf. ibid., p366)

Without cardinal measurement the community indifference map does not obey the rules of the individual indifference map. With just ordinal measurement the various community curves intersect each other so that in moving from curve to curve the distribution of incomes varies.

That the analysis which follows below must hinge on cardinal measurement may seem to many a sufficient reason map then exists for a given distribution of income so that each indifference curve represents the same income distribution, but different curves different levels of welfare.

The distribution of real income for which the community indifference map is drawn is the distribution prevailing in the community at its present position of equilibrium. It is assumed that this distribution pattern is maintained as a matter of policy at all points on the indifference map. Changes in the equilibrium position are not permitted to upset this prevailing pattern. Only changes in policy will affect it.

With a community indifference map defined in this manner, a Marshallian offer curve can be derived from it in the same manner as a demand curve is obtained from an individual map. The only difference between the offer curve and demand curve is that the former need not, and usually will not, start from a point on an axis but may have its origin at some point within the effective region of the map. This point, of course, represents the initial quantities of X and Y which the community holds prior to trade. Therefore, it follows from the demand analysis that the more specialized the community,[9] <u>cateris paribus</u>, the

---

for not proceeding further. But as an indication of what the analysis would be like if cardinal measurement were possible, and for those who are not convinced that cardinal measurement should be ruled out, the discussion may be of some use. For a support of cardinal measurement cf. J.C. Weldon, "ANote on "Measures of Utility", <u>The Canadian Journal of Economics and Political Science</u>, May 1950, 16, 227-233.

9 i.e., the nearer the origin of the offer curve is to one of the axes.

greater will be the possibilities for trade. Also, the less easily saturated the community is by an import good, the greater the number of exchange rates at which trade can take place.

When X and Y are complementary or independent goods, the elasticity of the offer curve varies continuously from plus infinity at the origin to zero elasticity as the variable price falls to zero. There is an exception to this

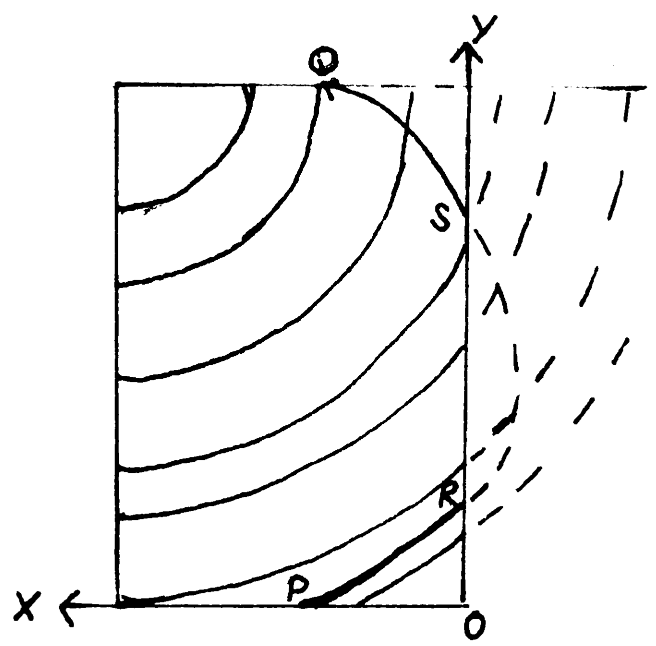


Fig. 25. - Offer curve for restricted supply of X.

similar to the special case illustrated in Figure 23 (p91) where the demand curve coincides in part with a section of an axis. It can arise here when a country is small so that the amount of goods it has to trade with is limited. Turning Figure 23 around and making the X axis the Y axis, the situation illustrated in Figure 25 arises. The offer curve PQ coincides with the Y axis for the range RS where the elasticity of demand for Y is unity. This occurs only because the elasticity of supply of X is zero.[10]

10 Cf. R.F. Fowler, "The Diagrammatical Representation of Elasticity of Supply", Economica, 1938, p225, where the formula for the elasticity of supply of the offer curve is given.

When X and Y are substitute goods it is possible for the offer curve to have a negative elasticity of demand.[11] The conditions under which this may arise are illustrated in Figure 26. The goods must be very strong

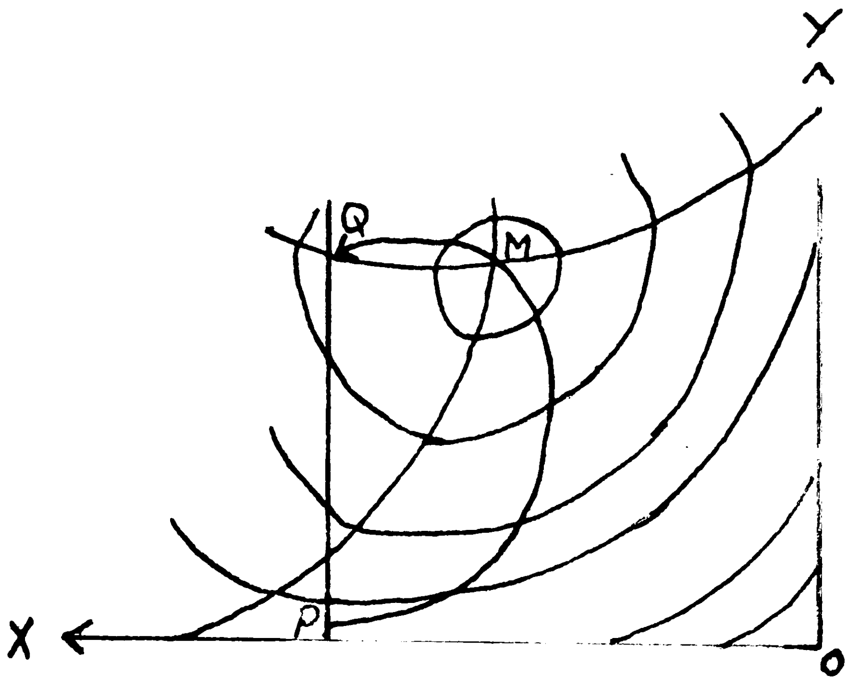
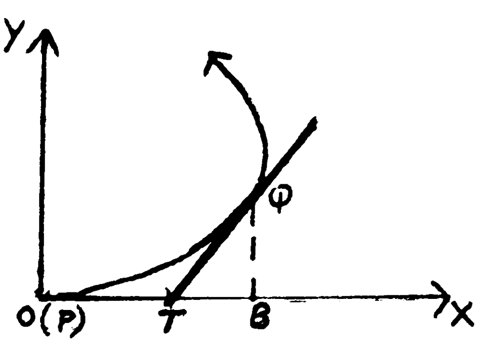


Fig. 26. - Offer curve with negative elasticity.

substitutes and the community must have such a large amount of one good, in this case X, that it is nearly saturated with it. At the same time it must possess very little of the other good. Somewhere within the range QM. Figure 26, the offer curve could have a negative slope, and so a negative elasticity; it all depends on the degree of curvature of the ridge line.[12] In this abnormal

---

11 All elasticities are measured with respect to the quantities of goods exported and imported. Therefore, to obtain the elasticity of demand from the offer curve as drawn in this chapter, the origin of the offer curve, P, must be taken as the point of intersection of the axes. The offer curve when referred to these new axes would appear as in the accompanying diagram. The demand elasticity at any point Q on this offer curve is given by the ratio PB/PT, where PT is the intercept on the X axis of the tangent to the offer curve at Q. When this intercept is negative, the demand elasticity is negative.



12 Had it been assumed that the ridge lines were linear, or that the substitution relation held all over the map so

situation a very favorable rate to the community may be too favorable in the sense that welfare would be greater if the rate were less favorable. This arises when the exchange line cuts the offer curve within the range QM. This is a real-

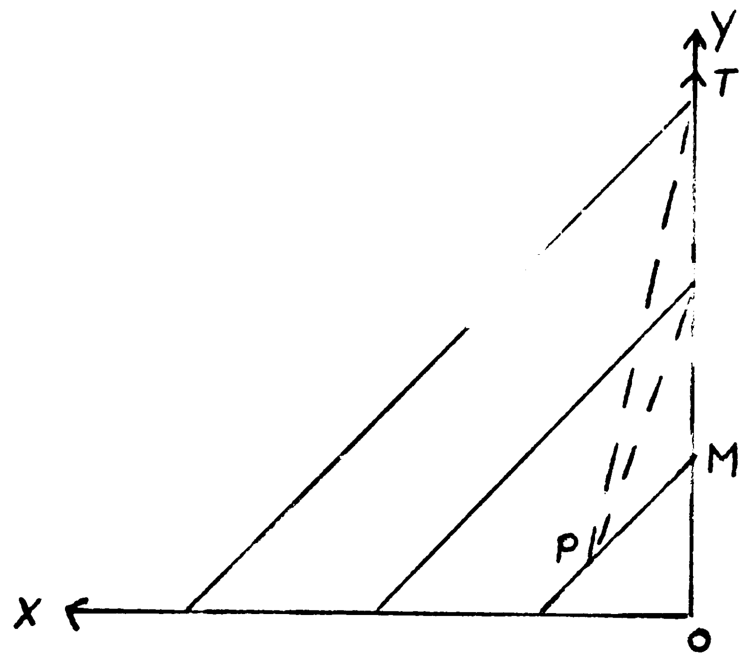


Fig. 27. - A perfectly elastic offer curve.

istic possibility because if a foreigner wishes to export a large quantity of a commodity to a community already in possession of a great bulk of close substitutes for the good, he must sell at an exchange rate very unfavorable to himself. Under these conditions the community may actually benefit from a tariff imposed against its exports if such a tariff does not alter the exchange rate too greatly.

The necessary and sufficient condition for a perfectly elastic offer curve is that X and Y be perfect substitutes for each other. When the initial quantity of such goods held by the community is represented, as in Figure 27, by the point P, the offer curve coincides with the indifference curve PM for the exchange ratio given by the slope of

---

that the ridge lines always have a negative slope, there would be no question about the offer curve having a negative slope between Q and M.

PM. For any exchange line steeper than PM the offer curve coincides with the Y axis. The complete offer curve is therefore PMT. In the range PM its demand elasticity is plus infinity while in the range MT it is plus one. If a foreign offer curve cuts it in the range MT there will be trade, but if it cuts it in the range PM it is doubtful if there will be trade. In the latter situation the domestic community does not stand to gain anything from exchange. It is a matter of indifference to it whether or not it trades.

Consider now the problem of the tariff distorted offer curve. The imposition of a tariff may be thought of as a change in distribution policy. The result of the tariff is to alter the prevailing distribution of incomes in the community because the incomes of importers are lessened and the proportion of the national income going to the government alters vis a vis the share going to the private sector of the economy. At the new equilibrium position for the change in policy the distribution of incomes is therefore different than the distribution at the old equilibrium position. On the assumption that the new equilibrium distribution of incomes is maintained for all indifference levels, all the community indifference curves are shifted in the same manner by the imposition of the tariff. The tariff distorted community indifference map then still obeys all the rules of an individual indifference map.

Since the imposition of the tariff shifts the offer curve to the left, it must affect the indifference curves by making them steeper at every point, as illustrated in Figure 28. The broken lines and curves constitute the tariff distorted map. This diagram indicates that the new Y ridge line must lie below the old Y ridge line, and that the

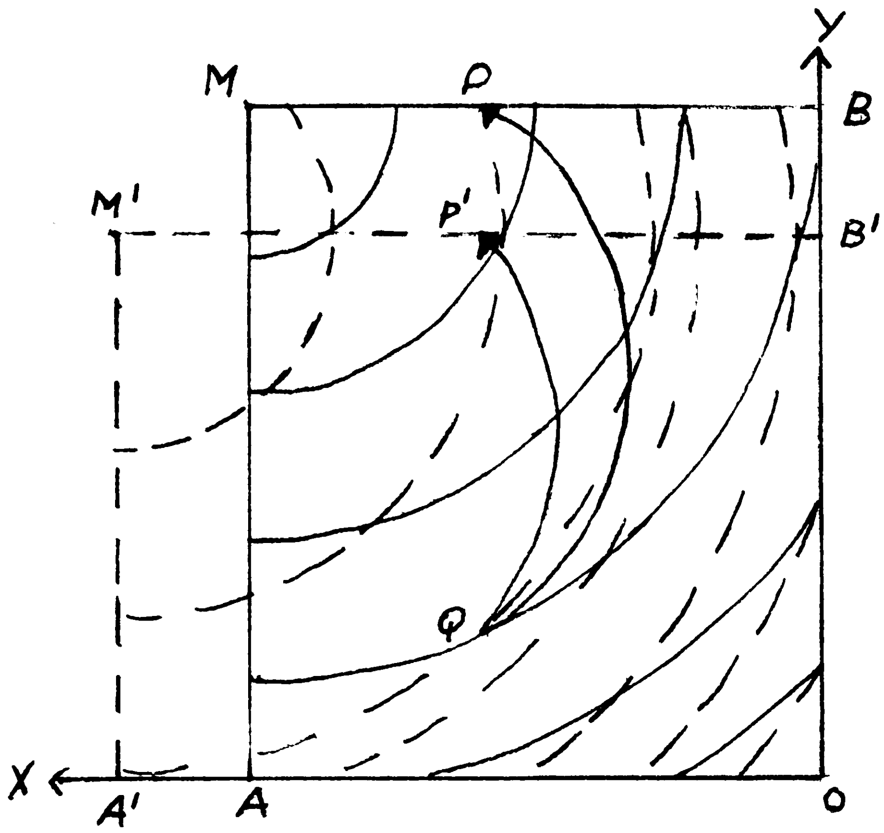


Fig. 28. - A tariff distorted offer curve.

new X ridge line must lie further to the left than the old one. The origin of the offer curve, Q, is not affected by the imposition of the tariff, but the free trade offer curve, QP, is shifted to the position of QP'. On the new map QP' joins all points of tangency between the new indifference curves and the price lines.

Figure 28 also shows that, under the conditions assumed, the effect of the tariff on the preference scale is to weaken the desire for imports, in this case Y, by causing the saturation points of Y to come sooner than they otherwise would. Similarly, the desire for the home good, X here, has been strengthened by the shift of the X ridge line.[13]

13 The tariff has these effects not because it actually alters preferences(i.e., individual indifference maps) but because it alters the given distribution of incomes.

The total effect of this change is to narrow down the possibilities of trade and to require higher or more favorable terms of trade if the community is to enter into trade.

CHAPTER VIII

SUMMARY AND CONCLUSIONS TO PART I

The extensions of contour analysis as it applies to the problem of consumer choice has now been completed and some applications made of them. It is now time to review the results.

First, the major consequence of pushing the analysis of indifference curves to its logical conclusions is the formulation of a complete and general picture of the two dimensional indifference map. The second major result, made possible by a consideration of "unrealistic" areas, is refinement of an existing definition of commodity relations based on ordinal properties of the indifference map. In this field of commodity relations the neo-classical development of demand theory left a great void as far as two dimensional analysis was concerned. A definition to classify the wide variations between indifference maps in terms of prevailing relationships between commodities was sadly needed to render indifference curve analysis more fruitful and useful. The refinement of the definition presented here was only made possible by making use of the full indifference map and by examining the assumptions under which the complete map can be developed. Finally, a third significant result of the extensions lies in the possibility it provides for accounting for the origin of various unusual indifference maps. The analysis of Chapter V indicated that these uncommon maps all derive from the same fundamental experiment.

The applications of the above results to demand theory lead to some useful and practical conclusions. First, the use of the results make it possible to show clearly how a family of isoclines fits into the indifference map. This enables some definite conclusions to be made on the exact shape and path of the income curves, thus clearing up the great confusion which had surrounded this point. The examination of this matter reveals the existence of points of equilibrium where the marginal rate of substitution does not equal the price ratio, and accounts for this phenomenon. Of more importance is the geometrical illustration of the necessary conditions for a negative sloping income curve. From this illustration follows the obvious and simple conclusion that it is only a matter of prices and income which determines the inferiority of a particular good. The analysis of the demand curve was built onto the investigation of the income curve. From this some conclusions were drawn as to the necessary graphical conditions for the Giffen paradox.

The above conclusions drawn from the applications of the extended analysis of indifference curves give some indication of the usefulness of this sort of approach. However, even without such proof the value of the extended analysis might stand on the claim that the goal of a better conceived tool is a worthy end in itself.

PART II

EXTENSIONS OF ISOQUANT ANALYSIS

CHAPTER IX

ASSUMPTIONS AND DEFINITIONS BASIC TO PART II

Existing isoquant analysis can be extended in the same manner as indifference analysis. The purpose of this preliminary chapter is to set out the following:[1] first, postulates basic to an experiment from which existing isoquant analysis can be derived; secondly, definitions useful in the description of that experiment, and useful also in the extension of existing analysis; thirdly, a description of the experiment; and fourthly, the law of variable proportions. Under the third head certain secondary assumptions are set down. Some of these are useful for tabulating the results of the experiment; others are taken from existing analysis; still others are helpful for the extension of the analysis. Under the same head, a method is given for tabulating the results of the experiment, and for deriving from the tabulation the isoquant patterns associated with existing analysis. Again, under the third head, prices are introduced into the technological environment of the experiment. Under the last head, in addition to the law of variable proportions, the assumptions and conditions underlying the law are set down, along with an elasticity concept and proposition useful in applying the law. Again,

---

1 For convenience of reference, the scheme described below is set out in outline form, similar to that used in Chapter I.

under the last head, consideration is given to the geometrical representation of the law.

I - Postulates Basic to the Experiment

A. There is given, by technical considerations, a production function which expresses the maximum amount of some one product obtainable from any given set of inputs, all input units in the set being used.[2]

Comment: The production function relates physical quantities of inputs and output.

B. Some economic unit is associated with the given production function for the purpose of exploiting the productive opportunities represented by it.

Comment: This economic unit may be a head office or controlling unit. It is a composite entity consisting of all the fixed elements required by the production function.

C. A period of time elapses between the application of a set of inputs to production and the emergence of

---

2 The production function may be defined without this last condition attached, in which case input units making the product less that it otherwise would be, need not be used. Marginal productivity is then never negative. Cf. S. Carlson, A Study in the Pure Theory of Production, pp14-16. Also Samuelson, Foundations of Economic Analysis, p58.

However, the entrepreneur can always be induced to use excess input units if the input price can become negative. Owners of input units may very well pay to have them used rather than permit them to remain idle. Besides, input units are only input units if used in a productive process. It is therefore better to define the production function in terms which do not exclude negative marginal productivities, a realistic possibility, and in terms which make the concept of input most meaningful — i.e., in terms of goods and services used in a productive process.

the final product due to that set.[3]

> Comment: All input units in the set may be applied at the same moment to production, or they may be applied individually, or in groups, at different times. The final physical product due to this set may accrue all at once at the end of the production period, or it may accrue over a period of time.
>
> The time period of the productive process is measured from the time the first input unit of the set enters into production to the time that the last increment of product due to the input set accrues. This period of time, of course, varies with the input set considered. The problem of imputation is not considered here.

II -- Definitions

A. The term product qualitatively applies to all amounts of whatever results from a given transformation of any input set.

> Comment: Allunits of the product are homogeneous regardless of the date at which they emerge from the productive process. Two units of the product accruing at different times in the process are physically identical.

B. The term input qualitatively applies to amounts of a good, or service, which, when used in a productive process, either yield some product or enable other inputs to yield a greater amount of product than they otherwise would.[4] Quantitatively, units of an input

---

3 The time element introduced here will be required later for making the analysis of production analogous to the analysis of demand.

4 Any amount of a good or service satisfying at least one of the conditions of the definition is an input. The last condition of the definition can be made clearer by taking an example. Suppose that Y makes no direct and active con-

are units of some amount of a good or service to which the qualitative definition of the term input applies.

