

CAPITAL, GROWTH AND WELFARE: ESSAYS IN THE THEORY OF ALLOCATION

by

FRANCIS M. BATOR

S.B., Massachusetts Institute of Technology, 1949

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August 20, 1956

Signature redacted

Signature of Author: _____

Signature redacted

Certified by: _____

Signature redacted

Accepted by: _____

Chairman, Departmental Committee
on Graduate Students

Professor Ralph E. Freeman
Chairman,
Department of Economics and Social Science
Massachusetts Institute of Technology
Cambridge, Mass.

Dear Professor Freeman:

I herewith submit, in partial fulfillment of the requirements for the degree of Doctor of Philosophy, a dissertation entitled Capital, Growth and Welfare: Essays in the Theory of Allocation.

Sincerely yours,
Signature redacted

Francis M. Bator

August 20, 1956
Cambridge, Mass.

CAPITAL, GROWTH AND WELFARE: ESSAYS IN THE THEORY OF ALLOCATION

Francis M. Bator

Submitted to the Department of Economics and Social Science
on August 20, 1956, in partial fulfillment of the
requirements for the degree of Doctor of Philosophy.

The dissertation consists of three essays in the theory of resource allocation. The first, entitled "The Simple Analytics of Welfare Maximization", is a complete treatment, in six parts, of the problem of welfare maximization in its "new welfare economics" aspects. Part I is a rigorous diagrammatic determination of the "best" configuration of inputs, outputs, and commodity-distribution for a two-input, two-output, two-person situation, where all functions are of smooth curvature and where neo-classical generalized diminishing returns obtain in all but one dimension -- returns to scale are assumed constant. A geometric device is suggested which simplifies one step of the derivation. Part II identifies the "price-wage-rent" configuration embedded in the maximum problem which would ensure that decentralized profit and preference maximizing behavior by atomistic competitors would lead to the maximum-welfare position. Part III explores the requirements on initial factor ownership if market-imputed income distribution is to be consistent with the commodity-distribution required by the maximum-welfare solution. Part IV, consists in comments on some technical ambiguities, e.g. the presumption that all tangencies are internal; also on a number of feasible extensions: more inputs, outputs and households; elasticity in input supplies; joint and intermediate products; external interactions. The discussion is still stationary and neo-classical in spirit. Then, in Part V, the consequences of violating some of the neo-classical curvature assumptions are examined. Attention is given to the meaning, in a geometric context, of the "convexity" requirements of mathematical economics and to the significance of one important type of non-convexity: increasing returns to scale, for "real" market allocation, for Lange-Lerner type "as if" market allocation, and for the solubility of a maximum-of-welfare problem. Part VI contains some brief remarks on possible dynamical extensions. A historical note on the seminal literature is appended.

Library (Don) File 21.1736

The second essay, "Elements of the Pure Economics of Social Overhead Capital", is an attempt to sort out the various mutually reinforcing and overlapping elements that underlie the notion of "social overhead": indivisibility, durability, "external" interaction, non-appropriability, the "public", non-exhaustible quality of the output ... It consists of four parts. Part I is an introductory

examination of the various qualities by which "social overhead" capital has been identified and of the criteria against which the consequences of social overhead phenomena are to be tested. Part II explores the content of what is the most significant defining quality: increasing returns to scale and/or indivisibility, and traces its implications for efficient allocation in a stationary and static context of perfect information. Part III consists, first, in an expository digression on the doctrine of external economies; second, it suggests a particular ordering of externality phenomena. This last is designed to clarify the ways in which externalities associate with "social overhead" and to identify the links between the externality and the indivisibility aspects. Use is made, in the ordering, of the recently revived "public good" notion of public expenditure theory. Part IV suggests directions for future research.

The third essay, "On Capital Productivity, Input Allocation and Growth", consists in an application of allocation theory to some problems of economic growth. Specifically, Parts I and II treat some often misunderstood aspects of the relationship between the interest rate, the marginal productivity of capital and "the" capital-output ratio. In Part III, the problem of optimal input combinations is subjected to scrutiny. The specific issue concerns the desirability of capital intensive processes in countries whose capital is scarce relative to labor. Part IV consists in some general comments on input allocation as a maximizing problem and on the strategy of theorizing about economic growth.

The mode of analysis, throughout the three essays, is that of modern allocation theory: it provides the analytical techniques and fixes -- for better or worse -- the level of abstraction. It is the link by which the essays are joined.

Thesis Supervisor: Robert M. Solow
Associate Professor
of Economics

ACKNOWLEDGMENTS*

This dissertation was undertaken in connection with my research responsibilities as a member of the staff of the Center for International Studies, M.I.T. I am grateful to the Center, and in particular to Prof. M. F. Millikan, for making this work possible.

My outstanding personal debt is, again, to Mr. Millikan. I am much indebted, also to Prof. R. M. Solow, who was unstinting with his time and thought and support, and to my friend and colleague Prof. R. S. Eckaus. Finally, the work itself attests my intellectual obligation to Professor Paul Samuelson.

* I should like to acknowledge, also, the superior services of Miss Regina Winder, in connection with the manufacture of the typescript.

TABLE OF CONTENTS

<u>Chapter One</u>	The Simple Analytics of Welfare Maximization . . .	page 1
Part I	Inputs, Outputs and Commodity Distribution	2
	1. From Endowments and Production Functions to the Production Possibility Curve	4
	2. From the Production Possibility Curve to the Utility Possibility Frontier	5
	3. From the Utility Possibility Frontier to the Constrained "Bliss Point"	9
	4. From "Bliss Point" to "Best" Inputs, Outputs and Commodity Distribution	11
Part II	Prices, Wages and Rents	12
Part III	Income-Distribution and Initial Endowments	17
Part IV	Some Extensions	21
	1. Many Households, Many Goods, Many Factors	21
	2. Corner Tangencies	21
	3. Leisure-Work Choices	23
	4. Joint Products	24
	5. Intermediate Products	24
	6. Diminishing Returns to Scale	24
	7. External Interaction	27
Part V	Non-Convexities	30
	1. Kinks and Corners	31
	2. Convexity Defined	31
	3. Increasing Returns to Scale	33
	a. and Average Cost Curves	34
	b. and Marginal Cost Curves	37
	c. and the Production Possibility Curve	39
Part VI	Dynamics	44
Appendix	A Historical Note on the Literature	50
Diagrams	Figures 1, 2-a,b, 3-a,b	53

<u>Chapter Two</u>	Elements of the Pure Economics of "Social Overhead Capital"	56
Part I	On Definition and Identification	61
	Test of Market Efficiency	62
	Identification	64
	The Efficiency Test Again	66
	What About Growth Potential ?	70
Part II	Indivisibilities and Increasing Returns to Scale	72
	A. Rudiments	72
	Durability	72
	Lumpiness	75
	More on Indivisibility	77
	B. Indivisibility and Production Functions	79
	C. Indivisibility, Efficiency and Markets	83
	"Apples and Bridges"	85
	Modifications	94
	Institutional Implications	98
	On Optimum Conditions Again	101
	Durability	104
	D. Indivisibility in Scale and Social Overhead	108
Part III	External Economies and Social Overhead	112
	A. A Digression on External Economies	112
	By Way of Some History	112
	The Modern Formulation: Apples and Honey	116
	Comments	122
	Some Queries	126
	B. Externalities: An Ordering	129
	Type 1: Organizational Externalities	129
	Type 2: Technical Externalities	133
	Type 3: Public Good Externalities	138
	Comments	141
	1. On Type (1)	141
	2. Prof. Meade's "atmosphere"	142
	3. Prof. Samuelson on Types (2) and (3)	143
	4. Exclusion	144
	5. Arrangements to offset	144
	6. On Blends	147
	C. Back to Social Overhead	147
Part IV	On the Economics of Social Overhead Capital: Agenda	150
Diagrams	Figures 1 - 11	158

<u>Chapter Three</u>	On Capital Productivity, Input Allocation and Growth	167
Part I	Interest Rate, Marginal Productivity of Capital and the Capital-Output Ratio	168
Part II	Interest Rate, Capital-Output Ratio and Distribution	178
Part III	Labor Productivity and Input Allocation	180
Part IV	Commentary	188
<u>Bibliography</u>	197

CHAPTER ONE

THE SIMPLE ANALYTICS OF WELFARE MAXIMIZATION

THE SIMPLE ANALYTICS OF WELFARE MAXIMIZATION

It appears, curiously enough, that there is nowhere in the literature a complete and concise non-mathematical treatment of the problem of welfare maximization in its "new welfare economics" aspects. It is the purpose of this exposition to fill this gap for the simplest statical and stationary situation.