> Comment: The term input does not apply to broad composite quantities such as land, labor, and capital but to units — homogeneous with respect to efficiency — of a measurable economic good or service. Only units of productive services which are perfect substitutes for each other are grouped together to form one input.

C. The saturation amount of an input is that amount for which the total product is greater than the product due to any other amount of the input, all other input quantities kept constant.

> Comment: The saturation quantity of any input may, and in general does, vary with changes in the fixed amounts of the other inputs.

III - Description of the Experiment

A. The significant points in the description of the experiment are set down below.

1. The economic unit associated with the given production function is presented with various input sets for the purpose of determining the maxi-

---

tribution to production, but that it enables a given amount of X to yield more product than it does without the aid of Y. Then Y is an input and has the effect of shifting the total product curve of X. If this curve is shifted to the left, without any alteration in its height, then the given amount of X has to lie to the left of the saturation amount for it to yield more product after the shift. For if the given amount of X lies to the right of the saturation amount, then after the shift it yields less product. Y in such a case is not an input. Therefore, whether or not Y is here an input depends upon the quantities of the inputs in the fixed collection.

mum amount of product each input set can produce.

> Comment: Each input set contains only quantities of goods and services to which the term input can be applied. Zero amounts are not excluded from the input sets; only those non-zero amounts to which the qualitative definition of input does not apply are excluded. Also excluded, of course, are all amounts of all inputs which do not appear as variables in the production function.

2. All possible relevant input sets are used in the experiment.

3. Economic considerations such as input and product prices are abstracted from for the time being.

> Comment: Once the experiment is set up and the results tabulated, such economic considerations can be introduced.

4. The technological and organizational environment of the chosen economic unit determines the output each set of inputs will yield.

B. Certain secondary assumptions are useful for tabulating the results of the experiment.

1. The fixed economic unit makes some definite contribution to the productive process.

> Comment: As a result of this assumption the production function is not linear and homogeneous.

2. The economic unit is an indivisible unit completely adaptable in form.

> Comment: Adaptability refers to the ease with which the form of the economic unit can be adapted, qualitatively but not quantitatively, for operations with varying quantities of any

input.[5]

Since the economic unit is indivisible and completely adaptable, the law of variable proportions applies to a proportionate increase in all inputs.[6]

3. Measurement is cardinal so that the output due to different input sets can not only be ranked, but can also be measured absolutely.

<u>Comment</u>: Output uncertainty is nil. For the given productive process it is possible to state, <u>ex ante</u>, that output will be doubled, say, if all inputs are doubled, and that output will be a certain absolute amount. In some productive processes, e.g., agricultural processes, only the first part of the above statement can be made.

4. The inputs are continuously divisible and the productive process continuously variable.

<u>Comment</u>: The continuity assumption is made here for theoretical convenience.

5. The results of the experiment conform to the law of variable proportions.

<u>Comment</u>: This law is accepted here as being useful for the analysis to follow. It may be taken as an established fact, but it is safer to consider it as a secondary assumption not absolutely essential to the analysis.

The law can be assumed to hold here because measurement is taken to be cardinal. It has no counterpart in utility theory because ordinal measurement is assumed there.

C. Two observations can be made on the hyper figure which emerges from plotting the experimental data in hyper space.

---

5 Cf. G. Stigler, "Production and Distribution in the Short Run", <u>Readings in the Theory of Income Distribution</u>, p121.

6 The law of variable proportions is discussed below under section IV of this chapter.

1. Only one hyper figure results from plotting the data in hyper space.

   Comment: This follows from the assumption of cardinal measurement. The hyper figure has one more axis than the number of inputs appearing in the production function as variables.

2. The hyper figure only exists for those quantities of goods and services which are definable as inputs.

   Comment: The figure does not extend indefinitely in all directions but is restricted in size so as to exclude some positive, as well as all negative, quantities of goods and services.

D. The two dimensional isoquant map is derived from the hyper figure by taking a plane through some point on the figure parallel to the plane containing any two input axes, denoted here only as X and Y. This derived map will be identical to the isoquant maps dealt with by the existing analysis if the following restrictions[7] are imposed.

1. The point on the hyper figure through which the cross-sectional plane is taken cannot lie on a zero product contour of the hyper figure.

   Comment: For certain pairs of inputs special cross-sectional patterns arise when the selected point lies on a zero pooduct contour. It is to rule out these special cases,

---

7 These restrictions are relaxed later in Chapter XII below.

for the time being, that this restriction is imposed.

2. The origin of the derived isoquant map cannot be a point in hyper space lying on a zero product contour of the hyper figure.

3. The X and Y inputs cannot be perfectly related to any other inputs unless they are related to each other.

E. The following assumptions apply only to the relevant sections[8] of the derived isoquant maps, these being the sections with which existing analysis deals.

1. All input quantities, other that X and Y, are fixed at the values associated with the point in hyper space representing the origin of the isoquant map. These quantities are treated as a fixed collection completely adaptable in form.[9]

<u>Comment</u>: Since all input quantities must be

---

8 i.e., those sections relevant for realistic analysis.

9 This means that the form of the fixed collection of inputs is adjustable so that the collection can be combined with any input quantities of X and Y. To take an example suggested by Professor Stigler, let the fixed collection of inputs be a fixed quantity of shovels, measured in value terms, and let the variable inputs be represented by the single variable labor. Then, if each unit of labor requires a shovel in the productive process, the form of the shovels must be adjusted as the quantity of labor is increased so that each shovel becomes smaller or less durable while the total value of the shovels remains constant. Cf. "Production and Distribution in the Short Run", <u>op</u>. <u>cit</u>., p121.

used, the fixed collection of inputs also has the characteristic of an indivisible element. Therefore, the law of variable proportions holds when X and Y are varied in proportion.

2. The form of either variable input (X or Y), when temporarily fixed, is completely adaptable to varying quantities of the other input.

Comment: Thus, when Y (or X) is held fixed at some value on the isoquant map, and X (or Y) alone is varied, the fixed amount of Y (or X) is completely adaptable as to its form. This fixed amount is also indivisible since all Y (or X) units in it must be used. Therefore, the law of variable proportions holds when X (or Y) is varied by itself.

3. The iso-product curves have the same properties as the indifference curves with respect to slope and curvature. That is, their first derivative is negative and their second, is positive.

4. Each isoquant is continuous and designates the constant amount of product due to the various amounts of X and Y represented by points along that curve.

Comment: The amount of product which each isoquant denotes is that product due to the combination of various amounts of X and Y with a fixed collection of inputs. The product due to the presence of the fixed elements is ignored, so that for zero X and Y the product is taken to be zero.[10]

---

10 i.e., the product due to X and Y is measured as a deviation from the amount of product due to the fixed elements and not from zero product.

5. The isoquants appear in an ascending order of absolute amounts of product so that the marginal productivity of X, and of Y, is positive.

F. The analysis of the isoquant map can be extended by making certain assumptions about the nature of the production surface for which the isoquants are the product contours.

1. The surface rises monotonically to a maximum height as X alone, or Y alone, is increased from zero.

   Comment: The surface rises in accordance with the law of variable proportions.

2. There are just two ridge lines on the surface. These are formed by the partial maximum heights which exist for different combinations of X and Y.

   Comment: The projections of these ridge lines on to the isoquant map form the loci along which the iso-product curves turn parallel to one of the axes. These ridge lines play the same important role in the present analysis as they did in indifference analysis.

3. Beyond the ridge lines the surface falls monotonically until it reaches X, or Y, values for which total product due to X alone, or to Y alone, is zero.

   Comment: This follows from the condition set out in the first basic postulate that all amounts of an input must be used. Excessive units of an input interfere with the productive process when an attempt is made to use them. That is, their marginal productivity is negative.

4. The surface terminates at those X and Y values which are so great that the total product which can be imputed to X alone, or Y alone, is zero.

> <u>Comment</u>: Further units of X, or Y, then sub$tract from the total product of the fixed elements. Such units are not covered by the definition of an input and so are excluded from the experiment. Therefore, the hyper figure, and so the production surface, cannot exist for X and Y amounts containing such units.[11]

---

11 The exclusion of such amounts is arbitrary as it depends on a definition. The excluded amounts are not, and can not be, used in the productive process so that the product associated with such amounts is, in effect, zero. Therefore, all such amounts are designated by points lying on the XY plane. For these amounts the plane is "empty" as in the corresponding situation of indifference analysis.

If the hyper figure is not arbitrarily restricted, and if the origin of the isoquant map, and corresponding surface, is associated in hyper space with a point denoting zero product, then the production surface falls beyond the ridge lines until it meets the XY plane everywhere. In this case the surface can not extend below the XY plane, so that it is terminated when it meets the plane. Beyond this the XY plane is "empty" in a non-arbitrary sense. The restriction imposed on the hyper figure by the definition of an input has the consequence of preventing the production surface from meeting the XY plane everywhere. It thus adds to the "empty" area. Therefore, in this case, only a portion of the "empty" area depends for its emptiness on a definition, in contrast to the indifference situation where all the emptiness of the plane depends on a definition.

If, in the above unrestricted illustration, the origin of the isoquant map is associated in hyper space with a point denoting some amount of product, then the production surface falls beneath the XY plane beyond the ridge lines. The depth of the surface beneath the XY plane indicates a deduction from, instead of an addition to, the total product denoted by the origin in hyper space. The surface therefore falls beneath the plane until it reaches a depth at all points which represents the negative amount of product just sufficient to balance the positive

G. The introduction of prices, when made on the basis set out below, requires no alteration in the conditions or representation of the experiment or its results.[12]

1. Markets are perfect so that the economic unit can sell any amount of product at any given price and can also obtain all amounts of inputs it requires at any given input prices.

> Comment: The same sort of relationship exists between marginal productivity and input price as between marginal utility and commodity price. That is, if the marginal productivity of an input is positive, its price is positive; if it is negative, its price is negative.
>
> By fixing input prices at all conceivable levels, the experimentalist

---

amount of product denoted by the origin in hyper space. The surface becomes parallel to the XY plane at such a depth and remains parallel to it for all further X and Y values up to infinity. Some surface then lies either over or under the XY plane so that there is no "empty" area.

It is desirable to restrict the hyper figure so that quantities of goods and services are excluded which are so large that output would be greater if none of the goods or services were used at all. An economic unit will only utilize an amount of a good or service when that amount yields more product than the zero amount (i.e., more than zero product). It would never utilize an amount yielding less product than the zero amount. The restriction may be arbitrary but it is meaningful.

12 When economic considerations are introduced attention must be given to the problem of selecting the most relevant cross-section of the hyper figure for interpretation of events in the real world. This cross-section is obtained by taking prices as given, at any given moment, in the markets and permitting these to determine the equilibrium quantities for inputs. These quantities fix a point on the hyper figure through which the cross-sectional plane should be taken. The plane may be parallel to any two input axes.

can cause the economic unit to select and use all the various input sets, each set being used in a particular price situation.

2. In any input set those inputs whose amounts are so excessive that their marginal productivity is negative pay to be employed in the productive process.

3. The economic unit selects the output which maximizes revenue for the given product price, and then selects that set of inputs which produces that output for the least cost at the given input prices.

IV - The Law of Variable Proportions

A. The law[13] may be stated[14] as follows: If, without change in methods of production, successive physical units of an input, or given combination of inputs, are added to a fixed collection of other inputs, the total physical product obtained varies in magnitude through three phases:

1. During the first phase, total product always

---

13 The law is based on empirical evidence. That is, "it is a universal phenomenon that no combination of two elements of production can be conceived to which the principal of diminishing physical outputs does not apply." (J.D. Black, Introduction to Production Economics, p286)
Acceptance of the law adds to the preciseness of the tabulation of the experimental data. Although this is useful for the analysis it is not essential.

14 This statement of the law is a condensation of Professor Cassels' statement of it. Cf. "On the Law of Variable Proportions", Readings in the Theory of Income

increases at a percentage rate greater than the rate of increase in the varible input.

Comment: The end of this phase is marked by the point where the rate of increase in the product is equal to the rate of increase in the variable input.

2. During the second phase, the total product always increases at a percentage rate less than the rate of increase in the variable input.

Comment: The end of this phase is marked by the point where total product reaches a maximum.

3. During the third phase, total product decreases absolutely until it is reduced to zero.

B. Four conditions underly the above statement of the law. It will be noticed that the first two involve assumptions already set down above.

1. The variable input can be broken up into small separable homogeneous units.

2. The fixed collection of inputs is an indivisible collection completely adaptable in form.

3. The context of the law is strictly static and purely physical.

4. The third phase of the law is the converse of the first.[15]

Distribution, pp107-8. For a more summarily statement cf. F.H. Knight, Risk, Uncertainty and Profit, pp98-9.

15 i.e., the first and third phases are complementary. Cf. Stigler, The Theory of Prices, p125.

Comment: In the third phase, the proportion of the variable input in relation to the fixed elements is so excessive as to be harmful. In the first phase, the proportion of the fixed elements in relation to the variable is also so excessive that it is harmful. Therefore, the only economically relevant phase described by the law is the second.[16,17]

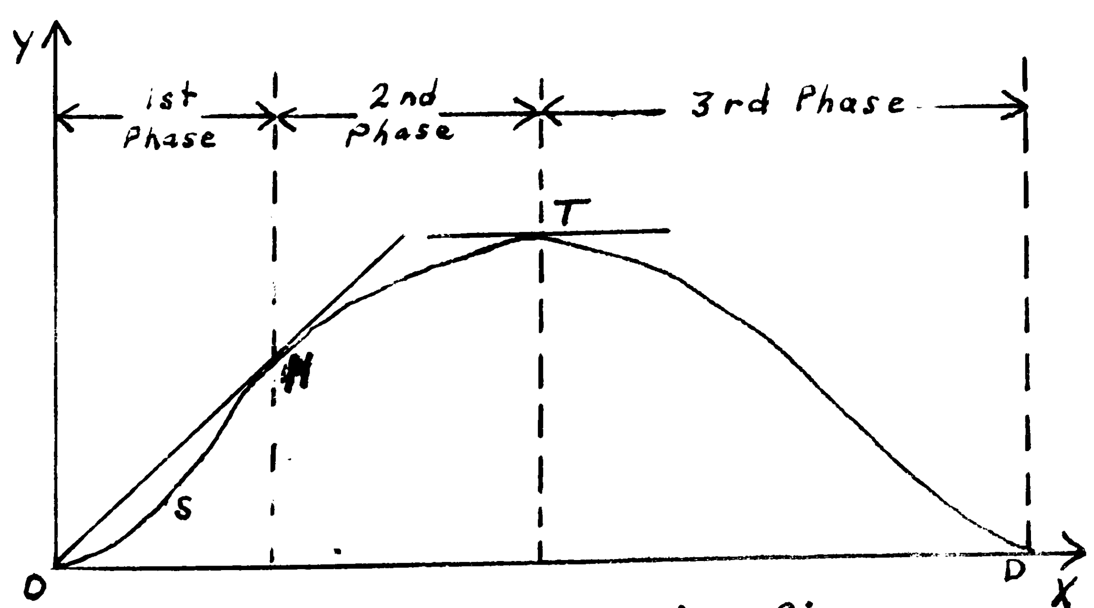
C. The law is reflected in the shape of the total product curve for a variable input when some fixed elements are present.[18] The significant points in

---

16 It is the phase within which diminishing average and marginal returns prevail. Cf. Cassels, "On the Law of Variable Proportions", op. cit., pp107-8.
Due to the fact that only the second phase is economically relevant, two separate and distinct redundant areas appear on the isoquant map. One of these exists next to the axes and the other, beyond the ridge lines.

17 The "law" is "merely a way of expressing the fact that the inputs in a process are not perfectly substitutable one for the other." (Boulding, Economic Analysis, 1st ed, p489) And "it is the fact that some proportions of inputs are better than others which gives rise to the law." (Ibid., p491)

18 There is no graphical illustration of the law which is universally applicable. The graphical hypothesis most widely accepted is that put forward by Professor Knight. (Stigler, The Theory of Prices, p123) A refinement of Knight's graphical illustration, adopted from Cassels, "On the Law of Variable Proportions", op. cit., p108, is presented in the accompanying figure.



In this figure the Y axis represents product and the X axis, some variable input. The point S is a point of inflection denoting where marginal productivity for the input is a maximum. At N average product is a maximum, while at T total product is maximized. The first phase described by the law holds between the origin and N, the second phase,, connection with such a curve are noted below.

1. A total product curve can be derived from the hyper figure by taking any plane through the origin of the figure containing the product axis, or by taking a plane through any point on the figure, but the origin, parallel to a plane containing the product axis and some one input axis.

   Comment: The usual type of total product curve is derived in the latter manner. If derived in the former manner, the curve may be for some variable combination of two, or more, inputs.

2. The total product curve illustrates how total product varies when one input, or combination of inputs, is varied, all other input quantities held fixed.

   Comment: The exact shape of the curve depends on the location of the cross-sectional plane from which it derives. However, in all cases, the total product curve rises monotonically to a maximum and then falls off monotonically "but in some cases it will rise rapidly and fall away rapidly; in others, rise rapidly, remain nearly flat for a long time, and then fall away slowly; in others rise somewhat more slowly and decline either slowly or rapidly, *etc*."[19]

---

between N and T, and the third phase between T and D. The third phase is more prolonged than the first since the proportions between the fixed inputs and the variable input are less affected by each additional unit of the variable input. That is, the excessiveness of the variable input is not as harmful to production as is the excessiveness of the fixed inputs. (Cassels, "On the Law of Variable Proportions", *op. cit.*, pp108-9)

19 Black, *op. cit.*, p286.

3. Variations in total product, due to variations in the variable input, are measured from the amount of product due to all the fixed input quantities, and not from zero product.

Comment: When the variable input is zero the total product curve cuts the input axis.

4. The total product curve, in general, cuts the product axis at the zero mark.

Comment: The total product curve starts at the origin when this is adjusted so as to eliminate the product due to the fixed inputs.[20]

D. In distinguishing between the different phases of product variation described by the law, an elasticity concept, the function coefficient,[21] is of great use.

1. The function coefficient is given by the ratio of the percentage change in product to the

---

20 Professor Cassels argues that the total product curve does not start at the origin "because before the ratio of the excessive factor to the deficient one becomes infinite (as it is at the origin) the product is reduced to zero". ("On the Law of Variable Proportions", op. cit., p109) This statement is incomprehensible to the writer as is Professor Knight's claim that the curve cannot start from the origin for if it did it would later have to become asymptotic to the X axis instead of cutting it, which it clearly must do for a sufficiently large input amount. (Cf. Risk, Uncertainty and Profit, p102n)
From the analysis presented in the following chapters it is obvious that in all but a few special cases the total product curve must start from the origin.

21 Carlson, op. cit., pp22-3.

percentage change in the variable input which brings about the change in product.[22]

> <u>Comment</u>: During the first phase of the law the function coefficient of the variable input is greater than one. During the second phase it is less than one. During the third phase it is negative. Thus, the coefficient is one when average product for the variable input is maximized, and it is zero when total product is a maximum.

E. At any point on the hyper figure the sum of the function coefficients of all the inputs at that point gives the total elasticity of production at that point, where this latter term is defined as the ratio of the change in product to the proportionate change in all inputs responsible for the given variation in total product. This may be set out as follows:

1. $\frac{\lambda}{p}\frac{dp}{d\lambda} = \sum_{i=1}^{n} \frac{\partial p}{\partial x_i}\frac{x_i}{p}$, where p represents amounts of product, x_i, amounts of the <u>i</u>th input, λ, the factor by which all inputs are increased ($\lambda > 1$) in a fixed proportion from a given basic position,[23]

---

22 This is the elasticity of production in a restricted sense. The elasticity of production is calculated by varying all the inputs in the same fixed proportion, whereas the function coefficient is calculated by varying just one input, all others held constant. Therefore, the function coefficient is really a partial elasticity of production. The sum of the function coefficients of all the inputs gives the total elasticity of production for a proportionate increase in all inputs. For a fuller discussion of the elasticity of production cf. Allen, <u>Mathematical Analysis</u>, p263.