Part I consists in a rigorous diagrammatic determination of the "best" configuration of inputs, outputs, and commodity-distribution for a two-input, two-output, two-person situation, where, furthermore, all functions are of smooth curvature and where neo-classical generalized diminishing returns obtain in all but one dimension -- returns to scale are assumed constant. A geometric device is suggested which simplifies one step of the derivation. Part II identifies the "price-wage-rent" configuration embedded in the maximum problem which would ensure that decentralized profit and preference maximizing behavior by atomistic competitors would lead to the maximum-welfare position. Again, primary reliance is on the diagrammatics of Part I. Part III explores the requirements on initial factor ownership if market imputed, or "as if" market imputed, income distribution is to be consistent with the commodity-distribution required by the maximum-welfare solution. Part IV, consists in brief comments on some technical ambiguities, e.g. the presumption that all tangencies are internal; also on a number of feasible (and not so feasible) extensions; more inputs, outputs and households; elasticity in input supplies; joint and intermediate products; external interactions. The discussion is still

stationary and neo-classical in spirit. Then, in Part V, the consequences of violating some of the neo-classical curvature assumptions are examined. Attention is given to the meaning, in a geometric context, of the "convexity" requirements of mathematical economics and to the significance of non-convexity for "real" market allocation, for Lange-Lerner type "as if" market allocation, and for the solubility of a maximum-of-welfare problem. Finally, Part VI contains some brief remarks on possible dynamical extensions. A note on the seminal literature concludes the paper.¹

I

Take, as given:

- (1) Two inelastically supplied, homogeneous and perfectly divisible inputs, labor-services (L) and land (D). This "Austrian" assumption does violate the full generality of the neo-classical model: elasticity in input supplies would make diagrammatic treatment impossible.
- (2) Two production functions, $A = F_A(L_A, D_A)$, $N = F_N(L_N, D_N)$, one for each of the two homogeneous goods: apples (A) and nuts (N). The functions are of smooth curvature, exhibit constant returns to scale and diminishing marginal rates of substitution along any isoquant (i.e. the isoquants are "convex" to the origin).
- (3) Two ordinal preference functions, $U_X = f_X(A_X, N_X)$ and $U_Y = f_Y(A_Y, N_Y)$ -- sets of smooth, convex to the origin indifference curves -- one for X and one for Y. These reflect unambiguous and consistent preference orderings for each of the two individuals (X & Y) of all conceivable combinations of own-consumption of apples and nuts. For convenience we adopt for each function an arbitrary numerical index: U_X and U_Y to identify the indifference curves. But the functions have no inter-personal implications whatever and for any one individual they only permit of statements to the effect that one situation is worse, indifferent or better than an other.

¹ Anyone familiar with the modern literature will recognize my debt to the writings of Prof. Samuelson. Reference is to be made, especially, to Chapter VIII of Foundations of Economic Analysis; to "Evaluation of Real National Income", Oxford Economic Papers, Jan. 1950; and to "Social Indifference Curves", Quarterly Journal of Economics, February 1956.

We do require consistency: if X prefers situation α to situation β and β to γ , then he must prefer α to γ ; indifference curves must not cross. Also, satiation type phenomena and Veblenesque or other "external" effects are ruled out.

(4) A Social Welfare Function, $W = W(U_X, U_Y)$, that permits a unique preference ordering of all possible states based only on the positions of both individuals in their own preference fields. It is this function that incorporates an ethical valuation of the relative "deservingness" of X and Y.

Our procedure is as follows: Taking the given fixed endowment of L and D and the production functions, we derive via an Edgeworth-Bowley "box diagram" the production-efficient locus of maximal production-possibilities. Then for each point on that production possibility curve (in output-space, with A and N on the axes), we derive, using the indifference functions and the same box diagram technique, the utility possibility curve in the $U_X - U_Y$ space associated with that point of specified apples & nuts. One can derive such a utility possibility curve for each point on the production possibility locus, i.e. for each feasible apple-nut combination, and then map out their frontier "envelope". By means of a simple but apparently novel¹ device, however, it is possible to find the envelope point associated with each apple-nut combination directly, without deriving the curves. At any rate, we derive an envelope to the utility-possibility curves associated with each production point: this envelope is the grand Pareto-efficient utility possibility frontier that characterizes the whole situation. Where this curve touches the highest contour of the Social Welfare Function, we have the maximum welfare position. From this, working backwards, we can identify the maximum-welfare values of all the variables: labor input into apples (L_A), labor input into nuts (L_N), land input into apples (D_A),

¹ Any claim of novelty, even for such a trivial little device, is risky. Perhaps I am simply revealing an ignorance of the literature. Most of the geometric techniques used are, of course, well-known.

land input into nuts (D_N), total production of apples (A) and nuts (N), the distribution of apples and nuts between X and Y (A_X, N_X, A_Y, N_Y). Further, in Parts II and III, we shall be able to identify the price-wage-rent and initial factor-distribution configurations that would make decentralized profit and preference maximizing behavior by atomistic competitors lead to such a maximum welfare position.

1. From Endowments and Production Functions to the Production Possibility Curve

Construct a rectangular "box", as in Fig. 1, with horizontal and vertical dimensions just equal to the given supplies, respectively, of L and D, and plot the isoquants for apples with the southwest corner as origin and those for nuts with origin at the north-east corner. (The isoquants for nuts appear flipped-over along the southwest-northeast diagonal). Every point in the box represents six variables, L_A, L_N, D_A, D_N, A, N . The problem of production-efficiency consists in finding that locus of points where any increase in the production of apples implies a necessary reduction in the output of nuts (and vice versa). The diagram shows that locus to consist in the points of tangency between the nut and apple isoquants (F-- F). Pick any feasible level of apple production then move along that apple isoquant until you reach the point of maximum nut production consistent with it -- that point will occur precisely at the point of tangency.

From this efficiency locus we can now read off the maximal obtainable combinations of apples and nuts and plot these in the

output (A-N) space. Given our curvature assumptions we get the nice smooth concave-to-the-origin Pareto-efficient production possibility curve F'-F' of Fig. 2-a. This locus, a consolidation of F-F in Fig. 1, represents input-output configurations such that the marginal rate of substitution (MRS) of labor for land in the production of any given amount of apples -- the absolute value of the slope of the apple isoquant -- just equals the marginal rate of substitution of labor for land in the production of nuts.¹

The slope (again neglecting sign) at any point on the production possibility curve of Fig. 2-a, in turn, reflects the marginal rate of transformation (MRT) at that point of apples into nuts. It indicates precisely how many nuts can be produced by transferring land and labor from apple to nut production (at the margin), with optimal reallocation of inputs in the production of both goods so as to maintain the MRS-equality requirement of Fig. 1. It is the marginal nut-cost of an "extra" apple -- or the reciprocal of the marginal apple-cost of nuts.

2. From the Production Possibility Curve to the Utility Possibility Frontier

Pick any point, δ , on the production possibility curve of Fig. 2-a: it denotes a specific quantity of apples and nuts. Construct an Edgeworth-Bowley (trading) box with these precise dimensions -- this is best done

¹ In marginal productivity terms, MRS, at any point, of labor for land in apple production -- the absolute value (drop all minus signs) of the slope of the apple isoquant (Fig. 1) -- is equal to

$$\frac{\text{Marginal Physical Product of Land}}{\text{Marginal Physical Product of Labor}} \quad \text{in apple production at that point.}$$

In the symbolism of the calculus $\left| \frac{\partial L_A}{\partial D_A} \right|_{\Delta A = 0} = \left(\frac{\partial A}{\partial D_A} \right) \div \left(\frac{\partial A}{\partial L_A} \right)$.

by dropping from \mathcal{S} lines parallel to the axes, as in Fig. 2-a: the resulting box will have just the desired dimensions. Then draw in X's and Y's indifference maps, one with the south-west the other with the north-east corner for origin. Every point in the box again fixes six variables: apples to X (A_X) and to Y (A_Y), nuts to X (N_X) and to Y (N_Y), and the "levels" of satisfaction of X and Y as measured by the ordinal indices U_X and U_Y which characterize the position of the point with respect to the two preference fields. For example, at λ in Fig. 2-a, $U_X = 300$, $U_Y = 200$. Note again, however, that this 200 is incommensurate with the 300: it does not imply that X is in some sense better off at λ than Y (or indifferent, or worse off).

The problem of "exchange-efficiency" consists in finding that locus of feasible points within the trading box where any increase in X's satisfaction (U_X) implies a necessary reduction in the satisfaction of Y, (U_Y). Feasible in what sense? In the sense that we just exhaust the fixed apple-nut totals as denoted by \mathcal{S} . Again, the locus turns out to consist of the points of tangency, S-S, and for precisely the same analytical reasons. Only now it is the marginal "psychic" rate of substitution of nuts for apples in providing a fixed level of satisfaction for X -- the absolute slope of X's indifference curve -- that is to be equated to the nut-apple MRS of Y, to the slope, that is, of his indifference curve.

From this exchange-efficiency locus, S-S,¹ which is associated with the single production point \mathcal{S} , we can now read off the maximal

¹ This is Edgeworth's "contract curve", or what Boulding has aptly called the "conflict" curve -- once on it, mutually advantageous trading is not possible and any move reflecting a gain to X implies a loss to Y.

combinations of U_X and U_Y obtainable from δ and plot these in utility (U_X-U_Y) space ($S'-S'$, Fig. 3-a). Each such point δ in output space "maps" into a line in utility space -- the U_X-U_Y mix is sensitive to how the fixed totals of apples and nuts are distributed between X and Y.¹

There is a possible short-cut, however. Given our nice curvature assumptions, we can trace out the grand utility possibility frontier -- the envelope -- by using an efficiency relationship to pick just one point from each trading box contract curve S-S associated with every output point δ . Go back to Figure 2-a. The slope of the production possibility curve at δ has already been revealed as the marginal rate of transformation, via production, of apples into nuts. The (equalized) slopes of the two sets of indifference contours along the exchange-efficiency curve S-S, in turn, represent the marginal rates of substitution of nuts for apples for psychic indifference (the same for X as for Y). The grand criterion for efficiency is that it be impossible by any shift in production cum exchange to increase U_X without reducing U_Y . Careful thought will suggest that this criterion is violated unless the marginal rate of transformation between apples and nuts as outputs -- the slope at δ -- just equals the common marginal rate of substitution of apples and nuts, as consumption "inputs", in providing psychic satisfaction.