23 i.e., output is a function of λ when the inputs are used in a fixed proportion.

and n, any arbitrary number denoting the number of inputs used in the productive process.[24]

> Comment: The left hand term of the expression is the total elasticity of production and the right hand term is the function coefficient of the input x_i.

2. $\frac{x_1}{dx_1}\frac{dp}{p}=\frac{x_2}{dx_2}\frac{dp}{p}=\sum_{i=1}^{2}\frac{\partial p}{\partial x_i}\frac{x_i}{p}$ is a special case of 1, applicable to isoquant analysis.[25]

> Comment: In the two dimensional case, the sum of the function coefficients of the two variable inputs gives the function coefficient of the compound homogeneous input resulting from varying the two variable inputs together always in the same fixed proportion.

---

24 The proof is as follows: Let the production function be $p=f(x_1, x_2, \ldots, x_n)$, where n is any arbitrary number. Now $dp=\sum_{i=1}^{n}\frac{\partial p}{\partial x_i}dx_i$. Since all inputs are increased in the same proportion, $\frac{dx_i}{x_i}=\frac{d\lambda}{\lambda}$, where $i=1, 2, \ldots, n$. Therefore, $dx_i=x_i\frac{d\lambda}{\lambda}$. By substitution, $dp=\sum_{i=1}^{n}\frac{\partial p}{\partial x_i}x_i\frac{d\lambda}{\lambda}$ or

$$\frac{dp}{d\lambda}\frac{\lambda}{p}=\sum_{i=1}^{n}\frac{\partial p}{\partial x_i}\frac{x_i}{p}. \quad \text{Q.E.D.}$$

25 The proof is as follows: $dx_i=0$, where $i=3, 4, \ldots, n$. Therefore, $dp=\frac{\partial p}{\partial x_1}\cdot dx_1+\frac{\partial p}{\partial x_2}\cdot dx_2$. When x_1 and x_2 are increased in the same proportion, $\frac{dx_1}{x_1}=\frac{dx_2}{x_2}$. Therefore,

$$dp=\frac{\partial p}{\partial x_1}\cdot\frac{x_1}{x_2}\cdot dx_2+\frac{\partial p}{\partial x_2}\cdot dx_2, \text{ or } \frac{x_2}{dx_2}\cdot\frac{dp}{p}=\frac{\partial p}{\partial x_1}\cdot\frac{x_1}{p}+\frac{\partial p}{\partial x_2}\cdot\frac{x_2}{p}=\frac{x_1}{dx_1}\cdot\frac{dp}{p}.$$

Q.E.D.

With the aid of the above, the analysis of the next two chapters makes use of isoquant patterns familar in existing analysis to extend that analysis. It will be recalled that to derive such isoquant patterns from the hyper figure certain restrictions must be imposed in taking the cross-sectional plane of the figure. These restrictions prevent the origin of the isoquant map from being any point in hyper space lying on a zero product contour of the hyper figure. They also prevent the X and Y commodities from being perfectly related to any other inputs but each other. The type of pattern derived depends on the selection of the axes and on the location of the origin of the pattern in hyper space. By dropping the above restrictions, some special isoquant maps arise which are not considered in existing analysis. These are discussed in Chapter XII. Chapter XIII is a digression on the isoquant patterns derived from the contour analysis of linear homogeneous production functions. Chapter XIV considers the applications which may stem from the extension of existing analysis.

CHAPTER X

EXTENSION OF ANALYSIS FOR INDEPENDENT INPUTS, PERFECT COMPLEMENTS AND PERFECT SUBSTITUTES

In this chapter three selections of the X and Y axes are considered in obtaining special plane cross-sections of the hyper figure. These special cross-sections yield some simplified isoquant patterns which may be used to begin the extension of the analysis. In the first selection inputs are chosen which are independent; in the next selection inputs are chosen which are complements; and in the last, inputs, which are perfect substitutes.

I - Isoquant Maps for Independent Inputs.

Consider an equilibrium point on the hyper production figure lying on a non-zero product contour. Suppose that there can be found a large number of other points on the figure which differ from the given point only by the amounts of X associated with them. Suppose further that planes taken through all these points parallel to one another and to the plane containing the product and Y input axes yield total product curves of identical shapes. All these total product curves attain their maximum height for a single Y value.[1] If a set of points can be found meeting

---

1 Along these curves maximum average product occurs for the same Y value on all of them. Also, they all indicate three phases of variation in total product, each phase holding for the same range of Y values on all the curves. However, the curves may differ from each other with respect to height.

these conditions with respect to X, then another set of points can be found on the hyper figure which meet corresponding conditions with respect to Y. The existence of such points can be taken to define the two inputs, X and Y, as being independent.

The isoquant map derived for two such inputs represents the contours of a production surface which is single-peaked. There is therefore a single point on the map representing a maximum total product, at which the ridge lines must intersect. Now since maximum product for X occurs always for the same X value independently of the value of Y, the X ridge line is a straight line parallel to the Y axis. Similarly, the Y ridge line is parallel to the X axis. These ridge lines can be represented by AMN and BML in Figure 29, overleaf, with M being the point of maximum product. The contour curves which do not cut these ridge lines must cut the axes, since X and Y are not indispensable for production. Thus, Figure 29 is similar to Figure 3 representing the indifference map for two independent commodities. The isoquant map, however, because of the law of variable proportions, contains additional loci which can not be found on the indifference map.

Consider further Figure 29 which represents the isoquant map for two independent inputs when all other input amounts are held fixed.[2] Moving along the Y axis from the origin, the total product ascribable to Y varies in the manner described by the law of variable proportions. Average product is maximized at some point such as F,

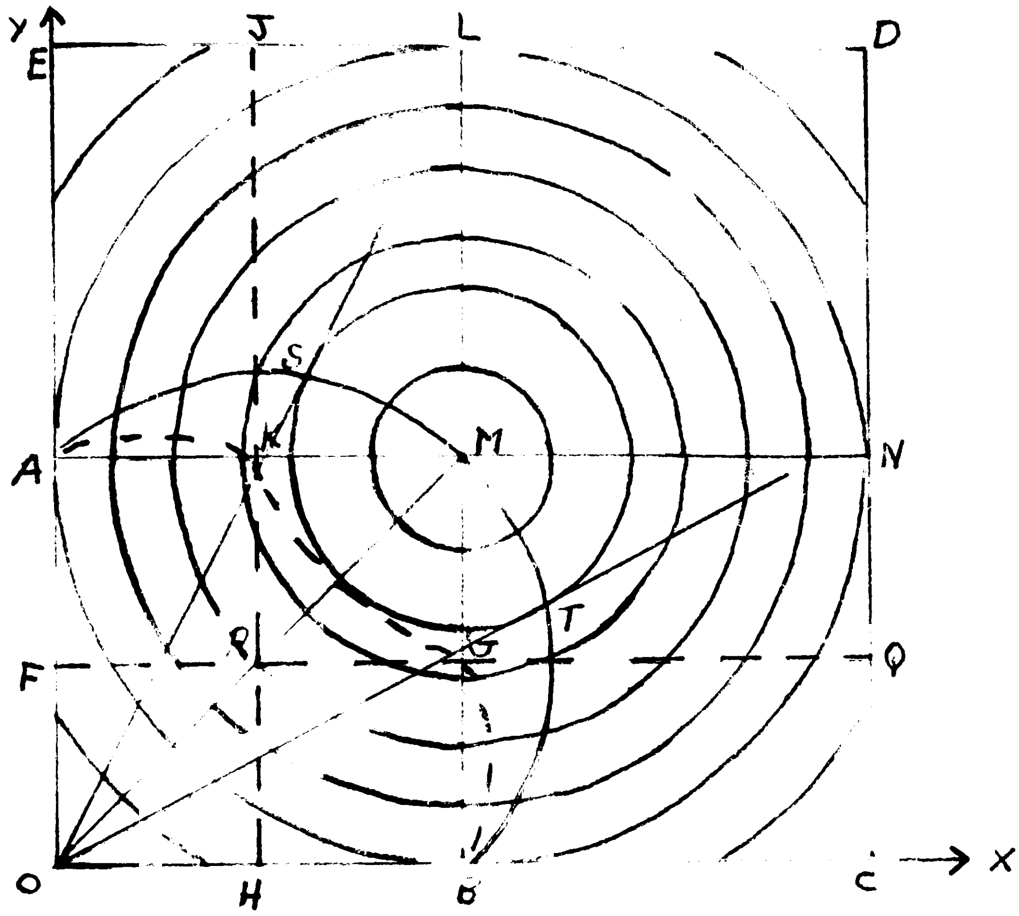


Fig. 29. - Isoquants for independent inputs.

and total product, at the point A. At E the total product of Y is reduced to zero again. It follows from the definition of independence used here that average product for Y is maximized at the same Y value independently of the value of X. The total product for Y also falls to zero again for the same Y value regardless of the value of X. Therefore

---

2 Figure 29, with minor additions and slight changes, is a reproduction of Professor Cassels' Figure 3, "On the Law of Variable Proportions", op. cit., p114. It differs radically from Dr. Carlson's diagrams, op. cit., Ch II, representing the product contours for two variable inputs and one fixed factor.

FGQ is the locus of maximum average product for Y along which the function coefficient of Y (E_y) is unity. The line ELD is the locus of zero total product for Y. It therefore forms the northern boundary of the map.[3] All the total product curves for Y, obtainable from the production surface associated with Figure 29, start from points on the X axis,[4] and end at points on ELD. For the range of Y values cut off by the X axis and FGQ the total product curves show increasing returns and E_y is greater than one. Between FGQ and AMN E_y is positive but less than one, and beyond AMN it is negative.

Similarly, when X alone is varied, Y constant, JKH is the locus of all points at which the function coefficient of X, E_x, is unity. The line BML is the locus of all points where E_x is zero. The line CND is the locus at which the total product curves of X terminate. This line then forms the eastern boundary of the map. These three loci along with the Y axis divide the map, with respect to X, into three sections: one section containing all points where E_x is greater than one; one, where it is less than one but positive; and one, where it is less than zero.

Now when X and Y are varied together along a straight line through the origin, such as OK, the effect on output

---

3 Cf. p125 above for the discussion on the termination of the production surface.

4 i.e., for zero Y values.

is the same as if units of a compound input, X-Y, (the proportions of X and Y in each unit being constant) were being combined with the fixed collection of inputs which were fixed in deriving the map. Therefore, as the amount of the compound input is increased from zero, total product varies according to the law of variable proportions.[5] On each straight line, or diagonal, through the origin there is a point where average product reaches a maximum for the compound input, and a point where total product for the compound input becomes a maximum.

The locus of all points on the diagonals at which total product is maximized is obtained by joining all points of tangency between the isoquants and the diagonals. The result is some such curve as the one in Figure 29 passing through the points A, S, M, T, and B.[6] The locus of maximum average product for variation of inputs along the diagonals must lie entirely inside of this curve, since the point of maximum total product always comes after the point of maximum average product. At points on this latter locus the function coefficient of the compound input is one. Therefore the areas in which the locus lies can be determined by considering the individual function coefficients since the sum of the X and Y coefficients gives the

---

5 Cf. Cassels, "On the Law of Variable Proportions", _op_. _cit_., p115.

6 Actually this locus only approaches the points A and B as a limit since it can not intersect the axes. Cf. note 8 below.

function coefficient of the compound input.[7] Thus the locus of maximum average product for variation in the compound inputs along the diagonals is also the locus of all points where the sum of E_x and E_y is one. The locus must then pass through the points K and G since at these points one of the coefficients is zero while the other is one. It can not pass through the area OFRH because everywhere there the sum of the coefficients is greater than one. Neither can it lie in the areas FAKR or HRGB since one coefficient is greater than one while the other is always positive.[8] In the regions SKM and MGT one coefficient is less than one while the other is negative. By the process of elimination, it follows that the locus must pass through the areas ASK, GBT and RKMG. In the first two areas one coefficient is greater than unity while the other is negative so that it is possible for their sum to be one. Within RKMG both coefficients are positive but less than one.

It is now possible to form a general idea of the path

7 Cf. p133 above.

8 At the zero amount of an input the function coefficient of that input can not be calculated because the first derivative of the total product curve does not exist at zero. That is, the total product curve begins at the origin so that it has a right derivative but no left derivative. Both must exist for the curve to have a derivative at zero. When this derivative does not exist the function coefficient can not be calculated. Therefore any loci involving the function coefficients can not cut the axes.

traced out by this locus of maximum average product for input variations along the diagonals. A suggested path is given by the broken curve passing through the points K and G. It approaches A and B as terminal points,[9] and it is probably convexed to the origin in the area RKMG, the convexity being less than that of the isoquants which it intersects.[10]

Therefore, the area KMG in Figure 29 is the economically relevant area of the map for perfect competition. For quantities of inputs beyond the ridge lines one marginal productivity is always negative. In the area KRG the marginal productivities of both inputs are positive and diminishing, but returns to "scale" are increasing,[11] this latter condition being incompatible with perfect competition. Only within the area KMG are returns to "scale" decreasing at the same time that the marginal productivities of both inputs are positive. This

---

9 The locus can only cut the ridge lines at the points K and G.

10 In general, the locus will not coincide with an isoquant within this area because the value of the average product varies for each diagonal. On one of these diagonals, probably the one passing through the origin and M, the maximum average product should be greater than the maximum average product on any of the other diagonals. The locus, to reflect this, has to be less convexed to the origin than the isoquants.

11 i.e., returns along the diagonals where both inputs are varied together in a fixed proportion.

makes it the effective region of the isoquant map which is more restricted in area than the effective region of the indifference map.[12]

III - Isoquant Map for Perfect Complements.

Consider now a selection of the input axes such that X and Y are perfect complements in the sense that the effective range of substitution between the two inputs is zero. For two such inputs the plane cross-section of the hyper figure yields an isoquant pattern such as the one represented in Figure 30, overleaf. This pattern is similar to the indifference pattern for perfect complements. The X and Y ridge lines coincide to form the straight line OD. As X and Y are varied along this diagonal in a fixed proportion, total product varies in the manner described by the law of variable proportions. There is some point R on this diagonal where average product for the variable compound input is a maximum. Of course, at M total product is a maximum while at D, as at O, it is zero. Therefore, the isoquant passing through the points O and D is an isoquant of zero product for X

12 The area is more restricted in the case of production for two reasons. One of these is the law of variable proportions which confines the effective region to the area RKMG in Figure 29. The second reason is the assumption, usually made in the analysis of production theory, that the economic unit can purchase inputs without limit, there being no effective money barrier to its operations. However, if such a barrier does exist then RKMG, and not just KMG, is the effective region.

and Y. It thus forms the boundary of the map. Quantities of X and Y beyond this boundary are excluded. Such quantities add nothing to total product but actually lessen the total product yielded by the fixed collection of inputs associated in hyper space with the origin of the map.[13] With the definition of inputs used here such

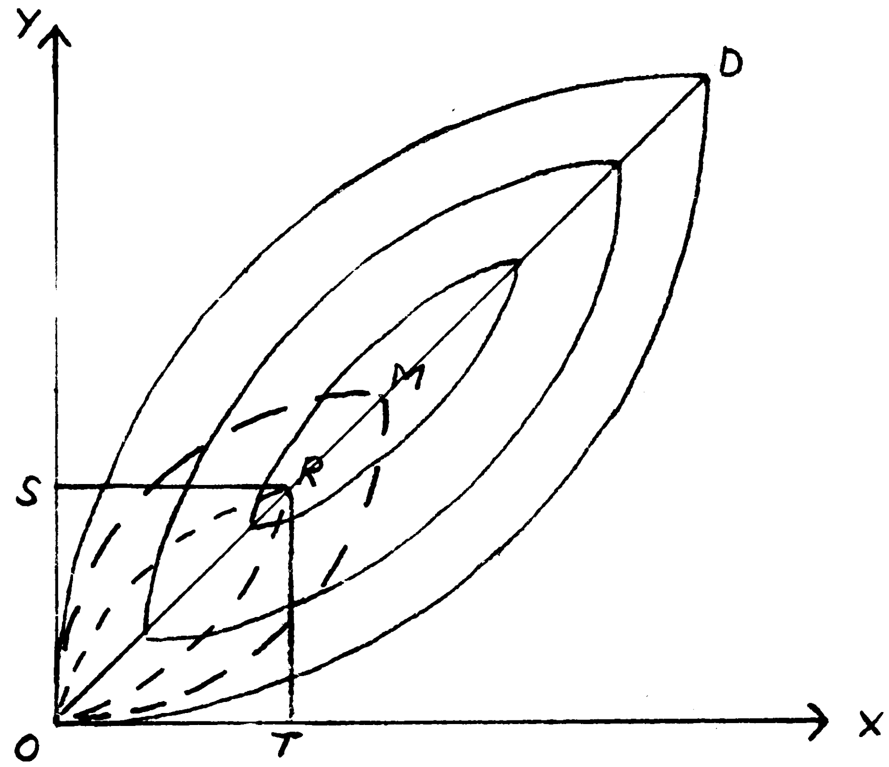


Fig. 30. - Perfect complements.

quantities can not be classed as inputs and so must be excluded.

---

13 Since it is assumed here that the fixed collection of inputs yields some total product, contours would exist in the excluded area if the hyper production figure was not restricted so as to be non-existent for points in that area. Such contours would have negative numbers attached to them signifying a deduction from, in contrast to an addition to, the total product of the fixed collection. Only when the fixed collection yields no total product is the "emptiness" of the excluded area in Figure 30 independent of the restrictions imposed on the hyper figure.

In the situation depicted in Figure 30, the exclusion of negatively numbered isoquants is arbitrary as it is based on a definition of terms. However, the definition and consequent exclusion is meaningful and relevant enough.

Along any diagonal through the origin the law of variable proportions applies so that the map can be divided into regions where returns along the diagonals are increasing absolutely and relatively, increasing absolutely but decreasing relatively, and decreasing absolutely. The loci of maximum total and average product for input variations along these diagonals are the broken curves passing through the origin and the points M and R, respectively. These loci are located in the same manner as the corresponding loci in the situation of independent inputs.

For variations in X (Y) alone, Y (X) fixed, the variation in total product does not correspond to the variation described by the law of variable proportions. This observation follows from the attempt to apply the proposition on elasticity sums to points on the diagonal OD. Consider the segment OR. The sum of the function coefficients, if these exist, at any point on this segment must be greater than one. However, if the total product curves of X and Y take on the form associated with the "law", then both the X and Y function coefficients would be zero at all points on OR. The sum of the coefficients would then be zero along OR. Therefore, the total product curves of X and Y can not be of the usual form but must have kinks, or be discontinuous, at the ridge lines. The function coefficients then do not exist at the ridge lines,[14] except at the point M. Only along a straight line passing through M and parallel to either axis does the law of variable proportions apply.[15] Along all other straight lines parallel to either axis, the "law" does not apply. All total product curves associated with such lines have kinks. All that can be said about these kinked total product curves is that those lying above the area OSRT must show increasing returns, both absolutely and relatively, from zero product up to maximum product.[16]

Along the segment RM, the function coefficient of the compound input is positive but less than one, while beyond M it is less than zero. Since to the left of RM the marginal productivity of Y is negative, and to the right, that of X is negative, the effective region of this map is the segment RM.

---

14 The first derivative of the total product curves does not exist at the ridge lines. Without a value for this derivative the function coefficients can not be calculated.

15 The total product curve associated with such a straight line does not begin at the zero value for the variable input. Rather it starts at some positive non-zero value for the input.

16 Within ORT E_x is negative and hence E_y must be greater than one in order that the sum of the two can be greater than one within the area marked off by the broken curve OR. The same reasoning applies to the area OSR.

III - Isoquant Map for Perfect Substitutes.

If two physically distinct inputs can be found which are perfect substitutes for each other, select their axes as the X and Y axes for the purpose of taking a plane cross-section of the hyper production figure. The resulting iso-

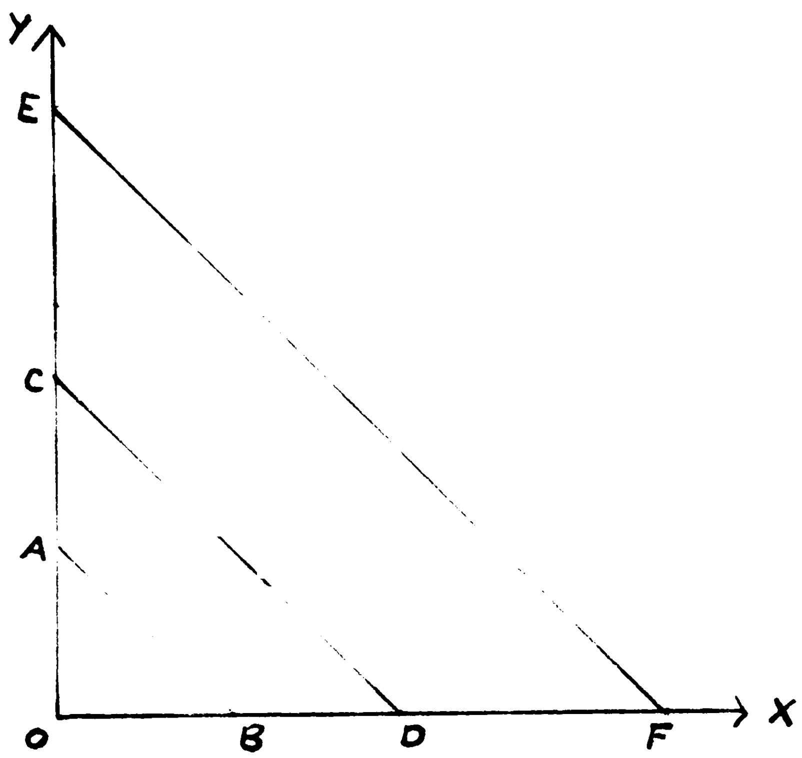


Fig. 31. - Perfect Substitutes.

quant map will have the general appearance of the one drawn in Figure 31. For variations in Y along the Y axis, average product is a maximum at some point A, and total product, at C. B and D are similar points for variations in X along the X axis. EF is an isoquant of zero total product for X and Y and thus forms one of the three boundaries of the map. The other two boundaries are, of course, the axes.

The line CD is an isoquant of maximum total product for X and Y. As such it is the locus of maximum total product for variations in X and Y along diagonals through the origin. CD also represents the two ridge lines of the map. The locus of maximum average product for input variations along the diagonals is the isoquant AB, which is also the loci of maximum average product for X (Y fixed) and for Y (X fixed). Since AB represents these three loci, the individual function coefficients do not exist at points on AB. If they did exist at such points each coefficient would equal one along AB, and the sum of the two coefficients would be two. Yet AB is defined to be the locus of all points where the sum of the coefficients, if it exists, equals one. Therefore, the function coefficients can not exist for points on AB. The total product curves for X and Y associated with the area OAB have kinks at the X and Y values along AB. The law of variable proportions, as a result, only holds when the inputs are varied along diagonals through the origin, including the axes. The total product curves for X and Y which begin at points lying on BF and AE never show a phase of increasing returns. The effective region of the map is thus ABDC.

CHAPTER XI

DEFINITIONS OF INPUT RELATIONS AND EXTENSION OF ANALYSIS FOR PARTIALLY RELATED INPUTS

In the last chapter three special plane cross-sections corresponding to three critical input relations were considered. In this chapter plane cross-sections corresponding to more general input relations are considered. In the determination of these cross-sections axes are selected so that partial substitutes or complements are represented. In order to make such a selection some definition of input relations must first be decided upon. Clearly, the definition must be in terms of some property of the isoquants, or associated production surface, which alters in a calculable manner for different selections of the commodity axes. This problem is approached by considering various possible definitions[1] with a view to using the most appropriate of them for the derivation of a typical isoquant map from the hyper figure. The later sections of the chapter are given to an examination of this map.

I - Definitions of Input Relations

The classical treatment of commodity relations suggests the possibility of using the sign of the first cross-derivative[2] of the production function to classify input relations. If this cross-derivative is zero everywhere, then a change in the quan-

---

1 The only attempts to define input relations in the presence of fixed factors with which the writer is familiar appear in Hicks, Value and Capital, Ch VII; Black, *op. cit.*, Chs. XI-XII; and Cassels, "On the Law of Variable Proportions", *op. cit.*, pp111-113. The latter two attempts represent incomplete formulations.

tity of Y (or X) does not affect the marginal productivity of X (or Y). The inputs can then be labeled independent. If the cross-derivative is positive everywhere, then an increase in the amount of Y (or X) raises the marginal productivity of X (or Y) and the inputs can be termed complementary. If the cross-derivative is negative everywhere then the inputs can be called substitutes. This provides a good criterion as long as the cross-derivative has the same sign for all combinations of X and Y having positive marginal productivities. However, such constancy in sign is a special case only, for it is possible for an increase in X (or Y) to affect the marginal productivity of Y (or X) in an irregular manner. Then the sign of the cross-derivative varies for different combinations of X and Y, yet in such a way as not to upset the law of variable proportions.[3] Thus, X and Y may be complementary at one point on their isoquant map and substitutes at another if these terms are defined according to this criterion. For this reason the criterion is unsatisfactory.

---

2 i.e., the sign of $\partial p/\partial x \partial y$, where p represents product.

3 A geometrical interpretation of the cross-derivative is given by considering two marginal productivity curves for X (or Y) when Y (or X) is held fixed at two slightly different values. Thus, if for the higher fixed value, the corresponding marginal curve lies everywhere above the first, then the sign of the cross-derivative is everywhere positive. However, it is possible for this marginal curve to lie above the first curve at some points, below it at others, and to co-incide with it at still others while still retaining a shape compatible with the law of variable proportions.

The Hicksian definition of commodity relations can be applied to technical relations. That is, output can be treated as a constant and the relation between inputs defined in terms of what occurs when the inputs are varied in such a way as to keep output constant.[4] On the isoquant map the only relationship possible within the effective region would then be substitution. The degree of substitution would of course vary from zero to infinity depending on the inputs considered. This approach is as unsatisfactory here as it is in indifference curve analysis.

Hicks himself does not apply his definition of commodity complementarity to input relations. Instead, he develops a definition of technical complementarity by treating output as a variable in contrast to his method of treating utility as a constant.[5] It is an empirical definition based on the nature of demand for inputs. Although the nature of this demand depends upon the isoquant pattern the definition can not be geometrically incorporated into two dimensional isoquant analysis because there are now three unknowns instead of just two.[6]

---

4 This is apparently the definition Professor Keirstead adopts when he writes: "These two factors...may be technically related to one another in one of three ways. They may be absolutely complementary or non-substitutable, they may be perfectly substitutable, or they may be partly or imperfectly substitutable." (op. cit., pp210-211)

5 Cf. Value and Capital, Ch. VII.

6 That is, the output that will maximize revenue is unknown as well as the amounts of X and Y which will produce that output at the lowest cost. To solve for these three unknowns three dimensions are required. Two dimensional analysis cannot give the complete solution.

The ridge line definition of utility theory can be used here to classify input relations. To judge its relevance consider a simple illustration.[7] Take the total product curve arising from the application of seed to a fixed amount of land. Then suppose scarecrows are added as a second variable input and that in this way the amount of seed lost to birds is reduced by a certain percentage depending on the number of scarecrows used. If, for a given number of scarecrows, this percentage is 10%, then 9 bushels of seed will now have the same effect that 10 had before. The total product curve is therefore compressed horizontally to nine-tenths its former length. The point of maximum total product is thus shifted to the left.[8] This means that the addition of scarecrows has the effect of enabling seed to combine more efficiently,but less intensely, with land.[9] If the point of maximum total product is shifted to the right it means that seed can be applied more intensively to the land.

---

7 This illustration is summarized from Cassels, "On the Law of Variable Proportions", op. cit., p112.

8 When the point of maximum product is shifted to the left, Professor Cassels calls the scarecrows "augmentative" to the variable input of seed and "attenuative" to the fixed factor, land. (ibid., pp112-3) But there is no need to introduce a new terminology here for this phenomenon. The old terms of complementarity and substitution can be used if they are defined in the same manner as in utility theory. There will be no confusion then in using the same terms in both places.

9 i.e., less seed now has to be planted to obtain any given product.

In general, for a given fixed collection of inputs, the addition of Y may either enable X to combine more intensively with the fixed elements, or less intensively; or it may have no effect on X's capacity to combine with the other inputs. These results are reversable. That is, additions of X affect the capacity of Y in the same manner as additions of Y affect X.[10] Of course, the slopes of the ridge lines reflect these results. If the slopes are positive it signifies that the input relation is one which enables X and Y to work more intensively with any fixed elements. This relation can then be defined as complementarity, after the corresponding relation in utility theory. If the ridge line slopes are negative, it signifies that the input relation is one which restricts the intensity with which X and Y may be worked with the fixed elements. The relation can then be defined as substitution.

This way of defining input relations is not an abstract one divorced completely from reality. If a fairly large increase in output, say 50 per cent, can only be brought about by utilizing more of both inputs,[11] then, on a common sense basis, the inputs would be defined as complements. This condition is reflected on the isoquant map by the spread between the ridge lines. The greater the positive slopes of both ridge lines the smaller the spread between them. Thus the

---

10 Cf. Black, <u>op</u>. <u>cit</u>., p306.

11 i.e., output could not be increased 50 per cent by using more of just one input.

inputs are more likely to be complements on this common sense basis. If, on the other hand, output could be increased by 50 per cent by merely increasing the amount used of either input, the amount of the other input held constant, then the inputs are substitutes. This is reflected by a large spread between the ridge lines. They will only be so spread if they have negative slopes. Thus, the slope of the ridge lines reflect in a common sense manner the effects of input relations and so provide a suitable criterion for classifying such relations.

Consider now what is likely to be the general shape of these ridge lines when the input relation is imperfect. They might be expected to behave in the same way as the utility ridge lines since, under the given assumptions, the two cases are more or less similar. This is in fact the case, but it is possible to be more precise here about the ridge line shapes because of the law of variable proportions. This "law" has some bearing on the form of the ridge lines because "the effect of the varying element of production on the output is influenced by the relation between the two fixed elements."[12] Let us consider some illustrations.

First, take the case of land, seed and fertilizer. "If an amount of fertilizer were applied to the land beyond the point of dininishing total outputs for fertilizer applied to land, the...[curve of total output for seed input]...would be greatly affected thereby. An amount of fertilizer near to

12 Black, op. cit., p296.

the points of diminishing total outputs for fertilizer will give an appreciably different set of curves for varying seed inputs than any amount of fertilizer less than this."[13] Again, consider the effect of water upon the effectiveness of a variable input applied to growing crops. "The curves of outputs for varying amounts of seed or fertilizer or labor applied to a growing crop will be higher, and reach total diminishing outputs later, in a moderately wet year than in a dry year.... If, on the other hand the year is too wet... the curves will all be lower, and reach...the point of maximum product for the variable input...sooner than if the moisture supply had been moderate."[14]

The above illustrations suggest a curvilinear form for the ridge lines, with the sign of the ridge line slopes changing with the sign of the marginal productivities of the inputs. It is postulated here as being highly probable that for the range of, say, Y values for which the marginal productivity of Y is increasing, each point on the ridge line of X is shifted to the left, or to the right, by a greater marginal amount than the preceding point as Y is increased. Then, when the marginal productivity of Y begins to fall, each higher point on the X ridge line is shifted less, in a marginal sense, than the point just below it. Finally, when the marginal productivity of Y reaches zero, Y loses its ability to shift the X ridge line. When its marginal productivity becomes

---

13 Ibid., p297.

14 Ibid.

negative, its effect on the X ridge line is reversed so that the direction of shift changes. The input X would have a similar effect on the ridge line of Y.

Thus, in the case of a positive sloped ridge line, defining a complementary relation, the shape of the locus would at first indicate an increasing degree of complementarity, then a constant degree at some point, and beyond that a decreasing degree up to the point of maximum product. Beyond this the relationship would change to one of substitution. In the case of substitutes, there is the same variation in the strength of the relationship within the effective region and the reversal of the relation in the redundant areas.[15]

The location of the origin of the isoquant map in hyper space, as well as the axes selected, has a bearing on the shape of the ridge lines. As that point varies the relationship between the two inputs, as reflected in the slopes of the ridge lines, may very well change.[16] That is, for one fixed collection of inputs, X and Y may be partial substitutes while for a totally different collection of inputs, X and Y

---

15 The only cases where the degree of the relationship is constant throughout, so that the ridge lines are linear, are the cases of neutrality, perfect complements and substitutes. For partially related cases it is assumed that the degree of the relationship varies within the area where both marginal productivities are positive.

16 "...if several elements are associated in a productive process, the effect on the output of varying any one of them is dependent upon its relation to all the other elements and also upon the effects of the other elements upon each other." (Black, op. cit., p298)

may be partial substitutes, input relations being defined in the same way in the two cases. But once the original of the isoquant map is fixed in hyper space, the shape of the ridge lines for any two inputs is fixed by the nature of the two inputs, and by their relation to the fixed elements.

II - A Typical Isoquant Map.

With the aid of the ridge line definition of input relations a typical isoquant map can be derived from the hyper production figure. That is, it is now possible to select as the X and Y axes those axes associated with inputs which are not related in any special and rigid manner. Consider a selection of two inputs which are partially related in a complementary manner at a given equilibrium point on the hyper figure. Through the equilibrium point pass a plane parallel to the plane containing the two chosen axes, designated as X and Y. The resulting cross-section yields the isoquant map illustrated in Figure 32.

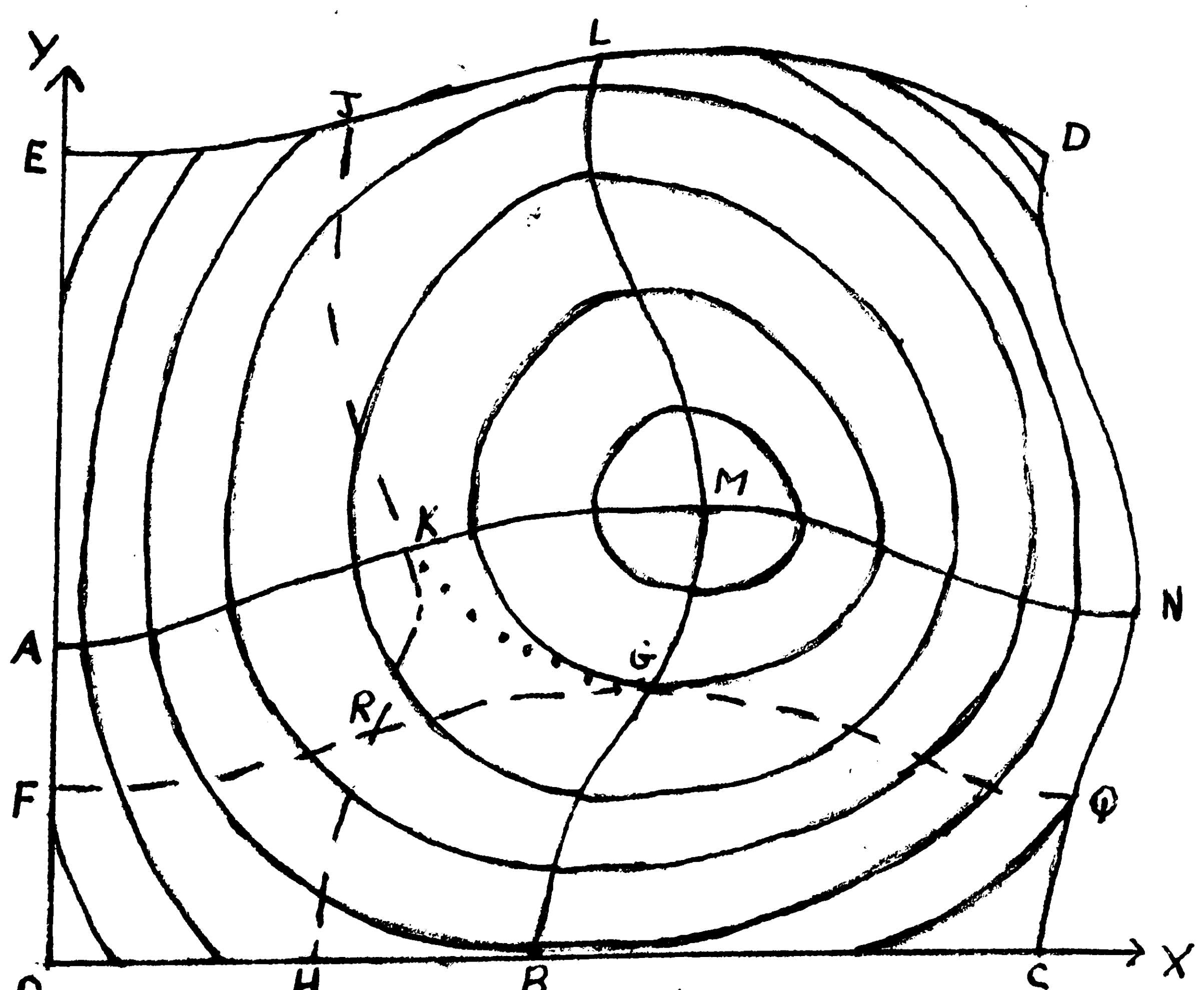


Fig. 32. - Isoquant map for partial complements.

This map is almost identical with the indifference map for partially complementary goods illustrated by Figure 13 in Chapter IV. One of the major differences between the two lies in the additional loci of maximum average product which can be introduced into Figure 32. Two of these loci are FRGQ and HRKJ, drawn roughly parallel to their respective ridge lines because "the later that total outputs begin to diminish, the later also that average outputs are likely to begin to diminish."[17] The other additional locus of maximum average product is of course the locus of maximum average product for variations along the diagonals. This locus must pass through the points K and G in some such manner as the dotted curve lying between these points. This dotted curve marks the inner boundary[18] of the economic relevant area the other two boundaries of this area being KM and GM. The boundaries of the entire map are the two axes and the curves ELD and CND. These latter two curves probably would have the same general shape as the ridge lines since the later maximum total product comes the later also that output is likely to be reduced to zero again. The points E, D, and C are points representing the same output as the origin.

---

17 Black, op. cit., p302.

18 When the ridge lines are curvilinear Professor Cassel argues that "the determination of the precise location of this inner boundary line would naturally be a matter of greater difficulty" than when they are linear. He concludes that it would lie in the same area of the map but that it need not pass through the points K and G in Figure 32 claiming that it could intersect the ridge lines at points between K and M and G and M. (Cf. "On the Law of Variable Proportions", op. cit., pp117-8) This of course is erroneous. The locus must always pass through K and G no matter what shape the ridge lines have because at these two points the sum of the function coefficients is always one.

The isoquant map of Figure 32 is probably the most general type of isoquant pattern for two dimensional analysis, since complementarity, as defined here, is likely to be the dominant input relation.[19] For a substitution relation the isoquant map would only differ from the one in Figure 32 in that all the loci and boundaries would slope in the opposite direction. This does not apply to the dotted locus KG which would not change very much in shape with a change in the relationship.