¹ Each point in utility space, in turn, maps into a line in output-space. Not just one but many possible apple-nut combinations can satisfy a specified U_X-U_Y requirement. It is this reciprocal point-line phenomenon that lies at the heart of Prof. Samuelson's proof of the non-existence of community-indifference curves such as would permit the derivation of demand curves for apples and nuts. The psychic "community" MRS between A and N for given fixed A and N, e.g. at δ in Fig. 2-a, would surely depend on how the A and N are distributed, i.e., on which U_X-U_Y point on S-S is chosen. Hence the slope of a "joint" X-Y indifference curve at δ is not uniquely fixed by A-N. See Op. cit. #3.

If, for example, at δ one can get two apples by diverting resources and reducing nut-output by one, a point on S-S where the (equalized) marginal rate of substitution of apples for nuts along indifference curves is, e.g. one to one, permits the following "arbitrage" operation. Shift land and labor so as to produce two more apples and one less nut. Then, leaving X undisturbed take away one nut from Y and replace it by one apple. By our $MRS = 1$ assumption both X and Y are left indifferent: U_X and U_Y remain unaltered. But we have left over an extra apple; this permits raising U_X and/or U_Y , hence the initial situation was not on the U_X-U_Y frontier.¹

To be on the grand utility possibility frontier (B-B of Fig. 3-a), then, $(MRT)_{\delta}$ must equal the (equalized) MRS of the indifference contours along the S-S associated with δ . This requirement fixes the single U_X-U_Y point on S-S that lies on the "envelope" utility-possibility frontier, given the output point, δ . Pick that point on S-S, in fact, where the joint slope of the indifference curves is exactly parallel to the slope at δ of the production-possibility curve. A glance at Fig. 2-a will show this point to be at δ' . δ' gives the one "efficient" U_X-U_Y combination associated with the A-N mix denoted by δ . This U_X-U_Y combination can then be plotted as δ'' , in Fig. 3-a.²

¹ The above argument can be made perfectly rigorous in terms of the infinitesimal movements of the differential calculus.

² Never mind, here, about multiple optima. These could occur even with our nice curvature assumptions. If, for example, both sets of indifference curves show paths of equal MRS that coincide with straight lines from the origin and, further, if the two preference functions are so symmetrical as to give an $(S-S)_{\delta}$ that hugs the diagonal of the trading box, then either every point on $(S-S)_{\delta}$ will satisfy the MRS - MRT criterion, or none will. For discussion of these and related fine points see Parts IV and V.

Repetition of this process for each point on the production possibility curve -- note that each such point requires a new trading box -- will yield the grand utility possibility frontier of Pareto-efficient input-output combinations, B-B. Each point of this frontier gives the maximum of U_X for any given and feasible level of U_Y and vice versa. The frontier is a frontier in the sense that the situation precludes any (north-easterly) movement beyond it.

3. From the Utility Possibility Frontier to the "Constrained Bliss Point".

But B-B, the grand utility-possibility function, is a curve and not a point. Even after eliminating all combinations of inputs and outputs that are non-efficient in a Paretian sense, there remain a single-dimensional infinity of "efficient" combinations: one for every point on B-B. To designate a single best configuration we must be given a Bergson-Samuelson Social Welfare Function that denotes the ethic that is to "count" or whose implications we wish to study. Such a function -- it could be yours, or mine, or Mossadegh's, though his is likely to be non-transitive -- is intrinsically a-scientific.¹ There are no considerations of economic efficiency that permit us to designate Crusoe's function, which calls for many apples and nuts for Crusoe and just a few for Friday, as economically superior to Friday's. Ultimate ethical valuations are involved.

Once given such a Welfare-function, in the form of a family of social indifference contours in utility space, as in Fig. 3-b, the problem becomes

¹ Though it may provide the anthropologist or psychologist with interesting material for scientific study.

fully determinate.¹ "Welfare" is at a maximum where the utility-possibility envelope frontier B-B, touches the highest contour of the W-function. In Fig. 3-b, this occurs at Ω .²

Note the unique quality of that point Ω . It is the only point, of all the points on the utility frontier B-B, that has normative or prescriptive significance. Pareto-efficient production and commodity-distribution -- being on F'-F' and also on B-B -- is a necessary condition for a maximum of our kind of welfare function, but is not a sufficient condition.³ The claim that any "efficient" point is better than "inefficient" configurations that lie inside B-B is indefensible. It is true that given an "inefficient" point, there will exist some point or points on B-B that represent an improvement; but there may well be many points on B-B that would be worse rather than better. For example, in terms of the ethic denoted by the specific W-function of Fig. 3-b, Ω on B-B is better than any other feasible point. But the efficient point ξ is distinctly inferior to any inefficient point on or north-east of W_2 . If I am X, and my W-function, which reflects the usual dose of self-like, is the test, "efficient" B-B points that give a high U_Y and a very low U_X are clearly less

¹ These curves cannot be transposed into output-space. They are not community indifference curves which would permit the derivation of demand schedules. See f.n. on p. 7 above.

² If there are several such points, never mind. If the "ethic" at hand is really indifferent, pick any one. If it doesn't matter, it doesn't matter.

³ Note however, that Pareto efficiency is not even a necessary condition for a maximum of just any conceivable W-function. The form of our type function reflects a number of ethically loaded restrictions, e.g. that individuals' preference functions are to "count", and count positively.

desirable than lots of inefficient points of higher U_X .¹

4. From "Bliss Point" to "Best" Inputs, Outputs and Commodity-Distribution

We can now retrace our steps. To Ω on B-B in Fig. 3-b, there corresponds just one point, Ω' , on the production possibility curve $F'-F'$ in Fig. 2-b. (We derived B-B, point by point, from $F'-F'$). Ω' fixes the output mix: A and N. Then, by examining the trading box contract curve $S_{\Omega} - S_{\Omega}$ associated with Ω' of $F'-F'$, we can locate the one point on $S_{\Omega} - S_{\Omega}$ where U_X and U_Y correspond to the coordinates of Ω in utility space. The equalized slope of the indifference curves will at that point, Ω'' , just equal the slope of $F'-F'$ at Ω' . Ω'' fixes the apple-nut distribution implied by the maximum of W: A_X, A_Y, N_X , and N_Y . Further, we can now locate the point Ω''' on the Pareto-efficient input locus, F-F of Fig. 1, that corresponds to Ω' of $F'-F'$. It fixes the remaining variables, the factor allocations: L_A, D_A, L_N , and D_N . The maximum welfare configuration is determinate. We have solved for the land and labor to be used in apple and nut production, for the total output of apples and nuts, and for their distribution between X and Y.

¹ Note, however, that no consistency requirements link my set of indifference curves with "my" W-function. The former reflects a personal preference ordering based only on own-consumption (and, in the more general case, own services supplied). The latter denotes also values which I hold as "citizen", and these need not be consistent with maximizing my satisfaction "qua consumer". X as citizen may prefer a state of less U_X and some U_Y to more U_X and zero U_Y . There is also an important analytical distinction. X's preference function is conceptually "observable": confronted by various relative price and income configurations his consumption responses will reveal its contours. His W-function, on the other hand, is not revealed by behaviour, unless he be dictator. In a sense only a society, considered as exhibiting a political consensus, has a W-function subject to empirical inference (cf. the third Samuelson reference). The distinction -- it has a Rousseauvian flavor -- while useful, is of course arbitrary. Try it for a masochist; a Puritan ...

II

The above is antiseptically independent of institutional context, notably of competitive market institutions. It could constitute an intellectual exercise, engaged in by the often invoked man from Mars, in how "best" to make do with given resources. Yet implicit in the logic of this purely "technocratic" formulation, embedded in the problem as it were, are a set of constants which the economist will catch himself thinking of as prices. And wisely so. Because it happens -- and this "duality" theorem is the kernel of modern welfare economics -- that decentralized decisions in response to these "prices" by, or "as if" by, atomistic profit and satisfaction maximizers will result in just that constellation of inputs, outputs and commodity-distribution, that our maximum-of-W requires.¹

Can these constants -- prices, wages, rents -- be identified in our diagrammatic representations?² Only partially so. Two-dimensionality is partly at fault, but, as we shall see, a final indeterminacy is implied by the usual curvature assumptions themselves.³ The diagrams will, however, take us part way, and a little algebra will do for the rest.

The exercise consists in finding a set of four constants associated with the solution values of the maximum problem that have meaning as the price of apples (p_A), the price of nuts (p_N), the wage rate of labor (w),

¹ Note that this statement is neutral with respect to (1) genuine profit maximizers acting in real but perfectly competitive markets; (2) Lange-Lerner type bureaucrats ("take prices as given and maximize or Siberia"); or (3) technicians using electronic machines and trying to devise efficient computing routines.

² To avoid institutional overtones, the theory literature usually attempts verbal disembodiment and refers to them as shadow-prices. The mathematically oriented, in turn, like to think of them as Lagrangean multipliers.

³ These very assumptions render this last indeterminacy, that of the absolute price level, wholly inconsequential.

and the rental rate of land (r).¹

First, what can be said about w and r ? Profit maximization by the individual producer implies that whatever output he may choose as most lucrative, it must be produced at a minimum total cost.² The elementary theory of the firm tells us that for this condition to hold the producer facing fixed input-prices -- horizontal supply curves -- must adjust his input mix until the marginal rate of substitution (MRS) of labor for land just equals the rent to wage ratio. It is easy to see the "arbitrage" possibilities if this condition is violated. If one can substitute one unit of L for two units of D , and maintain output constant, with $w = \$10$ and $r = \$10$, it surely reduces total cost to do so and keep doing so until any further reduction in D by one unit has to be matched, if output is not to fall, by adding no less than one unit of L . In the usual diagrammatic terms, then, the producer will cling to points of tangency between the isoquants and (iso-expenditure) lines whose absolute slope equals $\frac{r}{w}$.