Let us now briefly summarize the assumptions and conditions lying behind the generalized isoquant map of Figure 32. First, the origin of the map is a point lying on some non-zero product contour of the hyper production figure. This implies that neither X nor Y are indispensable for production, for if they were, the origin would lie on the zero product contour of the hyper figure. Secondly, the X and Y inputs are assumed not to be perfectly related to any other input. They neither have perfect complements nor perfect substitutes. Thirdly, the hyper production figure is restricted in size, so that amounts of a good or service not meeting the definition

---

19 "Production requires the co-operation of three essential groups of factors: equipment, labour and raw materials. It is a significant fact of modern technique that <u>given the type of equipment in existence</u> there is a relation of strong complementarity between them: that is to say, the extent to which the proportions of these factors can be varied in production is highly limited." (N. Kaldor, "Stability and Full Employment", <u>Economic Journal</u>, 1939, p643. Italics are in the original)

That is, for any two inputs it is more probable that the ridge lines will have positive slopes than negative slopes since the possibility of varying the proportions of these inputs in production is more restricted in the former case.

of an input are excluded. Fourthly, whatever the amount of an input chosen for the productive process, the entire amount must be used up in the process. Thus marginal productivity can become negative, with the result that there is only one point of maximum total product on the map. Finally, the amount of product due to X and Y is measured, not as a deviation from zero product, but as a deviation from the output due to the fixed elements.

CHAPTER XII

ANALYSIS OF SOME SPECIAL ISOQUANT MAPS

The construction of this chapter is similar to that of Chapter V. This chapter is, like Chapter V, a miscellany bringing together and analyzing all those cross-sections of the hyper figure not previously covered. The examination of the hyper figure is completed by dropping the assumption that X and Y are not perfectly related to any other input, and by dropping all restrictions on the location of the point in the hyper figure through which the plane cross-sections are taken. That is, this point may lie anywhere on the figure. When it lies on the zero product contour of the hyper figure, problems connected with indispensable inputs and catalyst-type inputs arise. When the X and Y inputs have perfect complements or substitutes, other special cases arise. These latter cases are dealt with first.

When the X and/or Y inputs have perfect complements or substitutes, the possibilities are similar to those of Chapter V. For the analogy between the production and indifference instances to be complete, the X and Y inputs must not be indispensable for production. This means that the application of any small amounts of X and Y to the fixed elements brings about an immediate increase in output. Moreover, the inputs must be of the type which make a direct contribution to production. That is, they must yield some total product of their own accord in contrast to the type of inputs which merely enable other inputs to be more productive.[1]

It is only necessary to summarize these possibilities again. Where one input, Y, has a perfect complement which is missing from among the fixed elements, or a perfect substitute whose saturation amount appears among the fixed elements, and where the other input, X, is not perfectly related to any other input, the cross-sectional planes are all "empty", except when X and Y are partial complements. Then the special map of Figure 15B exists. If each input has a perfect complement missing from the fixed elements, or if each has a perfect substitute present at the saturation amount, or if one input has a perfect complement missing and the other a perfect substitute present at the saturation amount, all the cross-sectional planes for such inputs are "empty". The two special cases of Figures 16B[2] and 22 may

---

1 An example of this catalyst-type input is scarecrows. A special case, discussed below, arises when the origin of the isoquant map lies on the zero product contour of the hyper figure and one of the inputs is of the "scarecrow" type.

2 Professor Black (op. cit., pp300-2) presents an analysis of an isoquant map similar to the one in Figure 16B. In his figure Y represents commercial fertilizer and X, seed. The fixed elements consist of a given quantity of land, the nature or composition of which is not carefully defined so that this situation is not considered by Professor Black to be a special one. Actually, however, the fixed elements consist not just of land but of land containing an excess amount of natural fertilizer which is a perfect substitute for commercial fertilizer. Hence the reason for Professor Black's unorthodox isoquant pattern. Twenty years after the appearance of his Production Economics this explanation is suggested in one of his recent books, but it is not connected with the type of pattern under discussion. Cf. Black and others, Farm Management, pp394-7.

Professor Black's diagram does serve the useful purpose of demonstrating that such special cases as Figure 16B do have realistic content in production problems.

also arise here. The former exists if one input, Y, has a perfect substitute which is present among the fixed elements at an amount greater than the saturation amount and if Y and X are partial complements. The latter of the two possibilities may develop if the perfect substitute for Y is present at an amount just less than saturation and if Y and X are partial substitutes. Finally, to complete this summary, if the non-zero amounts of a perfect complement for X or Y, or for both, are present among the fixed elements, the cross-sections for the XY inputs have "empty" spaces in the neighborhood of the axes just as in the corresponding indifference situations.

For the next batch of special cases it is a matter of indifference whether or not X and Y are perfectly related to other inputs, but it might be less confusing if perfect relations are ruled out. So for this section of the analysis suppose that X and Y have no perfect relations and that both are of the input type which yield total product directly on application to the productive process. Now consider what may happen if the origin of the XY isoquant map is a point which lies on a zero product contour of the hyper production figure. This may mean one of three things, or any combination of the three. First, it may mean that the quantity of all other inputs are fixed at their zero amounts.[3] Secondly, it may

---

3 i.e., the origin of the isoquant map is at the origin of the hyper figure.

mean that although some input quantities are fixed at non-zero values, some essential input (other than X and Y) is missing so that the fixed elements can produce no output. Finally, it may mean that X and Y are indispensable for production and that all partial substitutes for them are not present among the fixed elements in sufficient quantities to permit production without X and Y. The latter two meanings are the ones of interest here.[4] The analyses associated with them are unique to production theory since in indifference analysis no element is essential for utility[5] so that any collection of consumer goods other than the "zero" collection yields some utility.

Consider first the implications of the second of the above meanings. If some essential input is missing for which neither X nor Y is a partial substitute, the XY plane is "empty". However, if Y, say, is a partial substitute for the missing essential input, and if X and Y are partial complements or independent, then combinations of X and Y yield some product after a minimum amount of Y is present. The resulting isoquant pattern is normal, except for the fact that the isoquants can not cut the X axis. However all the isoquants which do not cut the Y ridge line cut the Y axis. If X and Y

---

4 Where the origin of the isoquant map coincides with the origin of the hyper figure all the XY cross-sections may be "empty" unless there exist two inputs which, along with the economic unit, can alone turn out some product. In this case, unless X or Y alone can work with the economic unit to produce the product, both inputs will be indispensable for production.

5 The obvious exception of perfect complements is ignored.

are partial substitutes for each other and for the missing essential input, then output arises whenever a minimum amount of X or Y is mixed with the fixed elements. This fact is reflected in the nature of the zero isoquant which is a curve, instead of a point, cutting both axes. Below this isoquant the plane is "empty" while above it the plane is covered in the usual fashion with contour curves.

The third interpretation, given in connection with the isoquant origin lying on a zero product contour of the hyper figure, has the same implications for the isoquant pattern as the second interpretation just dealt with. That is, if either X or Y is indispensable for production, the isoquants are debarred from cutting the axis of the non-essential input, unless X and Y are partial substitutes. In that case some of the higher isoquants might cut the non-essential input axis. If both X and Y are indispensable for production, then the isoquants can not cut either axis,[6] except for the qualifications noted when X and Y are partial substitutes.

The restriction barring perfect relations can now be relaxed, and the consequences for the above analysis of indispensability noted. If X and Y are inputs which have perfect complements or perfect substitutes, the property of indispensability has no effect in those cases where the XY plane is "empty". In the special situation represented by Figure 15B only the input which has no perfect relation can be

---

6 This is the situation illustrated by Professor Cassels' Figure 3, "On the Law of Variable Proportions", _op. cit._, p114.

indispensable. If it is indispensable, the isoquants of Figure 15B are debarred from cutting the axis of the dispensable input which is Y there. In the special situations of Figures 16B and 22 only the input without the perfect substitute can be indispensable. If it is indispensable, then in 16B all isoquants are prevented from cutting the dispensable input axis which is the Y axis there. In Figure 22 probably only the lower isoquants are prevented from cutting the dispensable input axis, Y, because of the substitution relationship between X and Y. For those cases of perfect complementarity for which the isoquant map has "empty" spaces the existence of indispensability increases, or adds to, these "empty" areas.

The final set of special situations occurs when not only the origin of the isoquant map lies on the zero product contour of the hyper figure but when at least one of the inputs is of the catalyst-type.[7] If both inputs are of the catalyst-type then the XY plane is completely "empty".[8] The plane is not "empty" however if only one of the inputs is of the catalyst-type while the other is of the more normal type which can create product whether or not the fixed elements yield any output. For example, in the illustration of the

---

7 i.e., types of inputs which make no direct contribution to production.

8 This is only so if the origin of the map lies on the zero product contour of the hyper production figure. If the origin lies on a non-zero contour of the hyper figure the XY plane is covered with iso-product curves in the usual way. This is because the total product associated with the fixed elements is increased by the use of catalyst inputs. It is only when this product is zero that this type of input can make no contribution to production.

scarecrows and seed used in Chapter XI, the presence of the scarecrows adds nothing to total product but only serves to shift the total product curve for seeds so that maximum total product for seeds comes for smaller seed quantities. However, the value of the total product is not affected thereby. The situation is illustrated in Figure 38 where X is a catalyst-type of input and Y is a normal type of input. Since

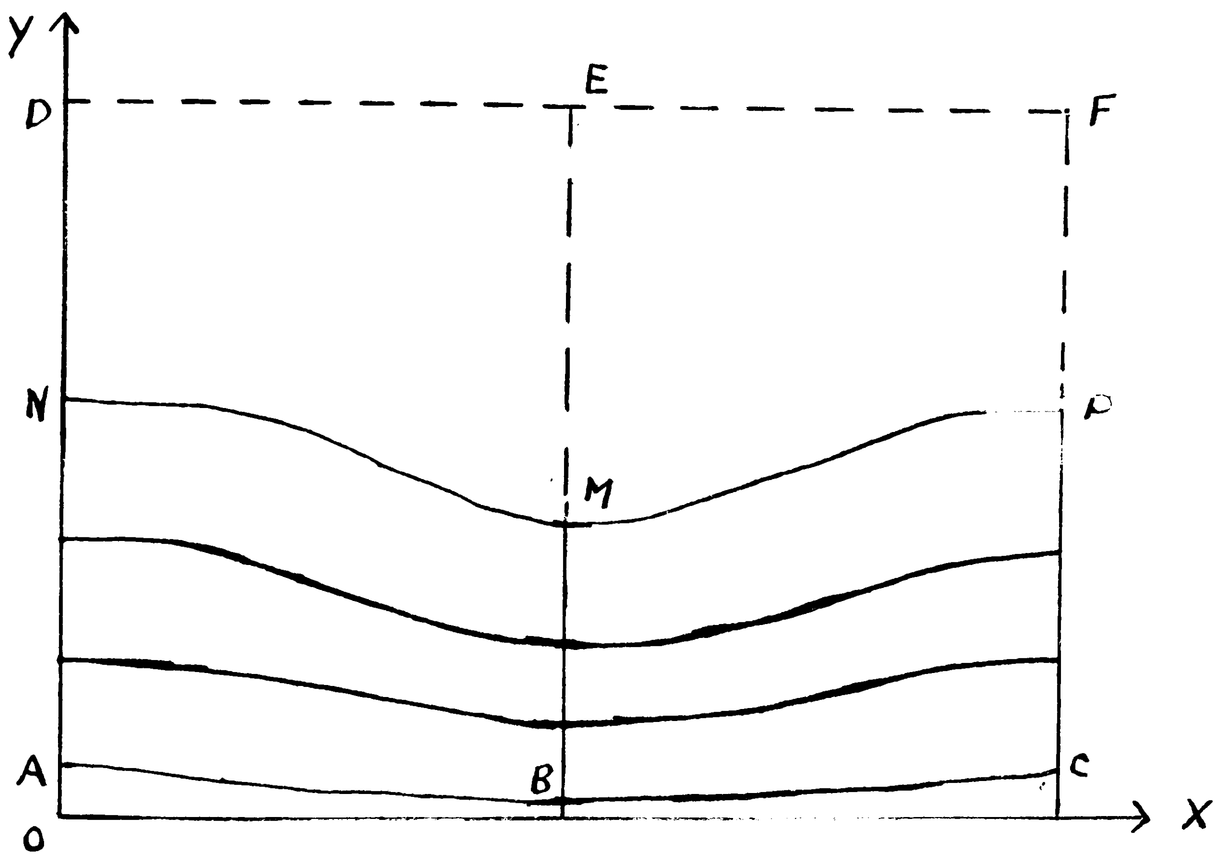


Figure 38. - Isoquants for scarecrows and seed.

the origin of this map lies on a zero product contour of the hyper figure, additions of X, when Y is fixed at zero, do not result in any output. However, when Y is fixed at OA, say, additions of X result in that amount of Y producing a greater output than it does alone. That is, X is a partial substitute for Y. This is indicated by the bulge of the isoquants towards the X axis.[9] The line MB is the locus of all points where the isoquants turn parallel to the X axis and indicates the X value for which product is maximized when X alone is varied. Beyond BM X becomes excessive, which interferes with its ability to replace Y. As X is increased this excessiveness grows to the point where X can no longer replace Y. This occurs at the line CP. Beyond this boundary additional X makes no contribution to production and may even hamper it. Therefore such additional X is not an input so that CP forms the eastern boundary of the map.

When X is fixed and Y is varied output increases in the manner described by the law of variable proportions. NMP is the locus where total output for Y is a maximum, and it is also the isoquant of maximum output. If the map extended beyond this isoquant the line DEF would be the locus where the total product for Y drops to zero again. The X axis is the zero isoquant.[10] The isoquants in the neighborhood of the X axis are very nearly straight lines parallel to the axis. There are two reasons for this. First, the isoquants

---

9 If X and Y were independent then the isoquants would be straight lines parallel to the X axis. The isoquant pattern would then be similar to the indifference pattern discussed on p54n1 above. If X and Y were partial complements the isoquants would bulge away from the X axis. In both these cases X is not an input because in the former case it makes no contribution to production at all while in the latter case it actually hampers the productive efforts of Y.

10 Strictly speaking, only the point O is connected with the hyper figure and all other points on the X axis do not appear on the hyper figure since X is not an input when Y is zero.

can not cut the X axis, and secondly, when the amount of Y is small, X has little effect on its output. Only as Y, and so output, increases does the influence of X become sufficient to make the isoquants bulge towards the X axis. Thus the higher the isoquants the greater the bulge.

The northern boundary of the map is formed by the isoquant of greatest product, NMP, for when the amount of Y is fixed between the points N and D additions of X reduce the output produced by such a Y quantity. As X is increased output becomes less and less until it reaches a minimum for points along ME. Therefore, in this region X is not an input so that the area above NMP is "empty". The point M is the point where product is maximized when X is free.

Figure 38 reveals that unless the cost of X is very low relative to the cost of Y (which means that Y is scarce and expensive and that X is plentiful and cheap), X will not be used at all. The iso-cost curves must have very little slope if they are to be tangent to the isoquants at points to the right of the Y axis. Therefore, in this situation, X is only employed when it is necessary to conserve the essential input Y.

This isoquant pattern does not exist if either X or Y has a perfect complement missing from among the fixed elements or a saturation amount of a perfect substitute present. If amounts of a perfect complement exist for either input in the fixed elements then an "empty" space appears next to the axis of the input having the perfect complement.

This type of pattern is unique to the production hyper figure. It is not yielded by any cross-section of the hyper preference figure for there are no consumer objects analogous to scarecrows whose presence alone in the consumer's budget has no desireable effect but which has the effect of making other consumer objects more desireable.[11]

All (input) plane cross-sections of the hyper production figure have now been examined. The most general pattern is that illustrated by Figure 32. Less general patterns are those of Chapter X. Special cases similar to those represented by Figures 15 B and C, 16 B and 20 in indifference analysis, also arise here. The phenomenon of indispensability is peculiar to production theory, although it has the same effect on isoquant maps as the existence of thresholds had on indifference maps. As the analysis suggests, this phenomenon is not frequently encountered. Also peculiar to production theory is the situation involving the catalyst-type input. Finally, there are those XY cross-sections of the hyper production figure which are completely blank. These are probably more numerous in production theory than they are in utility theory.

---

11 Such objects can be conceived of but the writer can think of no practical examples or realistic counterparts for such conceptions outside the realm of perfect complements.

CHAPTER XIII

A DIGRESSION ON THE CONTOUR ANALYSIS OF LINEAR HOMOGENEOUS PRODUCTION FUNCTIONS

The preceeding analysis rules out constant returns to scale by placing the given production function into a technological context within which it can not be linear and homogeneous. As a result, the hyper figure previously dealt with corresponds to a non-homogeneous function. However, the technological environment of the production function may be set up so as to be consistent with homogeneity. In this chapter the technological environment is so set up. Certain plane cross-sections of the hyper figure then reflect constant returns to scale. These cross-sections differ greatly in appearance from any yet considered. It is the object of this chapter to analyze these plane cross-sections in order to expand upon and extend the existing contour analysis of linear homogeneous production functions.

To obtain the type of hyper production figure required for this object, it is only necessary to associate the given production function with an economic unit whose contribution to the productive process is negligible.[1] All other conditions of the basic experiment, laid out in Chapter IX, are unaltered. Therefore, since the production function is a continuous fun-

---

1 The view adopted here is that the case of constant returns to scale is the limiting case in which the contribution to production of the economic unit vanishes. This, of course, is Professor Hicks' view. Cf. Value and Capital, p96n.

ction of continuous variables,[2] a proportionate increase in all inputs leads to the same proportionate increase in product.[3] The production function is linear and homogeneous. From the hyper production figure associated with this function plane cross-sections can be derived which yield isoquant patterns similar in shape to those that can be derived from the hyper production figure associated with a non-homogeneous function. However, there is one type of cross-sectional pattern peculiar to the homogeneous hyper figure. This pattern arises when the cross-sectional plane is taken through the origin of the figure. The resulting isoquant maps so obtained, in contrast to all the others, reflect constant returns to scale.

---

2 i.e., the inputs are continuously divisible and the productive process continuously variable.

3 "If the amounts of all elements in a combination were freely variable without limit and the product also continuously divisible, it is evident that one size of combination would be precisely similar in its workings to any other similarly composed." (Knight, Risk, Uncertainty, and Profit, p98)

"If all agencies of production were infinitely divisible, if infinitesimal miniatures of each factor were possible...the productive set-up turning out a single commodity would operate forthwith at minimum costs. One combination would be as effective as any other and the question of size would lose significance." (M.M. Bober, "Theoretical Aspects of the Scale of Production", Economics, Sociology and the Modern World, Essays in Honor of T.N. Carver, N.E. Himes, ed., 1935, p79)

In brief, perfect divisibility and constant returns to scale are synonymous. Cf. F.H. Hahn, "Proportionality, Divisibility, and Economics of Scale: Comment", Quarterly Journal of Economics, Feb. 1949, p131.

There are three methods of passing a plane through the origin of the homogeneous hyper figure, all of them depending upon the selection of the two axes of reference. First, any two axes of the figure may be selected and the plane defined by these two intersecting lines taken as the cutting plane. Secondly, any two straight lines in hyper space, other than axes, perpendicular to each other and to the product axis at the origin of the hyper figure,[4] both lying in a positive quadrant of a hyper space, can define the cross-sectional plane. Thirdly, any one input axis and any other line in hyper space, not an axis, perpendicular to it and the product axis at the origin and lying in a positive quadrant can determine the desired plane. These are the methods of selection, of which the last two require further explanation. Any line in a positive quadrant of hyper space, perpendicular to the product axis at the origin, measures some homogeneous combination of two or more inputs. Along such a line the proportions in which the various inputs are combined remain fixed so that the line can be taken as an axis measuring amounts of some compound input variable. Thus, if two hyper lines, or one hyper line and one input axis, define the cutting plane, they also constitute the axes of reference for measuring amounts of two variables, each variable being either a single input or some compound mixture of inputs.