Reversing the train of thought, the input blend denoted by the point Ω''' in Fig. 1 implies a shadow $\frac{r}{w}$ ratio that just equals the MRS of labor for land in the production of both apples and nuts at that point Ω''' . $MRS_{\Omega'''}$ is given by the (equalized) slopes of the isoquants at Ω''' . The implicit $\frac{r}{w}$, therefore, must equal the slope of the line $R-W$ that is

¹ Since we are still assuming that all the functions have nice neo-classical curvature properties, hence that e.g. the production possibility curve, as derived, has to be concave to the origin, we can impose the strong condition on the constants that they exhibit optimality characteristics for genuine, though perfect, markets. It will turn out, however, that two progressively weaker conditions are possible, which permit of some non-convexities (e.g. increasing returns to scale), yet maintain for the constants some essentially price-like qualities. More on this in Part V.

² In our flow model, unencumbered by capital, this is equivalent to producing the chosen output with minimum expenditure on inputs.

tangent to (both) the isoquants at Ω'' .¹

The slope of R-W identifies the rent:wage ratio implied by the maximal configuration. Essentially analogous reasoning will establish the equalized slope of the indifference curves through Ω'' in Fig. 2-b, as denoting the $\frac{P_A}{P_N}$ ratio implied by the solution. X, as Y, to maximize his own satisfaction as measured by U_X , must achieve whatever level of satisfaction his income will permit at a minimum expenditure. This requires that he choose an apple-nut mix such that the psychic marginal-rate-of-substitution between nuts and apples for indifference just equal $\frac{P_A}{P_N}$. He and Y, will pick Ω'' only if $\frac{P_A}{P_N}$ is equal to the absolute slope of the tangent ($P_A - P_N$) at Ω'' . This slope, therefore, fixes the Ω -value of $\frac{P_A}{P_N}$.²

Note that this makes $\frac{P_A}{P_N}$ equal to the slope also of the production-possibility curve $F'-F'$ at Ω' .³ This is as it should be. If $\frac{P_A}{P_N} = 10$, i.e. if one apple is "worth" ten nuts on the market, it would be odd indeed, in our frictionlessly efficient world of perfect knowledge, if

¹ Again, absolute values of these slopes are implied throughout the argument. Recall from the footnote on p. 5 that the labor for land MRS: the absolute slope of the isoquants at Ω'' , as given by $\frac{ROA}{WOA}$, is equal to the $\left[\frac{\text{Marginal Physical Product of Land}}{\text{Marginal Physical Product of Labor}} \right]$ ratio. Our shadow $\frac{r}{w}$, then, turns out to be just equal to that ratio.

² The price-ratios relate reciprocally to the axes: $\frac{P_A}{P_N} = \frac{P_A^O}{P_N^O}$ in Fig. 2-b. Along, e.g. X-s indifference curve (U_X at Ω'') a rise in $\frac{P_A}{P_N}$, i.e. a steepening of $P_A - P_N$, results in a substitution by X of nuts for apples; ditto for Y.

³ Remember, in choosing the one point on $S_{\Omega} - S_{\Omega}$ that would lie on the envelope in utility space, we chose the point where the indifference curve slopes just equaled the marginal rate of transformation (Part I - 2, Par. 5-6.)

the marginal rate of transformation of nuts into apples, via production, were different from ten-to-one. Producers would not in fact produce the apple-nut combination of Ω if $\frac{P_A}{P_N}$ differed from MRT at Ω ¹.

We have identified the $\frac{W}{r}$ and $\frac{P_A}{P_N}$ implied by the maximum of W . These two constancies provide two equations to solve for the four unknown prices. Unfortunately this is as far as the two-dimensional diagrammatics will take us. None of the diagrams permit easy identification of the relationship between the input prices and the output prices. Yet such a relationship is surely implied. By the theory of the firm we know that the profit maximizing producer facing a constant price for his product -- the horizontal demand curve of the perfectly competitive firm -- will expand output up to where his extra revenue for an additional unit of output: price, just equals the marginal cost of producing that output.¹ And marginal cost, in turn, is sensitive to r and w .

It would be easy to show the implied price-wage or price-rent relationships by introducing marginal productivity notions. Profit maximization requires that the quantity of each input hired be increased up to the point where its marginal physical product -- the extra output due to an increase at the margin in that one input -- times the price of the extra output, just equals the price of the added input. Since these marginal physical productivities are determinate curvature properties of the production functions, this rule provides a third relationship, one between an output price and an input price.

¹ Never mind here the "total" requirement -- that this price exceed unit cost -- if the real life profit seeking producer is to produce at all. More on this in Part V.

Alternatively, given our assumption that production functions show constant returns to scale, we can make use of the Euler "product exhaustion" theorem. Its economic content is that if constant returns to scale prevail the total as-if-market-imputed income of the factors of production just "exhausts" the total value of the product. This means, simply, that $wL + rD = p_A A + p_N N$, and it provides a third relationship between w , r , p_A and p_N for the Ω -values of L , D , A and N .¹

At any rate, the maximal solution implies a third price-equation, hence we can express three of the prices in terms of the fourth. But what of the fourth? This is indeterminate, given the characteristics of the model. In a frictionless world of perfect certainty, where, for example, nobody would think of holding such a thing as money, only relative prices matter. The three equations establish the proportions among them implied by the maximum position, and the absolute values are of no import. If the $p_A : p_N : w : r$ proportions implied by Ω are 20:15:50:75, profit and satisfaction maximizers will make the input-output-consumption decisions required for the maximum-of- W irrespective of whether the absolute levels of these prices happen to be just 20:15:50:75, or twice, or one-half, or fifty times this set of numbers. This is the implication of the fact that for the maximum problem only the various transformation and substitution ratios matter. In all that follows we shall simply posit

¹ The condition also holds for each firm. In a competitive and constant returns to scale world the profit maximum position is one of zero profit: total revenue will just equal total cost. It should be said, however, that use of the Euler theorem to gain an input price - output price relationship involves a measure of sleight of hand. It is only as a consequence of the price-marginal productivity relationships of the preceding paragraph that the theorem assures equality of income with value of product.

that nuts are established as the unit of account, hence that $p_N = 1$. This then makes p_A , w and r fully determinate constants.¹

Summarizing: we have identified diagrammatically two of the three shadow-price relationships implied by the solution to the welfare-maximum problem and have established, in a slightly more roundabout way, the existence of the third. The purpose was to demonstrate the existence, at least in our idealized neo-classical set-up, of a set of constants embedded in the "technocratic" maximum-of-welfare problem, that can be viewed as competitive market prices.² In what sense? In the sense that decentralized decisions in response to these constants, by, or "as if" by, atomistic profit and satisfaction maximizers will result in just that configuration of inputs, outputs and commodity-distribution that the maximum of our W requires.

III

We have said nothing, so far, of how X and Y "pay" for their apples and nuts, or of who "owns" and supplies the labor and the land. As was indicated above, the assumption of constant returns to scale assures that at the maximum welfare position total income will equal total value of output, and that total revenue from the sale of apples (nuts) will just equal total expenditures for inputs by the producers of apples (nuts). Also, the "solution" implies definite "purchase" of

¹ For the possibility of inessential indeterminacies, however, see Part IV - 2

² On the existence of such a set of shadow prices in the kinky and flat-surfaced world of linear programming, see Part V, below.

apples and of nuts both by X and by Y. But nothing ensures that the initial "ownership" of labor-hours and of land is such that w times the labor-hours supplied by X, wL_X , plus r times the land supplied by X, rD_X , -- X's income -- will suffice to cover his purchases as required by Ω : $p_A A_X + p_N N_X$; similarly for Y. There does exist some Pareto-efficient solution of inputs, outputs and distribution that satisfies the "income - outgo" condition for both individuals for any arbitrary pattern of ownership of the "means of production" -- a solution, that is, that will place the system somewhere on the grand utility-possibility envelope frontier (B-B in Fig. 3-b). But only by the sheerest accident will that point on B-B be better in terms of my W -function, or Thomas Jefferson's, or that of a "political consensus", than a multi-dimensional infinity of other points on or off B-B. As emphasized above, only one point on B-B has normative, prescriptive significance: Ω ; and only some special ownership patterns of land and of labor services will place a market system with an "as imputed" distribution of income at that special point.¹

The above is of especial interest in evaluating the optimality characteristics of market institutions in an environment of private property ownership. But the problem is not irrelevant even where all non-human means of production are vested in the community, hence where the proceeds of non-wage income are distributed independently of marginal productivity, marginal-rate-of-substitution considerations.

¹ One way to evade this problem is to posit that whatever the productivity-imputed distribution of income, it is possible to achieve the Ω -implied distribution by costless lump-sum transfers. But nobody has yet invented such a non-distorting tax device.