---

4 The requirement that the selected lines of reference should intersect at right angles is more restrictive than necessary. It is imposed here so as to make the results comparable with the existing analysis which uses rectangular coordinates.

If the two variables represented by the axes defining the cross-sectional plane can not collectively produce some product then the plane is "empty". If they can produce collectively they may, or may not, be capable of producing individually. Existing analysis usually assumes that neither variable alone can yield any product but that both together can. The most general isoquant pattern of the textbooks reflecting constant returns to scale[5] take X and Y as not being perfectly related to each other either as substitutes or complements. If the two variables do not represent between them all the inputs, then the missing inputs are fixed at their zero amount and, by definition, none of the missing inputs can be essential ones. The isoquants are still assumed to be convexed to the origin when both marginal productivities are greater than zero. This standard textbook pattern is given by a cross-sectional plane through the origin of the homogeneous hyper figure. Figure 34, overleaf, is representative of this type of pattern.

Along any positive sloped straight line passing through the origin of this map, except the two axes, constant returns to scale must be exhibited. This means that any such diagonal must cut the isoquants in points where the tangents are parallel.[6] Thus, all straight lines through the origin, except the axes, are isoclines, and all isoclines converge on the origin.

---

5 The general type of isoquant map used in illustrating the analysis of constant returns can be found in the following places: Allen, <u>Mathematical Analysis</u>, p320; Boulding, <u>Economic Analysis</u>, 1st ed., p496, Fig. 63; Cassels, "On the Law of Variable Proportions", <u>op</u>. <u>cit</u>., p110, Fig. 2.

6 Cf. Allen, <u>Mathematical Analysis</u>, p320.

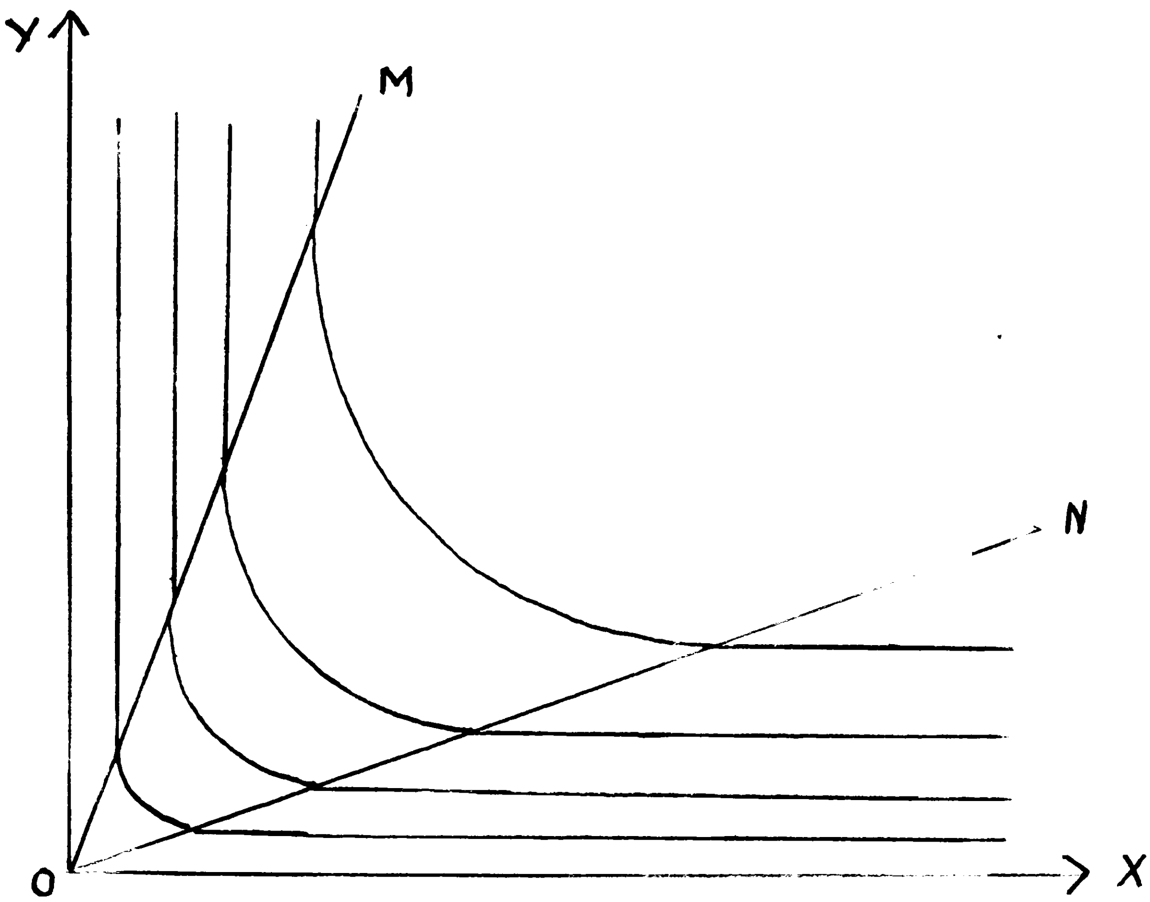


Fig. 34. - Isoquants reflecting constant returns to scale.

Since the ridge lines are only special kinds of isoclines they too are straight lines converging on the origin. In Figure 34, OM is the Y ridge line and ON, the X ridge line. The isoquants, of course, turn parallel to one of the axes at these ridge lines. The point of intersection of the ridge lines either denotes a maximum or minimum. Since this point here coincides with the origin it is obvious that it denotes a minimum. There is no locus of maximum product when both variables are varied together so that the isoquants never close. However, OM is the locus of maximum product when Y alone is varied and ON is the corresponding locus for variations in X. Above OM Y is excessive while below ON X is redundant. If excess units are not a nuisance,[7] and it is usually assumed that they are not in drawing this type of map, the isoquants remain parallel to one of the axes beyond the ridge lines in the manner indicated in Figure 34.[8] The axes therefore form the isoquant of zero product since, when one variable if fixed at zero, the other can not produce anything.

The manner in which total product varies when one of the variables is fixed at a non-zero value and the other varied may be determined by use of the elasticity proposition of Chapter IX. Along any diagonal the elasticity of production is unity. This means that at any point on the map the sum of the function coefficients of the two variables is one. Now in the area cut off by the Y axis and OM the function coefficient of Y is zero and therefore the function coefficient of X is one. That is, as X is increased from zero, Y fixed at some non-zero value, marginal and average returns are both constant and equal to each other.[9] Therefore, there is no

---

7 i.e., excess units do not have to be used.

8 To be consistent with the assumptions and definitions of Chapter IX the assumption that excess units are not a nuisance should not be used here. Figure 34 is drawn as it is for the purpose of following the textbook exposition as closely as possible. Below, the description of this type of cross-section is given as it would actually appear if the homogeneous hyper figure were constructed on the same terms as the non-homogeneous one.

9 The "fixed" amount of Y must, in this case, be treated as completely divisible so that some units of it may be disregarded or not used. Then Y is never excessive to X since redundant Y units can be ignored. If Y was excessive for small amounts of X there would be increasing returns to X at first and its function coefficient would be greater than one. The

first phase of increasing average returns as described by the law of variable proportions. The Y ridge line marks the end of this first stage of constant average returns. Between the two ridge lines total product varies, as X alone varies, in the manner described by the second phase of the "law". Marginal and average product declines until total product is zero at ON. Beyond ON total product remains at its maximized level as X is increased since the redundant X units are not used. The same three stages of constant average returns, diminishing average returns, and constant total product also occur when Y alone is varied.

When all units of an input amount must be used the isoquants become positively sloping beyond the ridge lines. The zero isoquant does not coincide with the axes but instead consists of two straight lines passing through the origin. One branch of this isoquant lies between the Y axis and Y ridge line and the other branch between the X axis and X ridge line.[10] The space between each branch of the zero isoquant and the nearest axis is "empty" space. The isoquants within the

---

function coefficient of Y would then have to be less than zero in which case the isoquants could not be parallel to the Y axis beyond the Y ridge line. Thus the parallelism of the isoquants implies not only that excess units of the variable input are ignored but also that excess units of the "fixed" input are ignored.

10 The two branches of this isoquant must be straight lines so that no diagonal through the origin will cut them. If any such diagonal were to cut one of these branches constant returns to scale would not exist along that diagonal.

redundant areas may have one of three different shapes. First, they may be straight lines parallel to the nearest zero isoquant branch. Secondly, they may curve continuously from a horizontal or vertical slope at one of the ridge lines to a slope equal, at infinity, to that of the nearest branch of the zero isoquant. And lastly, they may curve until their slopes are equal to the slope of one of the zero isoquant branches at a finite value, retaining that slope for all larger values.

When the slopes of the isoquants are positive within the redundant areas, the law of variable proportions summarizes completely the manner in which total product varies as one variable alone is increased.[11] Here, of course, the total product curves do not start at zero values but at positive values for the variable input since the zero isoquant deviates in a positive direction from the axes. Also the ridge lines now become the loci of maximum average product for variation in each variable individually.[12] Therefore, the effective region of the map here, as in the case where the isoquants are parallel, is the area lying between the two ridge lines.

More light can be thrown upon this type of isoquant map by considering two extreme cases which are also familiar textbook illustrations. First, consider the cross-section of the homogeneous hyper figure yielded by the plane through

---

11 Consideration of the elasticity proposition makes this clear. Because all inputs must be used the "fixed" input is actually indivisible since none of it can be ignored.

12 This may be proved by means of the elasticity proposition, or it may be proved by a more direct method. Cf. Allen, <u>Mathematical Analysis</u>, p321.

the origin when the axes of reference represent two variables which are perfect complements for each other. Such a plane produces the system of isoquants illustrated in Figure 36A, if the production function is defined so that excess units can be ignored. Here the ridge lines coincide so that the

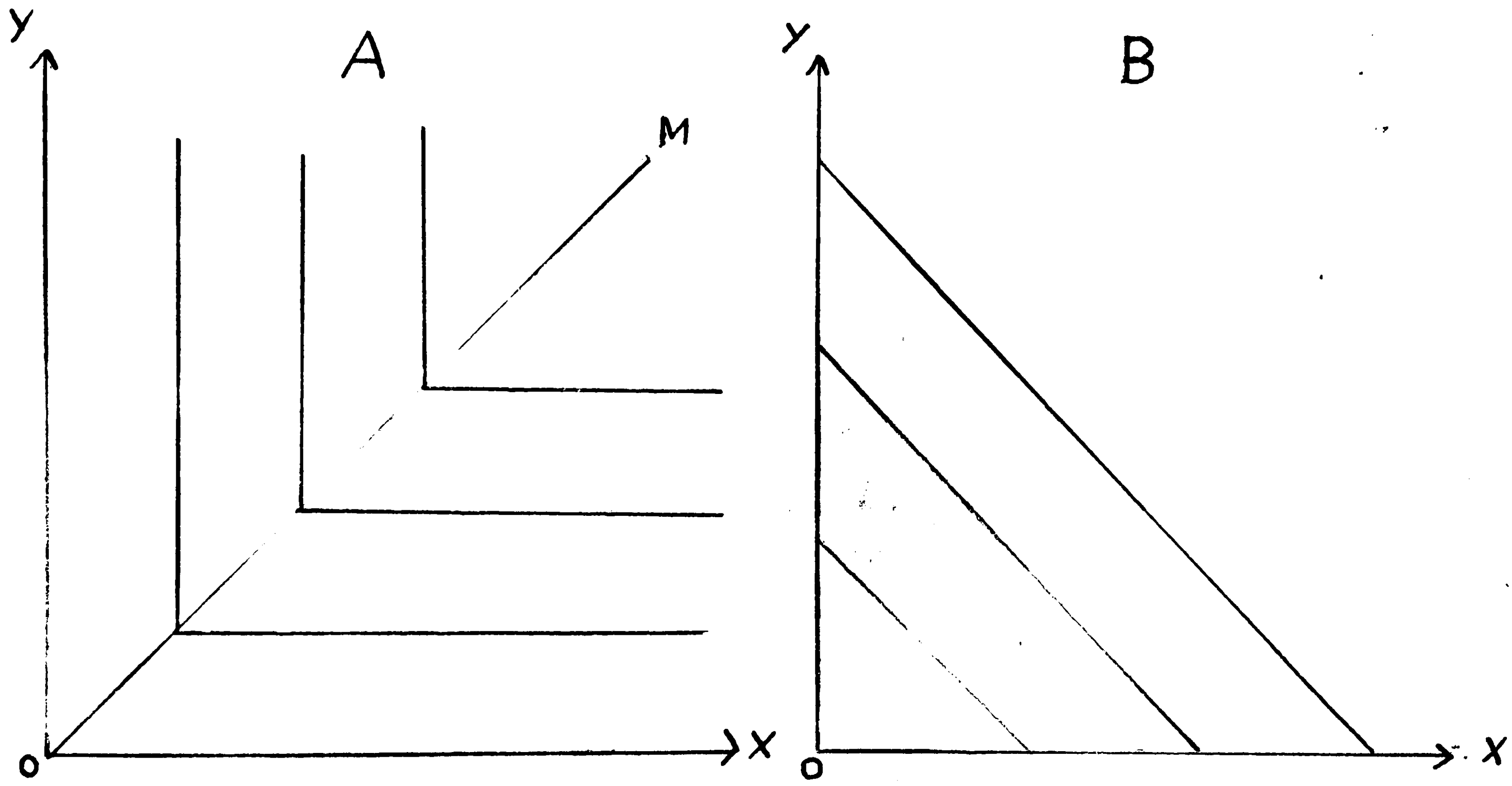


Fig. 36. - Isoquants for perfect complements and substitutes.

effective region of the map is the straight line OM along which, of course, constant returns to scale prevail. When one of the variables is increased from zero up, the other variable fixed at some non-zero value, average product is constant until OM is reached; beyond OM total product is constant. When the production function is defined so that all input units have to be used, the isoquants slope in one of the three ways described above and variation in each variable, individually, results in increasing average returns up to OM and beyond OM, in diminishing total product.

The other extreme case of the textbooks is the case of perfect substitutes. To obtain a cross-section of the homogeneous hyper figure which yields an isoquant map for two perfect substitutes, the variables selected must be capable of producing a product individually. Now there is not likely to be one input which, individually, can yield any product in a complex productive process without the help of other inputs. Therefore, one of the variables must be some mixture of several inputs combined in a given fixed proportion which represents sufficient inputs so that production is possible without the use of additional inputs. This compound input is represented by a line lying in a positive quadrant of hyper space, not an axis, passing through the origin of the hyper figure and perpendicular to the product axis. If there is another line in hyper space perpendicular to this line at the origin representing a compound input which is a perfect substitute for the first input,[13] then these two lines define a cross-sectional plane of the hyper figure, which yields an isoquant map for two perfect substitutes. The isoquant pattern is like that in Figure 35B. The isoquants are straight lines cutting both axes. The pattern is similar to the indifference pattern for perfect substitutes, but here there is no one isoquant of maximum value. There is only a point of minimum value and this is the origin. The ridge lines do not appear on the map at all. The effect is the same as if the ridge lines in Figure 34 were spread apart by rotating them

---

13 This requires that the given proportion in which the inputs in this mixture are combined be identical to the proportion contained in the first mixture. Also all the inputs found in one mixture either have to appear in the second or else be represented there by perfect substitutes.

about the origin until they disappeared from the positive quadrant. Because the ridge lines do not appear on the map constant returns to scale are reflected not only along all straight lines passing through the origin, including the axes, but also along all straight lines having positive slopes.[14] Hence the effect on returns is the same as if there were only one variable[15] since varying the proportions in which X and Y are combined does not interfere with the constancy of returns.

Besides the three types of isoquant patterns just discussed there are two other types which also reflect constant returns to scale. Neither of these are considered by existing analysis. One of these types arises when both axes represent economic variables which are not perfect substitutes yet which individually, as well as collectively, can produce some product. There are probably no two simple inputs which individually can produce an output in a realistic productive process when all other input amounts are zero. Therefore,

---

14 All such straight lines may be considered as isoclines since they cut the isoquants at points of equal slope.

15 The attempt to apply the elasticity proposition here also points to the conclusion that perfect substitutes should be considered as a single variable. If there are two variables there are two function coefficients whose sum is everywhere unity. However, if X alone is varied, Y fixed at some non-zero value, its function coefficient is one for all X values. Similarly, when Y alone is varied, its function coefficient is equal to unity for all Y values. Therefore, the sum of the two coefficients is two — a contradiction. X and Y must therefore be the same variable so that there is only one function coefficient which is everywhere equal to unity.

When perfect substitutes were discussed in the non-homogeneous case it was not necessary to use this argument although it could have been introduced there too. Any inconsistencies arising there from using the elasticity proposition were explained in terms of discontinuities in the total product curves. Cf. p147 above.

the two economic variables are taken to be compound inputs made up of some constant mixture of many inputs. The cross-sectional plane through the origin determined by the two hyper lines representing the appropriate economic variables yields a system of convexed isoquants which cut both axes. This type of isoquant pattern is illustrated by Figure 36.[16]

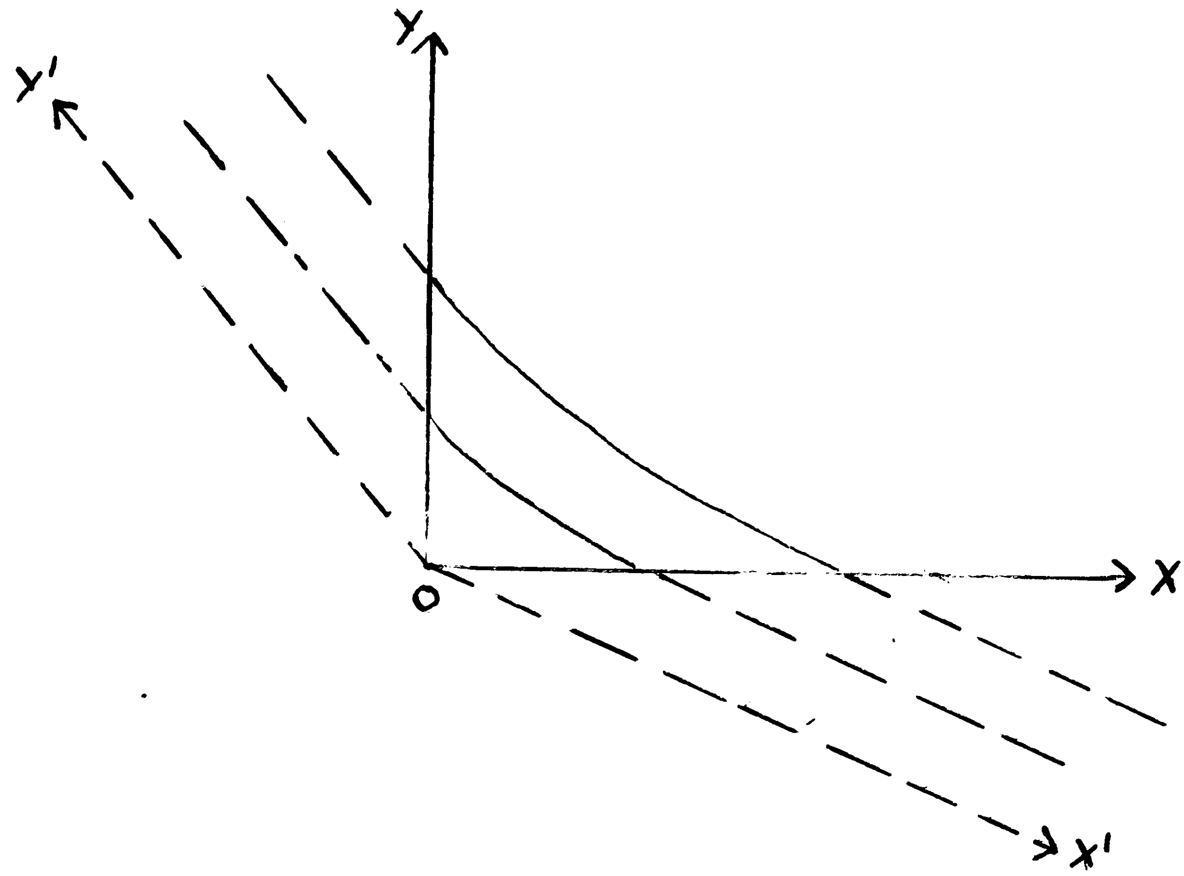


Fig. 36 - Constant returns along both axes.