If labor-services are not absolutely homogeneous -- if some people are brawny and dumb and others skinny and clever, not to speak of "educated" -- income distribution will be sensitive to the initial endowment of these qualities of mind and body and skill relative to the need for them. And again, only a very low probability accident would give a configuration consistent with any particular W-function's Ω .¹

Even our homogeneous-labor world cannot entirely beg this issue. It is not enough to assume that producers are indifferent between an hour of X's as against an hour of Y's labor-services. It is also required that the total supply of labor-hours per accounting period be so divided between X and Y as to split total wage payments in a particular way, depending on land ownership and on the income distribution called for by Ω . This may require that X supply e.g. 75% of total L : each man working $\frac{1}{2} L$ hours may well not do.²

But all this is diversion. For our non-institutional purposes it is sufficient to determine the particular L_X , D_X , L_Y and D_Y that are consistent with Ω , given market-imputed, or "as if" market-imputed, distribution. Unfortunately the diagrams used in Part I again fail, but the algebra is simple. It is required that:

$$wL_X + rD_X = p_A A_X + p_N N_X, \text{ and}$$

$$wL_Y + rD_Y = p_A A_Y + p_N N_Y,$$

¹ If slavery were the rule and I could sell the capitalized value of my expected lifetime services, the distinction between ownership of labor and that of land would completely disappear.

² All this is based on the "Austrian" assumption that labor is supplied inelastically; further, that such inelasticity is due not to external compulsion but to a peculiarity of the preference-fields of X and Y in relation to work-leisure choices. More than this, the W-function must not be sensitive to alternative L_X - L_Y mixes except as these influence income distribution.

for the already solved-for maximal Ω -values of A_X , N_X , A_Y , N_Y , p_A , p_N , w and r . Together with $L_X + L_Y = L$ and $D_X + D_Y = D$, we appear to have four equations to solve for the four unknowns: L_X , L_Y , D_X and D_Y . It turns out, however, that one of these is not independent. The sum of the first two: that total incomes equal total value of product, is implied by the Euler theorem taken jointly with the marginal productivity conditions that give the solution for the eight variables, A_X , N_X , A_Y , ... , that are here taken as known. Hence we only have three independent equations. This is as it should be. It means only that with our nice curvature assumptions we can, within limits, fix one of the four endowments more or less arbitrarily and still so allocate the rest as to satisfy the household budget equations.

So much for the income-distribution aspects of the problem. These have relevance primarily for market-imputed income distribution; but such relevance does not depend on "private" ownership of non-labor means of production. Note, incidentally, that only with the arbitrary "Austrian" assumption of fixed supplies of total inputs can one first solve "simultaneously" for inputs, outputs and commodity-distribution, and only subsequently superimpose on this solution the ownership and money-income distribution problem. If L_X , D_X , L_Y , D_Y , hence L and D were assumed sensitive to w , r , the p -s and household income levels, the dimensions of the production-box of Fig. 1, hence the position of the production possibility curve of Figs. 2, etc. would interdepend with the final solution values of L_X , D_X , L_Y and D_Y . We would, then, have to solve the full problem as a set of simultaneous equations, from the raw data: production functions, tastes (this time with an axis for leisure), and the W -function.

IV

We have demonstrated the solution of the maximum problem of modern welfare-economics in context of the simplest statical and stationary classical model. Many generalizations and elaborations suggest themselves, even if one remains strictly neo-classical and restricts oneself to a steady-state situation where none of the data change and no questions about "how the system gets there" are permitted to intrude. To comment on just a few:

1) The problem could well be solved for many households, many goods, and many factors: it has received complete and rigorous treatment in the literature. Of course the diagrammatics would not do; elementary calculus becomes essential. But the qualitative characteristics of the solution of the m by n by q case are precisely those of the 2 by 2 by 2. The same marginal rate of transformation and substitution conditions characterize the solution, only now in many directions. Nothing new or surprising happens.¹

2) The solution did skirt one set of difficulties that were not explicitly ruled out by assumption. We tacitly assumed that the two sets of isoquants would provide a smooth locus of "internal" tangencies, F-F, in the production box of Fig. 1; similarly, that we would get such an "internal" S-S in the trading boxes of Figs. 2a, b. Nothing in our assumptions guarantees that this should be so. What if the locus of maximum A-s for given feasible N-s, should occur not at points of strict tangency inside the box, but at what the mathematician would call corner-

¹ Rigorous general treatment of the $m \times n \times p$ situation does highlight a number of analytical fine points that are of interest to the pure theorist, e.g. the difficulties encountered if the number of factors exceeds the number of goods. But the qualitative economics is the same. For a full treatment from a non-normative point of view, see P.A. Samuelson, "Prices of Factors and Goods in General Equilibrium", Review of Economic Studies, XXI (1), No. 54.

tangencies along the edges of the box? Fig. 4 illustrates this possibility. The maximum feasible output of A, for $N = 6000$, occurs

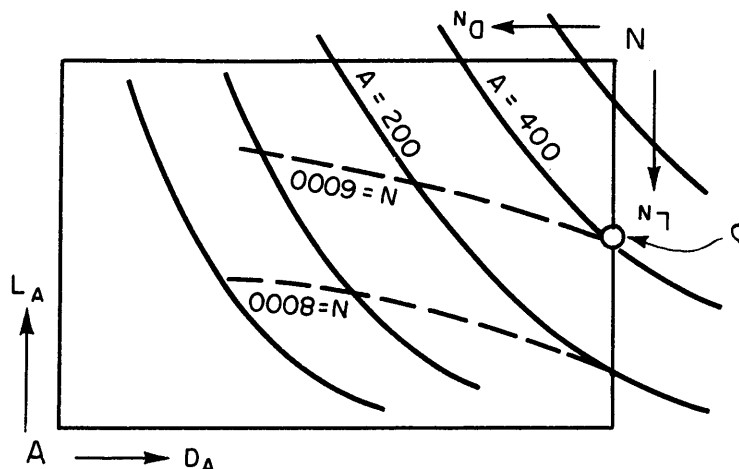


Fig. 4

at \bar{U} , where $A = 400$; but at \bar{U} the two isoquants are not strictly tangent (they touch but have different slopes). The economic meaning of this is simple. With endowments as indicated by the dimensions of the production box in Fig. 4, and with technology as denoted by the isoquants, it is not possible to reallocate inputs until the MRS of labor for land is the same in apple as in nut production. This is because apple technology (as depicted) is so land-using relative to nut production that the $\left[\frac{\text{marginal productivity of land}}{\text{marginal productivity of labor}} \right]$ ratio in apple production exceeds that in nut production even when, as at \bar{U} , all land is devoted to apples.

Space precludes further analysis of such corner-tangency phenomena. They reflect the possibility that the maximum-welfare solution may require that not every input be used in producing every output (e.g. no land in nut-production or no brain surgeons in coal-mining), and

may even render one of the inputs a "free good", so that its total use will not add up to the total available supply. Let it suffice to assert that by formulating the maximum conditions, not in terms of equalities of various slopes, but rather in terms of inequalities; by explicit statement of the proper second order "rate-of-change-of-slope" conditions; and by allowing inequalities in the factor-balance conditions, (e.g. $L_A + L_N \leq L$), such phenomena of bumping into the axes can be handled; further, that only inessential indeterminacies occur in the implied shadow-price configuration.¹

3) It is useful, and in a mathematical treatment not difficult, to drop the "Austrian" assumption of inelastically supplied inputs, and introduce leisure-work choices. The analytical effect is to sensitize the production possibility curve to the psychic sensibilities -- the preference functions -- of individuals. Note that the empirical sense of doing so is not confined to an institutional or ethical context of non-imposed choice.

¹ All this can perhaps be made clearer by two examples. In our Fig. 4, the essential requirement if A is to be at a maximum for $N = 6000$ is that the intersection at the boundary be as in Fig. 5a (~~next page~~) rather than as in Fig. 5b. The latter gives a minimum of A for $N = 6000$. The distinction between 5a and 5b is between the relative rates of change of the two MRS-s. The price indeterminacy implied by the maximum (5a) i.e. the fact that \bar{r} is consistent with an $\frac{r}{w}$ that lies anywhere between the two isoquants, turns out to be inessential. A second example concerns the theory of the firm. It has been argued that if the marginal cost curve has vertical gaps and the price-line hits one of these gaps, then the $M.C. = p$ condition is indeterminate, hence that the theory is no good. As has been pointed out in the advanced literature (e.g. by R.L. Bishop, in "Cost Discontinuities ... " *A.E.R.* XXXVIII, 1948, p. 60) this is incorrect: What is important is that at smaller than equilibrium output M.C. be less than price and at higher outputs M.C. exceed price. It is true, but quite harmless to the theory, that such a situation does leave a range of indeterminacy in the price that will elicit that level of output. Such phenomena do change the mathematics of computation. Inequalities cannot in general be used to eliminate unknowns by substitution.

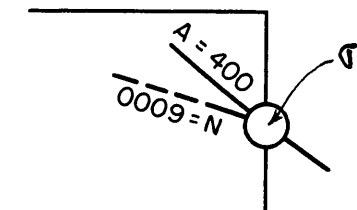


Fig. 5-a

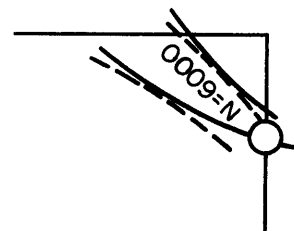


Fig. 5-b

4) We assumed away joint product -- lambs produce wool and mutton -- situations. This is convenient for manipulation but hardly essential; the results can be generalized to cover most kinds of jointness. It turns out, in fact, that in dynamical models with capital-stocks, one means for taking account of the durability of such stocks is to allow for joint-products. A process requiring a hydraulic press "produces" both stamped metal parts and a "one year older" hydraulic press.