Since constant returns to scale prevail along both axes in Figure 36, these axes are also isoclines, which in the absence of conventional type ridge lines, are also the two extreme isoclines. That is, the axes here set the limits to the variation in the slopes of the isoclines. They also serve to limit the variation in the slope of the isoquants. The

---

16 To the writer's knowledge the only person who has used this type of isoquant pattern, in print, is John S. Chipman, "Returns to Scale and Substitution", The Canadian Journal of Economics and Political Science, May 1950, p220, Fig. 3.

slope of each isoquant, Figure 36, varies between some negative value, at the Y axis, greater than minus infinity and some negative value, at the X axis, less than zero. When ridge lines of the conventional type appear on the isoquant map, as they do in Figure 34, the slope of each isoquant varies between minus infinity and zero.

Since the axes serve to set a limit to the variation in the slopes of the effective isoclines, they serve the same function as ridge lines. The two axes may then be considered to be ridge lines. If they are so considered then they must take on all the properties of the ridge lines. That is, they must not only set a limit to variations in the slopes of the isoclines and isoquants but they must be the loci of all points where the isoquants turn parallel to some lines of reference. There exists in hyper space two lines, lying in the XY plane, which may be considered as the geometrical axes to which the isoquants turn parallel. In Figure 36 the lines OX' and OY' designate such a set of geometrical axes. Neither line represents an economic variable since each is associated with some negative values.[17] Only in a geometrical sense can these lines be taken as the axes for the isoquant system of Figure 36 for which the economic axes are the ridge lines. Of course the only part of the isoquant system derivable from the hyper figure is the relevant part contained by the X and Y axes.

The use of geometrical, but economically meaningless,

---

17 i.e., these lines lie in a negative quadrant of hyper space.

axes in analyzing the type of isoquant pattern represented by Figure 36 suggest a generalization for explaining the variety of isoquant maps considered in this chapter. The degree of variation of all these isoquant patterns from the perfect complement pattern of Figure 36A can be calibrated by comparing the curvatures of the effective segments of the isoquants in the different maps. For perfect complements, the ridge lines coincide so that the curvature of the isoquants is zero. As the ridge lines spread apart the curvature of the isoquants decrease. When the ridge lines coincide with the economic axes, the geometrical axes form an oblique angle at the origin. The greater the angle formed by the geometrical axes at the origin the less the curvature of the isoquants. In the limit when the geometrical axes form a straight angle the curvature is reduced to zero. This, of course, is the situation of perfect substitutes.

Returning to the discussion of Figure 36, it will be noted that total product arising from increases in any one input alone never reaches a maximum. However, as X alone, or Y alone, is increased, total product increases at a dininishing rate so that average and marginal product decrease from the start.[18] Since total product never reaches a maximum

18 This follows from the elasticity proposition. Both function coefficients must be less than one so that their sum can be one. Since each coefficient must be less than one the total product curve for X, or Y, can not be a straight line passing through the origin. It can only be a curve concaved everywhere to the input axis. On such a curve the function coefficient is everywhere less than one and total product always increases at a diminishing rate.

the variable input never becomes excessive relative to the "fixed" input. This is the situation where the variables are such that each unit of the X variable, say, contains some Y inputs or substitutes for Y. Then as X increases the amount of "fixed" Y is supplanted so that X can not become excessive to Y.[19] Such excessiveness is also avoided if the "fixed" input is ignored. It can be ignored since it is not required by the variable input for production. That is, the variable input need not be combined with the "fixed" input, but both may be entered into the productive process on a separate and individual basis. Whichever explanation is accepted for the failure of the variable input to become excessive to the "fixed" one, the two inputs can be considered to be substitutes since neither input requires the presence of the other to produce.[20]

To generalize on input relations reflected by the isoquant patterns of this chapter, it may be said that where the economic axes serve also as the ridge lines, the inputs are substitutes. The less the variation in the slope of the isoquants, the stronger the substitution relation. Where the ridge lines

---

19 In the non-homogeneous situations where X and Y are substitutes X always becomes excessive to Y because, although X is a substitute for the "fixed" Y, it is not a substitute for the fixed contribution of the economic unit.

20 If the Johnson-Allen definition of complementarity is applied to the type of isoquant map under consideration, it would indicate that the inputs were, on its terms, substitutes. In this case all "non-normal" positions lie on the Y axis. Cf. Chapter III above.

have positive slopes but do not coincide with the economic axes, both inputs have to be used together to obtain a product so that, on a common sense basis the inputs are complementary. This complementary relation varies inversely with the degree of spread between the ridge lines. There seems to be no border line case of independence here.

One other special isoquant map reflecting sonstant returns to scale remains to be considered. In this pattern one variable is capable of producing various amounts of product by itself but the other variable is unable to produce anything by itself. To illustrate this situation suppose that the Y variable can produce nothing alone while the X variable can produce alone. This is the situation shown in Figure 37. Along the X axis constant returns to scale prevail while along the Y axis product is zero. OM is the Y

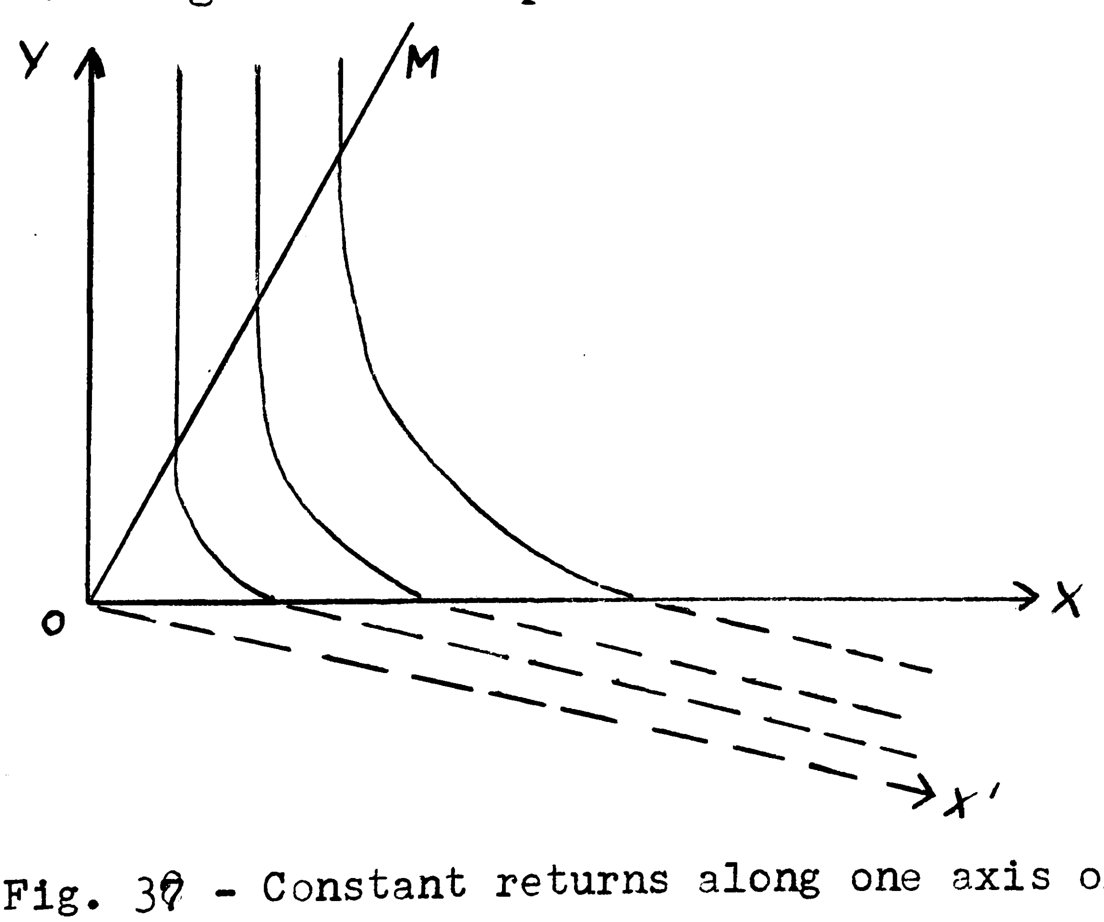


Fig. 37 - Constant returns along one axis only.

ridge line while OX can be considered to be the X ridge line. OX is then the locus of all points where the isoquants turn parallel to the geometrical axis OX', representing an economically meaningless variable. The relevant part of the map is contained by OM and the X axis, and the part of the map derivable from the hyper figure is contained by the X and Y axes. Nothing lying below the X axis is derivable from the hyper figure.

Total product increases as Y alone is increased from zero, but it increases at a diminishing rate[21] up to OM. When X alone is increased total product increases at a constant rate until OM is reached. Past OM total product increases either at a constant or diminishing rate,[22] but never reaches a maximum.

This digression on the linear homogeneous production function and the isoquant maps which derive from its geometrical representation reveals how restrictive the subject of constant returns to scale is for analysis. There is available for this type of analysis only those plane cross-sections of the homogeneous hyper figure which pass through the origin. Such cross-sections are few in comparison to the total number of cross-sections which can be taken from the figure. Furthermore, many of the plane cross-sections taken through the origin of the

---

21 This is because the Y function coefficient must be less than one.

22 If total product increases at a constant rate, the total product curve consists of two straight lines meeting at the X value, which, for the given Y value, lies on OM. The first straight line passes through the origin and has a steeper slope than the second straight line which makes up the other part of the total product curve.

hyper figure are blank, so that the number of isoquant maps reflecting constant returns to scale is definitely limited. Though these few isoquant maps exist conceptually, they may not exist at all in a realistic sense. It is a matter of controversy whether or not the production function can be linear and homogeneous. Although it is argued here that if there is no indivisibility, and all inputs are taken into account, the production function is homogeneous of degree one, it is a rather meaningless argument. "With regard to the first point, it is clear that labeling the absence of homogeneity as due to indivisibility changes nothing and merely affirms by the implication that 'indivisibility' does exist, the absence of homogeneity."[23] With regard to the second point, "it is a scientifically meaningless assertion that doubling all factors must double product. This is so not because we do not have the power to perform such an experiment.... Rather the statement is meaningless because it could never be refuted, in the sense that no hypothetically conceivable experiment could ever controvert the principle enunciated. This is so because if product did not double, one could always conclude that some factor was 'scarce'."[24]

Therefore, contour analysis of production theory should not devote much space to the relatively few iso-product maps which reflect constant returns when the vast majority of

---

23 Samuelson, Foundations of Economic Analysis, p84.

24 Ibid.

isoquant patterns are of the type reflecting varying returns. The analysis of the former should be relegated to a subordinate position in the more general analysis of the latter.

APPENDIX TO CHAPTER XIII

A NOTE ON SOME CONTROVERSIAL ISOQUANT PATTERNS

In Chapter XIII the point was made that if the influence of the economic unit in the productive process vanishes, and if all inputs are varied together in proportion, complete divisibility being assumed, it necessarily follows that the production function is linear and homogeneous. This is the position generally accepted by most economists. However there are a few who do not accept it, claiming that under the above assumptions returns to scale need not be constant when all inputs are varied in proportion.[1] There is then a body of analysis covering a type of isoquant pattern not discussed in Chapter XIII, and which is not too clearly discussed by those who support this analysis. It is the purpose of this note to consider such isoquant maps.

In deriving these maps it is assumed, as in Chapter XIII, that the fixed contribution of the economic unit to production is negligible. Also the production function is taken to be a continuous function of continuous variables, with all the necessary variables being accounted for. It is then supposed

---

1 Cf. E.H. Chamberlin, "Proportionality, Divisibility and Economics of Scale", Quarterly Journal of Economics, Feb. 1948. Also J. Lerner, "Constant Proportions, Fixed Plant and the Optimum Conditions of Production", Quarterly Journal of Economics, Aug. 1949.

Professor Boulding may also be said to adhere to this view since he introduces a non-linear isocline into a diagram conforming to all the assumptions taken here as implying constant returns. Cf. Economic Analysis, 1st ed, p507, Fig. 65. On the same kind of indirect evidence, Professor Allen may also be accused of holding the heretical view. Cf. Mathematical Analysis, pp284-7, Figure 78.

that in spite of these conditions the production function is non-homogeneous. The desired isoquant maps are then derived from the hyper figure associated with the non-homogeneous production function. Plane cross-sections of this figure taken through its origin, in the manner described in Chapter XIII, yield isoquant patterns which do not reflect constant returns to scale.[2]

Since constant returns to scale are ruled out, two possibilities[3] may arise when the variables of these isoquant maps are varied in a given proportion along a straight line through the origin. For such a variation in the variables there may eventually be diminishing returns to scale or else there will be increasing returns to scale for all output levels. In the former case there will be a single point of maximum product if the production function is so defined that all units of an input quantity must be used. At this maximum point the ridge lines intersect each other. This possibility is illustrated in Figure 38A, overleaf. The ridge lines are the pair of intersecting curves passing through the points O, M and D, with M being the single point of maximum product. The ridge lines, as well as all the other

---

2 These isoquant patterns are not the same as those yielded by the cross-sections of the hyper figure of Chapter IX when these sections are taken through its origin. This is because the economic unit makes some contribution to production in the former case. Here the economic unit does not take part in the productive process.

3 Omitting the trivial possibility of blank cross-sections.

isoclines, can not be straight lines. They must however converge on the origin as well as the point of maximum product.

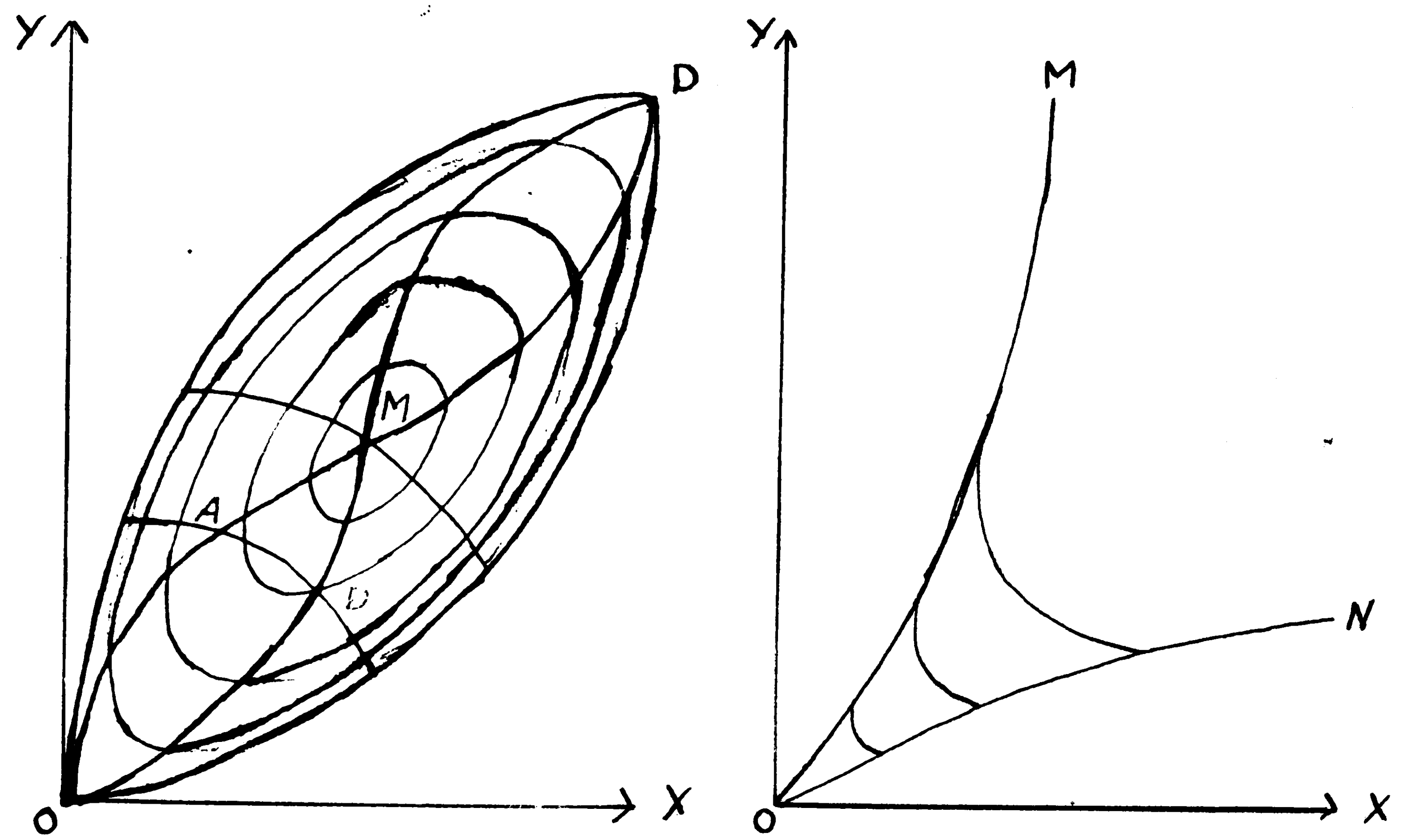


Fig. 38 - Isoquants for varying returns to scale.

They also converge on the point D which is a point of minimum output similar to O. Therefore the closed isoquant passing through the points O and D is the iso-product curve denoting zero product. The curve passing through the points A and D and cutting the isoquants is the locus of all points where returns to scale are constant.[4] The curve passing through the point M and cutting the isoquants is the locus where returns to scale cease. Thus returns to scale are increasing in the area of the map lying between the origin and the locus AD. Between AD and M returns to scale are decreasing. Therefore the relevant area of this map is contained by the

4 i.e., where the elasticity of production is equal to unity.

triangle AMD.

The possibility that returns to scale may be increasing for all output levels is illustrated in Figure 38B. Since there is no point of maximum product the ridge lines must curve in the manner illustrated by OM and ON because they can not be straight lines. This pattern, along with the one given by 38A, covers all the convexed curve systems yielded by plane cross-sections taken through the origin of the given hyper figure. Those who argue that the production function is not linear and homogeneous when all inputs are varied in proportion and complete divisibility exists are confined to these two types of isoquant maps. Their analysis is accordingly restricted.[5]

---

5 In the above situations all the isoclines must have positive slopes when the ridge lines have positive slopes. Only if the ridge lines take on a negative slope may a portion of some of the isoclines have a negative slope. It may be possible to draw Figure 38A in such a manner that the ridge lines in the vicinity of M have negative slopes. If this is so it is a very limited and exceptional phenomenon. Yet Mr. Lerner, "Constant Proportions, Fixed Plant and the Optimum Conditions of Production", op. cit., pp368-9, Fig. IV, draws into his isoquant map reflecting varying returns to scale a backward sloping isocline. He takes this to be a common occurrence if constant returns to scale do not exist under the usual conditions. As the above analysis indicates a backward sloping isocline is not a common occurrence.