5) In our system the distinction between inputs (L, D) and outputs (A, N) could be taken for granted. But the distinction is clear only in a world of completely vertically-integrated producers, all hiring "primary" non-produced inputs and producing "final" consumable goods and services. In a Leontief-like system that allows for inter-producer transactions and intermediate products, many outputs: electricity, steel, corn, beef, trucks, etc., are simultaneously inputs. It is of interest, and also feasible, to generalize the analysis to take account of e.g. coal being used not only to heat houses, but to produce steel required in the production of mining machines designed for the production of coal. Moreover, none of the essential qualitative characteristics of our maximum problem are violated by such generalization.¹

6) What if instead of assuming that production functions show constant returns to scale, we permit diminishing returns to proportional expansion of inputs? These could be due either to inherent non-linearities

¹ Analytically, this is done by designating all produced goods as $X_1, X_2, X_3 \dots$. The gross production of e.g. X_1 has two kinds of uses: It is partly used up as an input in the production of $X_2, X_3 \dots$ and perhaps of X_1 (the automobile industry is a major user of automobiles). What remains is available for household consumption. The production functions have X-s on the right as well as the left hand side.

in the physics and topography of the universe, or to the existence of some unaccounted for but significant input in limited, finite-elastic, supply.¹

Diminishing returns to scale, as distinct from increasing returns, do not give rise to serious trouble, either for the analytical solubility of the system, or for the market-significance of the intrinsic price-wage-rent constants. They do introduce some ambiguities, however. For one thing, the "value" of output will exceed the total of market-imputed income. This makes intuitive sense in terms of the "unaccounted-scarce-factor" explanation of decreasing returns; the residual unimputed value of output reflects the income "due" the "hidden" factor. If that factor were treated explicitly and given an axis in the production function diagram, returns would no longer diminish -- since, on this view, the relative inexpansibility of that input gave rise to decreasing returns to scale to begin with -- and the difficulty would vanish.²

In a market context, this suggests the explicit introduction of firms as distinct from industries. In our constant returns to scale world the number of apple (nut) producing firms could be assumed indeterminate.

¹ If "output" varies as the surface area of some solid body and "input" as its cubic-volume, a doubling of input will less than double output -- this is an example of the first kind. A typical example of the second is the instance where the production function for fishing does not include an axis for the "amount" of lake, hence where beyond a certain point doubling of man-hours, boats, etc. less than doubles the output. There is a slightly futile literature on whether the first kind could or could not exist without some element of the second. If every input is really doubled, so say the proponents of one view, output must double. The very vehemence of the assertion suggests the truth, to wit, that it is conceptually impossible to disprove it by reference to empirical evidence. Luckily, the distinction is not only arbitrary -- it depends on what one puts on the axes of the production-function diagram and what is built into the curvature of the production surface; it is also quite unimportant. One can think of the phenomenon as one will -- nothing will change.

² The fact that the "hidden scarce factor" view is heuristically useful does not, however, strengthen its pretension to status as a hypothesis about reality.

Every firm could be assumed able to produce any output up to A_{Ω} at constant horizontal unit cost. In fact, if we had a convenient way of handling incipient monopoly behavior, such as by positing frictionless entry of new firms, we could simply think of one giant firm as producing all the required apples (nuts). Such a firm would be compelled, nevertheless, to behave as though it were an "atomistic" competitor, i.e. prevented from exploiting the tilt in the demand curve, by incipient competitors ready instantaneously to jump into the fray at the slightest sign of profit.

With decreasing returns to scale, however, it is natural, at least in a context of market-institutions, to think of these as associated with the qualitatively and quantitatively scarce entrepreneurial entity that defines the firm but is not explicitly treated as an input. Then, as apple production expands, relatively less-efficient entrepreneurs are pulled into production -- the total cost curve of the "last" producer, and the associated shadow price of apples, become progressively higher -- and the intra-marginal firms make "profits" due directly to the scarcity value of the entrepreneurial qualities of their "entrepreneurs". The number of firms, their inputs and outputs, are determinate. The last firm just breaks even at the solution-value of the shadow-price.¹

At any rate, no serious damage is done to the statical system by decreasing returns to scale. From the point of view of actually computing a maximum problem the loss of linearity is painful, but the trouble

¹ More precisely, the "next" firm in line could not break even. This takes care of discontinuity.

is in the mathematics.¹

7) There is one kind of complication that does vitiate the results. We have assumed throughout that there exists no direct interaction among producers, among households, and between producers and households -- that there are no (non-pecuniary) external economies or diseconomies of production and consumption. The assumption is reflected in three characteristics of the production functions and the preference functions.

(a) The output of apples was assumed uniquely determined by the quantities of land and labor applied to apple production -- "A" was assumed insensitive to the inputs and outputs of the nut-industry; similarly for nuts. This voids the possibility that the apple production function might shift as consequence of movements along the nut production function, i.e. that for given D_A and L_A , A may vary with N , L_N and D_N . The stock example of such a "technological external economy" (or diseconomy) is the bee keeper whose honey output will increase, other things equal, if the neighboring apple producer expands his output (hence his apple blossom "supply").² The very pastoral quality of the example suggests that in a statical context such direct interaction among producers -- interaction that is not reflected by prices -- is probably rare. To the extent that it does exist,

¹ It should perhaps be repeated, however, that there remains considerable ambiguity about how the imbalance between income and outlay in decreasing returns to scale situations is best treated in a general equilibrium set-up.

² The other type of externality treated in the neo-classical literature, the type Prof. Viner labeled "pecuniary", does not in itself affect the results. It consists in sensitivity of input prices to industry output, though not to the output of single firms. External pecuniary economies (as distinct from diseconomies) do, however, signal the existence of either technological external economies of the sort discussed here, or of internal economies among supplier firms. These last reflect increasing returns to scale along production functions -- a most troublesome state discussed at length in Part V.

it reflects some "hidden" inputs or outputs (e.g. apple blossoms), the benefits or costs of which are not (easily) appropriated by market institutions.

It should be emphasized that the assertion that such phenomena are empirically unimportant is defensible only if we rule out non-reversible dynamical phenomena. Once we introduce changes in knowledge, for example, or investment in changing the quality of the labor force via training, "external" effects become very important indeed.¹ But on our stratospheric level of abstraction such considerations are out of order.

(b) The "happiness" of X, as measured by U_X , was assumed uniquely determined by his own consumption of apples and nuts. He was permitted no sensitivity to his neighbor's (Y's) consumption, and vice versa. This rules out not only Veblenesque "keeping up with ..." effects, but such phenomena as Y tossing in sleepless fury due to X's "consumption" of midnight television-shows; or X's temperance sensibilities being outraged by Y's quiet and solitary consumption of scotch. Nobody with experience of a "neighborhood" will argue that such things are illusory, but it is not very fruitful to take account of them in a formal maximizing set-up.

¹ The full "benefits" of most changes in "knowledge", of most "ideas", are not easily captured by the originator, even with strong patent and copyright protection. If, then, the energy and resources devoted to "creating new knowledge" are sensitive to private cost-benefit calculation, some potential for social gain may well be lost because such calculation will not correctly account for cost and benefit to society at large. All this is complicated by the peculiarity of "knowledge" as a scarce resource: unlike most other scarcities, just because there is more for you there is not necessarily less for me. As for training of labor: the social benefit accrues over the life-time services of the trainee; the private benefit to the producer accrues until the man quits to go to work for a competitor.

(c) X and Y were assumed insensitive, also, to the input-output configuration of producers, except as these affected consumption choices. Insensitivity to the allocation of their own working time is subsumed in the "Austrian" assumption, but more is required. Y's wife must not be driven frantic by factory soot, nor X irritated by an "efficiently" located factory spoiling his view.

There is still a fourth kind of externality: X's satisfaction may be influenced not only by his own job, but by Y's as well. Many values associated with job-satisfaction -- status, power, and the like -- are sensitive to one's relative position, not only as consumer, but as supplier of one's services in production. The "Austrian" assumption whereby U_X and U_Y are functions only of consumption possibilities, voids this type of interaction also.

Could direct interaction phenomena be introduced into a formal maximizing set-up, and, if so, at what cost? As regards the analytical solubility of some maximum-of-W problem, there is no necessary reason why not. The mathematics of proving the existence or non-existence of a "solution", or of a unique and stable "solution", or the task of devising a computational routine that will track down such a solution should one exist, may become unmanageable. But the problem need not be rendered meaningless by such phenomena.

Unfortunately that is saying very little indeed, except on the level of metaphysics. Those qualities of the system that are of particular interest to the economist -- (i) that the solution implies a series of "efficiency conditions", the Pareto marginal-rate of substitution

conditions, which are necessary for the maximum of a wide variety of W-functions, and (ii) that there exists a correspondence between the optimal values of the variables and those generated by a system of (perfect) market institutions cum redistribution -- those qualities are apt to vanish with "direct interaction". Such interaction destroys the "duality" of the system: the constants embedded in the maximum problem, if any, lose meaning as prices, wages, rents. They will not correctly account for all the "costs" and "benefits" to which the welfare function in hand is sensitive.¹

In general, then, most formal models rule out such phenomena. There is no doubt that by so doing they abstract from some important aspects of reality. But theorizing consists in just such abstraction: no theory attempts to exhaust all of reality. The question of what kinds of very real complications to introduce into a formal maximizing set-up has answers only in terms of the strategy of theorizing or in terms of the requirements of particular and concrete problems. For many purposes it is useful and interesting to explore the implications of maximizing in a "world" where no such direct interactions exist.

V

None of the above qualifications and generalizations violate the fundamentally neo-classical character of the model. What happens if we relinquish some of the nice curvature properties of the functions?

¹ It should not be concluded, however, that the different types of direct interaction are all equally damaging. All will spoil market performance, almost by definition: but some, at least, permit of formal maximizing treatment such as will yield efficiency conditions that are analogous to those of Part I -- conditions that properly account for full social costs and benefits.

(1) We required that the production functions and the indifference curves have well-defined and continuous curvatures -- no sharp corners or kinks such as cause indeterminacy in marginal rates of substitution. Such smooth curvatures permit the use of the calculus, hence are mathematically convenient for larger than 2 by 2 by 2 models. They are, however, not essential to the economic content of the results. The analysis has been translated -- and in part independently re-invented -- for a world of flat faced, sharp cornered, production functions: Linear Programming, more formally known as Activity Analysis, is the resulting body of theory.¹ All the efficiency conditions have their counter-parts in such a system, and the existence of implicit "prices" embedded in the maximum problem is, if anything, even more striking.²

(2) The above easing of the neo-classical requirement that functions be smooth is not only painless; in the development of analytical economics it has resulted in exciting new insights. Unfortunately, however, the next step is very painful indeed. In our original assumptions we required that returns to scale for proportional expansion of inputs be constant (or at least non-increasing) and that isoquants and indifference curves be "convex to the origin". These requirements guarantee a condition that the mathematicians call convexity. The violation of this condition, as by allowing increasing returns to scale in production -- due, if you wish, to the inherent physics and topography of the universe or to lumpiness and indivisibilities -- makes for serious difficulties.

The essence of convexity, a concept that plays a crucial role in

¹ Isoquants in such a set-up consist of linearly additive combinations of processes, each process being defined as requiring absolutely fixed input and output proportions. This gives isoquants that look like that in Fig. 6-c below.

² A little diagrammatic experimentation will show that the geometric techniques of Part I remain fully adequate.

mathematical economics, is rather simple. Take a single isoquant such as M-M in Fig. 6-a. It denotes the minimum inputs of L and D for the production of 100 apples, hence it is just the boundary of all technologically feasible input combinations that can produce 100 apples. Only points on M-M are both feasible and technologically efficient, but any point within the shaded region is feasible; nobody can prevent me from wasting L or D. On the other hand, no point on the origin side of M-M is feasible for an output of 100 apples: given the laws of physics, etc. it is impossible to do better. Mathematical convexity obtains if a straight line connecting any two feasible points does not anywhere pass outside the set of feasible points. A little experimentation will show that such is the case in Fig. 6-a. In Fig. 6-b, however, where the isoquants are of "queer" curvature -- MRS of L for D increases -- the line connecting e.g. the feasible points γ and ϕ does pass outside the "feasible" shaded area. Note, incidentally, that an isoquant of the linear programming variety, as in Fig. 6-c, is "convex"--

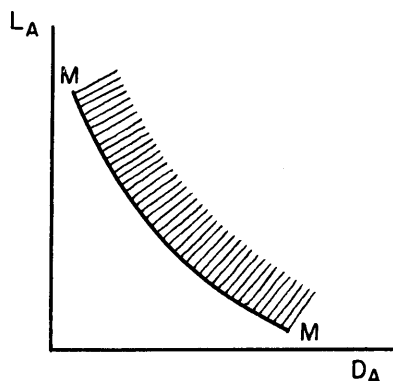


Fig. 6-a

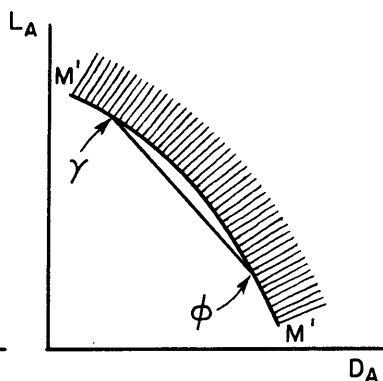


Fig. 6-b

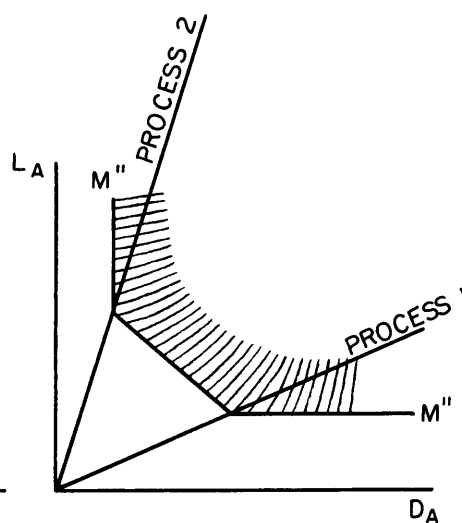


Fig. 6-c

this is why the generalization of (1) above was painless.¹

What kind of trouble does non-convexity create? In the case of concave-to-the-origin isoquants, i.e. non-convex isoquants, the difficulty is easy to see. Look back at Fig. 1 and imagine that the old nut-isoquants are really those of apple producers, hence oriented to the south-west, and vice versa for nuts. Examination of the diagram will show that the locus of tangencies, F-F, is now a locus of minimum combinations of A and N. Hence the rule that MRS-s be equalized will result in input-combinations that give a minimum of N for specified A.²

(3) This is not the occasion for extensive analysis of convexity problems. It might be useful, however, to examine one very important variety of non-convexity: increasing returns to scale in production. Geometrically, increasing returns to scale are denoted by isoquants that are closer and closer together for outward movement along any ray from the origin: to double output, you less than double the inputs. Note that the isoquants still bound convex sets in the L-D plane (they are still as in Fig. 6-a). But in the third or output dimension of a two input - one output production surface, slices by vertical planes through the origin perpendicular to L-D will cut the production surface in such a way as to give a boundary such as V-V

¹ It is important not to confuse mathematical convexity with curvature that appears "convex to the origin". Mathematical convexity is a property of sets of points, and the set of feasible output points bounded by a production possibility curve, for instance, is convex if and only if the production possibility curve itself is "concave to the origin" (or a straight line). Test this by the rule which defines convexity.

² A minimum, that is, subject to the requirement that no input be "wasted" from an engineering point of view, i.e. that each single producer be on the production function as given by the engineer.

in Fig. 7. It is evident that V-V bounds a non-convex set of feasible points -- try connecting any two points on V-V -- so the full three-dimensional set of feasible input-output points is not convex.

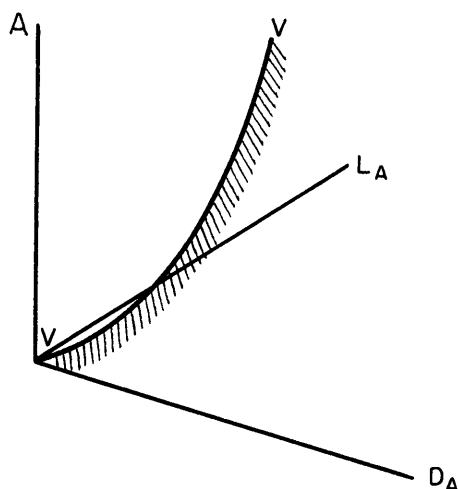


Fig. 7

The effect of such non-convexity in input-output space can be classified with respect to its possible implications for a) the slopes of producers' average cost (A.C.) curves; b) for the slopes of marginal cost (M.C.) curves; c) for the curvature of the production possibility curve.

(a) Increasing returns to scale and A.C. curves. It is a necessary consequence of increasing returns to scale that at the maximal configuration of inputs, outputs and input prices, producers' A.C. curves decline with increasing output. By the definition of increasing returns to scale at a given point \mathcal{T} of a production function, successive isoquants in the neighborhood of \mathcal{T} lie closer and closer together for movement "north-east"

along the ray from the origin through \mathcal{T} (Z in Fig. 8). As Fig. 8 is drawn, the ray Z happens also to correspond to an expansion path for the particular $\frac{r}{w}$ ratio denoted by the family of iso-cost lines $R'-W'$: each $R'-W'$ is tangent to an isoquant along Z. Given $\frac{r}{w} = |\text{tangent } \theta|$, a profit-maximizing apple producer will calculate his minimum total cost for various levels of output from input-output points along Z.

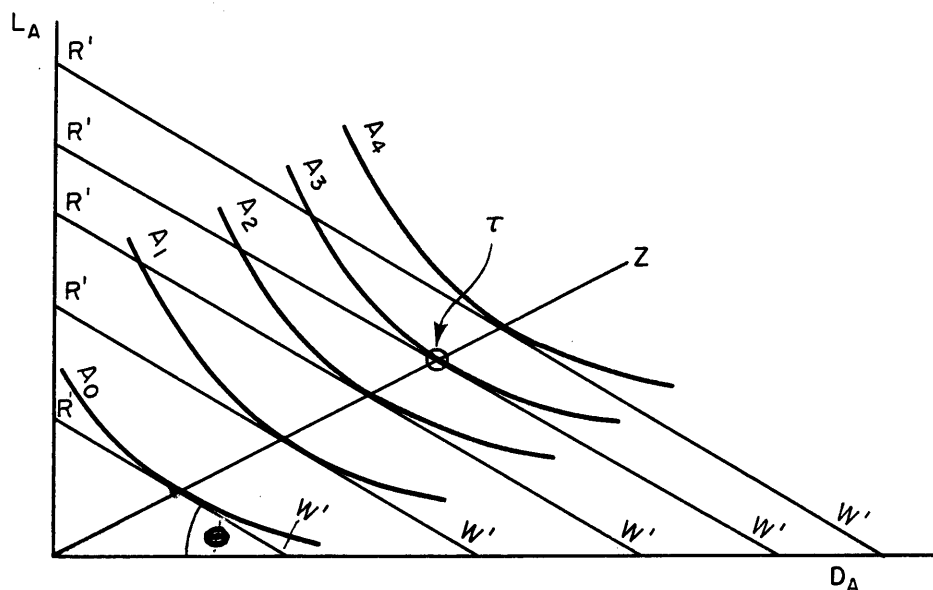


Fig. 8

But along Z the equal cost $R'-W'$ tangents in the neighborhood of \mathcal{T} lie closer and closer together for increasing output, as do the isoquants. This implies that the increase in total cost for equal successive increments in output declines. Ergo, the A.C. curve at \mathcal{T} for $\frac{r}{w} = |\text{tangent } \theta|$ must be falling.