CHAPTER XIV

APPLICATIONS OF EXTENSIONS

It now remains to investigate applications of the added knowledge of isoquants acquired in extending the analysis. The investigation is conducted in the following manner. First, comparison is made of the geometrical representations of indifference analysis and of the corresponding representations of isoquant analysis. Secondly, with the use of the above comparison, consideration is given to the conditions under which all the applications of the extended indifference analysis, considered in Chapter VII, also hold for the extended isoquant analysis. Finally, applications of the extended isoquant analysis are discussed under conditions for which all the indifference applications do not hold.

In applying the extended analysis of isoquants to production theory, it is helpful to draw first some general comparisons between the two types of contour analysis considered in this thesis. It should now be clear that there is a wider variety of isoquant maps than indifference maps. This is due mainly to the fact that while there are two kinds of hyper production figures there is only one kind of hyper preference figure.[1] There is the kind of hyper production figure associated with an economic unit contributing to

---

1 That type of hyper preference figure, discussed in Chapter VI, which is defined to exist for a period of time instead of for a moment of time is ignored here. The cross-sections derived from it do not differ materially from the more normal type of indifference maps.

production and there is the type associated with an economic unit not contributing to production. The hyper preference figure is associated with only one type of economic unit (individuals). It is not surprising therefore that the special type of cross-sections derived from the second of the above two mentioned hyper production figures, and discussed in Chapter XIII, has no counterpart in indifference analysis. Between the general indifference maps of Chapter IV and the isoquant maps of Chapter XIII there are no major points of comparison — there are only points of contrast.

All commodity plane cross-sections of the hyper preference figure have counterparts in the cross-sections of the non-homogeneous hyper production figure. The general isoquant map of Chapter XI is similar in all major aspects to the typical indifference maps of Chapter IV. It is this type of isoquant map which is used here in applying the extended isoquant analysis to production theory. All other isoquant patterns are ignored since they are only special cases.

Given the isoquant pattern, it is assumed in production theory that the economic unit associated with the hyper production figure attempts to operate at that level, designated by a point on the given map, which maximizes profit. This optimum point of operations, for the given isoquant pattern, depends on relative prices of the two variable inputs of the map and the price of the product. It always lies somewhere within the effective region of the map but, under the usual assumption of production theory, it is impossible to fix its exact location.[2] The usual assumption is that the economic unit is always able to reach the optimum scale of operations. This implies that there are no restrictions on outlay for X and Y. The value of this outlay varies with changes in relative prices or in product price. However, the equilibrium point always coincides with the optimum point. Contrast this with the situation in demand theory.

The individual is assumed to maximize something called utility subject to the restraints of a fixed outlay called his income. It is assumed that his income is not sufficient to permit him to attain the optimum point on his given indifference map. The location of this point is definite and fixed since it is given by the point of intersection of the ridge lines and is invariant to price changes. The equilibrium point does not coincide with the optimum point. However, since expenditure on X and Y is fixed the equilibrium point can be located on the indifference map once the prices of X and Y and the size of the fixed outlay is given. It is therefore obvious that if producer analysis is to parallel consumer analysis the expenditure of the economic unit on X

---

2 The exact location is given by the point of tangency between the iso-cost curve denoting the equilibrium outlay on X and Y and an isoquant. Since this outlay is not fixed in advance there are several iso-cost curves instead of just one. Therefore, mere consideration of the isoquants and iso-cost curves is not sufficient to determine the optimum point of operations.

and Y must be fixed so as to impose some restrictions on its level of operations.[3]

If such an effective limit on expenditures exists then the economic unit must budget its payments,[4] and the analysis of the consequences is analogous to demand analysis. The equilibrium point can be located with the aid of the isoquant map since outlay is fixed. Movements in the equilibrium point can be broken

---

3 Expenditure will be so fixed if the economic unit has an absolute money capital limit. Capital rationing undoubtedly exists in the real world and acts to prevent firms from attaining their optimum scale of output. Certainly in agriculture the nature of the risks involved and lack of capital prevents many farms from reaching their optimum scale. Cf. R.W. Rudd and D. L. MacFarlane, "The Scale of Operations in Agriculture", Journal of Farm Economics, V.XXIV, No. 2, May 1942.

If the limit on the scale of operations is sufficient to cause the equilibrium point to lie outside of the effective region of the isoquant map, there will be increasing returns when the two variable inputs are varied together although at the same time there may be decreasing returns to both inputs when varied singly. Thus the function coefficients of X and Y will both be less than one but their sum will be greater than one. In this situation then the effects of restrictions on the scale of operations will be reflected in the sum of these function coefficients. This suggests the analytical framework for the empirical testing for the presence of restrictions on investment in a competitive industry.

4 Budgeting is necessary because the productive process takes time. This consumption of time imposes an effective limit on the use of resources.

down into a substitution effect and a "resource effect",[5] this latter term applying to the effect which parallels the income effect of demand theory. Production analysis exactly parallels demand analysis. All applications of the extended indifference analysis to demand theory can be duplicated here for the application of the extended isoquant analysis to production theory.

When the two types of analyses are similar the demonstration of the possibility of inferior inputs follows along the same lines as the discussion of inferior goods. A backward sloping isocline indicates that for the appropriate relative costs, and scale of operations, larger outputs result in an absolute diminution in the amount of one input used. This may be the case where a larger output is obtained by technical methods such that certain types of labor or certain machines, which are better for smaller outputs, are diminished in quantity.[6] Also the possibility arises that the point of equilibrium may lie on one of the axes. At

---

5 Cf. H. Makower and W.J. Baurnol, "The Analogy Between Producer and Consumer Equilibrium Analysis", Economica, Feb. 1950, pp63-80. The derivation of the fundamental equation for this restrained maximum problem in production theory is given here. It is similar in composition to the corresponding fundamental equation of value theory.
Of course, if the limit imposed on the economic unit's use of resources is not effective so that the point of maximum profit can always be attained, then the "resource effect" does not appear and the analysis ceases to be parallel to demand analysis.

6 Cf. J. Lerner, "Constant Proportions, Fixed Plant and the Optimum Conditions of Production", op. cit., pp368-9.

such a point the ratio of the marginal productivities does not equal the price ratio. This situation is only possible when there are restrictions to scale since, otherwise, the boundaries of the effective region lie inside of the axes. Finally, demand curves for inputs can be derived in the same manner as demand curves for consumer goods. All the various varieties of demand curves for consumer goods can also be reproduced here for inputs.

All these uses of the isoquant analysis vanish when producer analysis is carried out on the textbook assumption that there are no restrictions on the resources available for production. When the optimum scale of operations can always be attained two-dimensional isoquant analysis can not, by itself, reveal how the demand for inputs vary as prices change. By resorting to algebra it can be proved that, in the absence of restrictions, the demand for an input, say X, <u>always</u> increases as its price falls. The demand for Y may either increase or decrease as the price of X falls. Also, output may either increase or decrease with a fall in the price of X.[7]

The only utility of the isoquant analysis, under the textbook assumptions, lies in its use to illustrate the expansion path of the firm. This path indicates how a firm expands its scale of operations under the assumption that it attempts to produce each successive output at the minimum

7 Cf. Hicks, <u>Value and Capital</u>, p93 and Mathematical Appendix, pp319-323. These results are due to the fact that the fundamental equation for production theory now has one less term than the corresponding fundamental equation of value theory.

cost. The validity of this assumption for portraying conditions in the real world has come under doubt[8] thus making questionable the usefulness of the concept of the expansion path. However, as the exact shape of the expansion path, under various conditions is not completely developed in the literature, it may be instructive to work out the details with the aid of the extended isoquant analysis.

---

8 Cf. J. Dingwall, "Equilibrium and Process Analysis in the Traditional Theory of the Firm", The Canadian Journal of Economics and Political Science, V.X,No.4, p431. Dr. Dingwall argues that in the expansion process the relevant curves are not the iso-product (or iso-revenue) curves but the iso-profit curves. The firm may expand in any manner whatsoever providing only that it cuts higher iso-profit curves. The expansion path of traditional analysis is then only a very special path based on a rather arbitrary assumption.

It seems to have escaped notice that as early as 1911, long before the conception of the expansion path had been worked out, Edgeworth examined the entire question of the manner in which an entrepreneur would proceed to the peak of his profit mound. Cf. "Contributions to the Theory of Railway Rates", Economic Journal, 1911, pp362-8. Here he discusses the expansion path of minimum costs and the two special paths considered by Dingwall which are derived by taking one variable input to be more fixed than the other one. These two special paths are the ridge lines of the profit mound. Edgeworth also considers a third kind of path, not considered by Dingwall, which is determined by the preference directions of the profit mound. Concerning these three types of expansion paths, he concludes, as does Dr. Dingwall, that: "A priori, I think it is not possible to say which of these types best represent the working of the managerial mind." (ibid., p367).

The expansion path is derived in the same way as the income curve of indifference analysis. When perfect competition prevails in both factors and product markets the derivation of both types of loci is similar. It is therefore only necessary to repeat here briefly the discussion on the income curve as it applies to the expansion path. For most output levels the expansion path is an isocline. The whole system of isoclines converge on the point of maximum product, while through any point on the axes, including the origin, only one isocline passes. So the expansion path starts from the origin. However, only for the cost ratio which defines the isocline passing through the origin does it coincide entirely with an isocline for all levels of output. Along this special expansion path amounts of both inputs are employed for every scale of operations. If the cost ratio differs from the ratio which defines this path, the expansion path coincides with one of the axes for a small range of output, and only one of the variable inputs is used in this range. If factor costs are not equal, the input used is most likely the cheapest one. If factor costs are equal, the most productive input is the sole one used. The typical expansion path then coincides with a portion of an axis and with an isocline and has a positive slope. However, when the inputs are partial substitutes the section of the expansion path coinciding with an isocline may have a negative slope.

This concludes all that need be said on the applications of isoquant analysis. The discussion has been brief because of the fact that, when production theory is formulated in the same manner as demand theory, the applications of indifference analysis holds in toto for isoquant analysis. Thus, most of the work of the chapter has already been done elsewhere. If production theory is not treated in the same manner as demand theory, then the application of isoquant analysis shrinks to the single use of illustrating the expansion path of the firm.

CHAPTER XV

SUMMARY AND CONCLUSIONS TO PART II

The consideration of the extensions of contour analysis as it applies to production theory comes to an end with the discussion of possible applications of the extensions. In summarizing the work done, two major consequences of the extensions can be noted.

The foremost consequence of extending isoquant analysis is the development of a complete and general view of two-dimensional iso-product maps and analysis. From this a full description of a typical isoquant map emerges. Although this description has not received much attention in the literature, it is the most relevant one for the theory of the firm. Most economic texts have been content to deal only with the special isoquant map reflecting constant returns to scale. A complete description of this special case is also considered here, a description which reveals the highly restricted notions of the text-book example. The full development of the isoquznt analysis also makes it possible to account for many unusual and apparently unrelated isoquant patterns, and enables the phenomonen of indispensability to be placed within its proper context.

The second major consequence of extending the analysis is the possibility it affords for obtaining a criterion, based on a property of the isoquant map, for defining input relations. The literature contains very little on this subject. Through the extension of the analysis it is possible to develop an analytical definition of technical substitution and complementarity on a basis which gives the same interpretation to these terms as is given to commodity substitution and complementarity.

The contributions of these consequences, when applied to production theory, suitably formulated, are similar to the contributions which follow from the applications of extended indifference analysis. If production theory is so formulated that these contributions do not apply the consequences are still useful since they render possible a clearer devolopment of the conception of the expansion path. However, the main contribution of the extensions lies in the better conceived analytical tool which emerges.

GENERAL CONCLUSIONS

A general survey of the ground covered in this study reveals one major idea linking the two seperate parts of the thesis. This idea is that the assumptions which underly the so-called relevant areas of indifference and isoquant maps should be extended to all sections of these maps, with a view to completing the existing analysis in a logically consistent manner. Apart from the logical aspects the idea is worthwhile on other accounts.

For one thing, extending the analysis has brought out many implicit assumptions and properties concerning contour maps, non-recognition of which can, and does, create analytical confusion. Then, in many cases, without a consideration of irrelevant areas, it is impossible to distinguish between varieties of isoquant maps or to draw comparisons between them and indifference maps. Also, it becomes possible to develop a definition of commodity and input relations by using the one criterion of ridge line slopes in both cases. This makes the two types of definitions comparable, which is desirable since the terms substitution and complementarity then mean the same thing in both demand and production theory. There is no need for recasting concepts in moving from one field to the other. However, the main utility of the discussion and consideration of irrevelant areas lies in the formulation of keener and sharper tools of analysis which result, and the consequences thereof. The more fully conceived indifference map makes it possible to illustrate and clear up many obscure and hastily formulated propositions of demand theory. The more detailed isoquant analysis serves the same purpose for production theory. But the mere presentation of more completely developed analytical tools is an end in itself of no little value, since a proper use of geometrical tools requires a knowledge of all parts and not just of the relevant parts.

Therefore, on the basis of this thesis it can be concluded that economists would be well advised to develop their analytical devices in a more precise and complete manner than is their usual custom. To stop at the bounds of so-called relevance in economic tool making and to concentrate entirely on the effective parts of the tools is an open invitation to muddled and confused methods of analysis. The main lesson to be drawn from this thesis is that "logical extemes are often illuminating in the field of Economic Theory".

BIBLIOGRAPHY

I - Bibliographical Note

A few of the works listed in this bibliography were not used in developing this thesis, and many were referred to only very briefly. Therefore, an asterisk (*) has been placed in front of those sources which played an essential role in the shaping of the ideas presented here or which proved of particular value. The list of works given below is meant to form a selected bibliography on the subject of contour analysis.

This selected bibliography is divided into three sections. The first séction includes those general works which contain material on both indifference and isoquant analysis. The second and third sections contain, respectively, references on indifference curves and on isoquants.

The following abreviations are used:

Am. Ec. Rev.	American Economic Review
C. J. E.	Canadian Journal of Economics and Political Science
E. J.	Economic Journal
J. P. E.	Journal of Political Economy
Q. J. E.	Quarterly Journal of Economics
Rev. Ec. Stud.	Review of Economic Studies

II - Selected Bibliography

GENERAL

*Allen, R.G.D., Mathematical Analysis for Economists, London, 1947.

*Boulding, K., Economic Analysis, New York and London, 1941.

*Hicks, J.R., Value and Capital, 2nd ed, Oxford, 1946.

Samuelson, P.A., Foundations of Economic Analysis, Cambridge, 1948.

Stigler, G., The Theory of Prices, New York, 1947.

INDIFFERENCE CURVES

Allen, R.G.D., "The Nature of Indifference Curves", Rev. Ec. Stud., 1934, I, 110-121.

_____, "A Note on the Determinateness of the Utility Function", Rev. Ec. Stud., 1934-35, II, 155-8.

_____, "Professor Slutsky's Theory of Consumer's Choice", Rev. Ec. Stud., 1936.

*_____, "A Comparison between Different Definitions of Complementary and Competitive Goods", Econometrica, 1934, II, 168-75.

_____, and Bowley, A.L., Family Expenditures, London, 1935.

*_____, and Hicks, J.R., "A Reconsideration of the Theory of Value", Economica, 1934, XIV, 52-76, 196-219.

Bernardello, H., "The End of the Marginal Utility Theory?", Economica, 1938.

_____, "A Note on the Determinateness of the Utility Function", Rev. Ec. Stud., 1934-35, II, 69-75.

*Bishop, R.L., "Professor Knight and the Theory of Demand", J. P. E., 1946, 141-157.

*Clark, J.M., "Realism and Relevance in the Theory of Demand", J. P. E., 1946, 352-3.

Davis, H.T., The Theory of Econometrics, Bloomington, 1941.

Edgeworth, F.Y., Papers Relating to Political Economy, I, 117.

Fisher, Irving, "Mathematical Investigations in the Theory of Value and Prices", Transactions of the Connecticut Academy of Arts and Science, 1892-95, IX, 1-124.

Georgesue-Roegen, N., "The Pure Theory of Consumer's Behavior", Q. J. E., 1935-36, L, 545-593.

Hart, A.G., "Pecularities of Indifference Maps Involving Money", Rev. Ec. Stud., 1941, VIII, 126-128.

Houthakker, H.S., "Revealed Preference and the Utility Function", Economica, 1950, XVII, 159-174.

Johnson, W.E., "The Pure Theory of Utility Curves", E. J., 1913.

*Knight, F.H., "Realism and Relevance in the Theory of Demand", J. P. E., 1944, LII, 289-318.

Lange, O., "The Determinateness of the Utility Function", Rev. Ec. Stud., 1934, I, 218-25.

Lerner, A.P., "The Diagramatical Representation of Demand Conditions in International Trade", Economica, 1934, 319-334.

Mosak, J.L., "On the Interpretation of the Fundamental Equation on Value Theory", Studies in Mathematical Economics and Econometrics, in Memory of Henry Schultz, (ed. O. Lange, F. McIntyre and T.O. Yntema) Chicago, 1942.

Norris, R.T., The Theory of Consumer's Demand, New Haven, 1941.

Samuelson, P.A., "Constancy of the Marginal Utility of Income", Studies in Mathematical Economics and Econometrics, in Memory of Henry Schultz,

_____, "Consumption Theory in Terms of Revealed Preference", Economica, 1948, XV.

*Schultz, H., The Theory and Measurement of Demand, Chicago, 1938.

de Scitovszky, T., "A Reconsideration of the Theory of Tariffs", Readings in the Theory of International Trade, Philadelphia and Toronto, 1949.

Thurstone, L.L., "The Indifference Function", Journal of Social Psychology, 1931, II, 139-167.

Wallis, W.A. and Friedman, M., "The Impirical Derivation of Indifference Functions", Studies in Mathematical Economics and Econometrics, in Memory of Henry Schultz, (ed. Lange et al) Chicago, 1942.

Weintraub, S., "The Foundations of the Demand Curve", Am. Ec. Rev., XXXII, 1942.

Weldon, J.C., "A Note on Measures of Utility", C. J. E., XVI, 227-233.

ISO-PRODUCT CURVES

Baurnol, W.J., and Makower, H., "The Analogy between Producer and Consumer Equilibrium Analysis", Economica, 1950, XVII, 63-80.

*Black, J.D., Introduction to Production Economics, 1926.

Bober, M.M., "Theoretical Aspects of the Scale of Production", Economics, Sociology and the Modern World, Essays in Honor of T.N. Carver, (ed. N.E. Himes) Cambridge, 1935.

Boulding, K., "The Theory of the Firm in the Last Ten Years", Am. Ec. Rev., 1942, XXXII.

Carlson, S., A Study in the Pure Theory of Production, London, 1939.

*Cassels, J.M., "On the Law of Variable Proportions", Readings in the Theory of Income Distribution, Philadelphia and Toronto, 1946.

Chamberlin, E.H., "Proportionality, Divisibility and Economies of Scale", Q. J. E., 1948.

Dingwall, J., "Equilibrium and Process Analysis in the Traditional Theory of the Firm", C. J. E., 1944, X, 448-463.

Edgeworth, F.Y., "Contributions to the Theory of Railway Rates", E. J., 1911, XXI, 362-368.

Hahn, F.H., "Proportionality, Divisibility, and Economies of Scale: Comment", Q. J. E., Feb. 1949.

Keirstead, B.S., The Theory of Economic Change, Toronto, 1948.

Knight, F.H., Risk, Uncertainty, and Profit, Boston, 1921.

Lerner, J., "Constant Proportions, Fixed Plant and the Optimum Conditions of Production", Q. J. E., Aug. 1949.

Stigler, G., "Production and Distribution in the Short Run", Readings in the Theory of Income Distribution, Philadelphia and Toronto, 1946.