Suppose the expansion path for $\frac{r}{w} = |\text{tangent } \theta|$ happened not to correspond to the ray Z, but only to cross it at \mathcal{T} . The intersection of A_4 with Z would not then mark the minimum-cost input-mix

for an output of A_4 , hence the increase in minimized total cost between A_3 and A_4 would be even less than in Fig. 8: the negative effect on A.C. would be reinforced. The point is, simply, that if for movement along a ray from the origin cost per unit of output declines, A.C. will decline even more should production at minimized total cost call for changes in the input-mix, i.e. departure from the ray Z.

What, then, if the maximum-of-W input-output combination required of this particular producer is denoted by the point \mathcal{T} ? It has just been shown that A.C. at \mathcal{T} is falling. A falling A.C. implies a marginal cost curve (M.C.) that lies below the average. But if \mathcal{T} is the \mathcal{N}'' -point, the shadow- p_A will just equal M.C. of \mathcal{T} . It follows that the maximum-of-W configuration requires $p_A < A.C.$: perpetual losses. Losses, however, are incompatible with real life (perfect) markets; hence where increasing returns to scale prevail correspondence between market-directed and W-maximizing allocation fails. In an institutional context where producers go out of business if profits are negative, markets will not do.¹

Increasing returns to scale has also a "macro" consequence that is associated with $p < A.C.$ For constant returns to scale, we cited the Euler-theorem as assuring that total factor incomes will just equal total value of output. In increasing returns to scale situations, total imputed factor incomes will exceed the total value of output:

$$rD + wL > p_A A + p_N N. \quad 2$$

¹ Needless to say, comments on market effectiveness, throughout this paper, bear only on the analogue-computer aspects of price-market systems. This is a little like talking about sexless men, yet it is surely of interest to examine such systems viewed as mechanisms pure and simple.

² The calculus trained reader can test this for, say, a Cobb-Douglas type function: $A = L_A^\alpha D_A^\beta$, with $(\alpha + \beta) > 1$ to give increasing returns to scale.

(b) Increasing returns to scale and M.C. curves. Where non-convexity of the increasing returns to scale variety results in falling A.C. curves, real life (perfect) markets will fail. What of a Lange-Lerner socialist bureaucracy, where each civil servant plant manager is instructed to maximize his algebraic profits in terms of centrally quoted "shadow" prices regardless of losses? Will such a system find itself at the maximum-of-W configuration?

It may or may not. If A.C. is to fall, M.C. must lie below A.C., but at the requisite Ω -output M.C.-s may nevertheless be rising, as for example at ξ in Fig. 9. If so, a Lange-Lerner bureaucracy making input and output decisions as atomistic "profit-maximizing" competitors but ignoring losses will make the "right" decisions, i.e. will "place" the system at the maximum-of-W. Each manager equating his marginal cost to the centrally quoted shadow price given out by the maximum-of-W solution, will produce precisely the output required by the Ω -configuration. By the assumption of falling A.C.-s due to increasing returns to scale

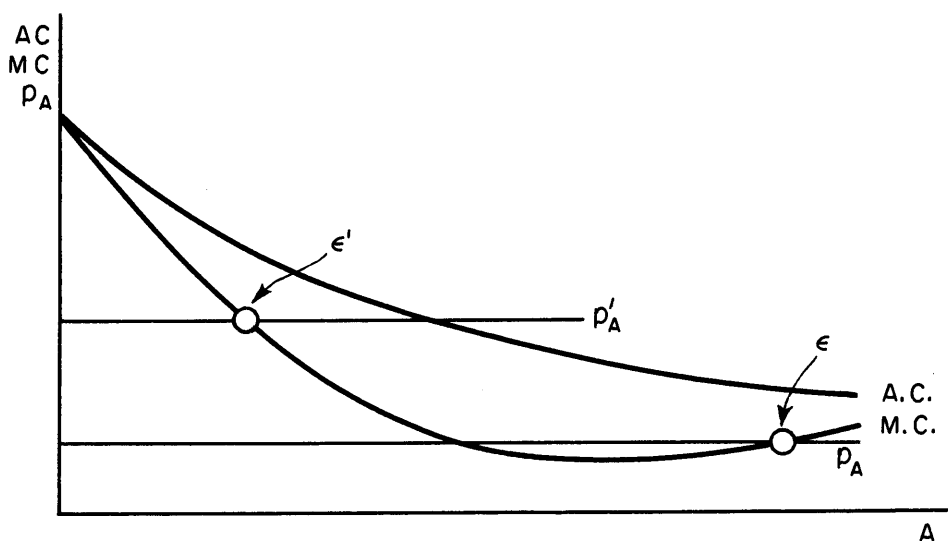


Fig. 9

either one or both industries will show losses, but these are irrelevant to optimal allocation.¹

What if for a maximum-of-W producers are required to produce at points such as ξ^1 , where $p = M.C.$ but $M.C.$ is declining? ² The fact that this implies $(A.C.) > (M.C.) = p$, hence losses, has been dealt with above. But more is implied. By the assumption of a falling $M.C.$ -curve, the horizontal price line at ξ^1 cuts the $M.C.$ curve from below, hence profit at ξ^1 is not only negative: it is at a minimum. A "real life" profit maximizer would certainly not remain there: he would be losing money by the minute. But neither would a Lange-Lerner bureaucrat under instruction to maximize algebraic profits. He would try to increase his output: "extra" revenue (p_A) would exceed his $M.C.$ by more and more for every additional apple produced. In this case, then, not only would real life markets break down; so would simple-minded decentralized maximizing of profits by socialist civil servants.³

¹ There is an ambiguity of language in the above formulation. If at the maximum-of-W configuration losses prevail, the maximum profit position "in the large" will not be at $p = M.C.$ but at zero output. Strictly speaking, a Lange-Lerner bureaucracy must be instructed to equate marginal cost to price or profit-maximize "in the small" without regard to the absolute value of profit. "Make any continuous sequence of small moves that increase algebraic profits, but do not jump to the origin". It is precisely the ruling-out of the zero output position, unless called for by $M.C. > p$ everywhere, that distinguishes Lange-Lerner systems from "real-life" perfect markets, both viewed as "analogue computers".

² This would necessarily be the case, for instance, with Cobb-Douglas type increasing returns to scale functions. Such functions imply ever falling $M.C.$ curves, for whatever $\frac{r}{w}$ ratio.

³ It is of interest to note that a falling $M.C.$ curve is simply a reflection of non-convexity in the total cost curve.

Paradoxically enough, the correct rule for all the industries whose M.C. is falling at the Ω -point is: "minimize your algebraic profits". But no such rule can save the decentralized character of the Lange-Lerner scheme. In a "convex" world the simple injunction to maximize profits in response to centrally quoted prices, together with $\left\{ \begin{array}{l} \text{raising} \\ \text{lowering} \end{array} \right\}$ of prices by the responsible "Ministries" according to whether supply $\left\{ \begin{array}{l} \text{falls short of} \\ \text{exceeds} \end{array} \right\}$ demand, is all that is needed.¹ Nobody has to know ex ante e.g., the prices associated with the Ω -point. In fact the scheme was devised in part as a counter to the view that efficient allocation in a collectivized economy is impossible due simply to the sheer administrative burden of calculation. With increasing returns to scale, however, the central authority must evidently know where M.C.-s will be falling, where rising: it must know, before issuing any instructions, all about the solution.

(c) Increasing returns to scale and the production possibility curve.

What is left of "duality"? Real life markets and unsophisticated Lange-Lerner systems have both failed. Yet it is entirely possible, even in situations where the Ω -constellation implies $A.C. > M.C.$ with declining M.C., that the maximizing procedure of Part I remains inviolate, and that the constants embedded in the maximum problem retain their price-like significance. To see this we must examine the effect of increasing returns to scale on the production possibility curve. There are two possible cases:

(i) It is possible for both the apple and the nut production functions to exhibit increasing returns to scale yet for the implied production possibility curve to be "concave-to-the-origin", i.e. mathematically convex (as in Fig. 2). While a proportional expansion of L_A and D_A by a factor of two would more than double apple output,

¹ Not quite all. Even in a statical context, the lump sum income transfers called for by Ω require central calculation. And if adjustment paths are explicitly considered, complex questions about the stability of equilibrium arise. (E.g. Will excess demand always be corrected by raising price?).

an increase in A at the expense of N will, in general, not take place by means of such proportional expansion of inputs. Examination of F-F in Fig. 1 makes this clear for the constant-returns-to-scale case. As we move from any initial point on F-F toward more A and less N, the $\frac{L_A}{D_A}$ and $\frac{L_N}{D_N}$ proportions change.¹

The point is that if, as in Fig. 1, land is important relative to labor in producing apples, and vice versa for nuts, expansion of apples will result in apple producers having to use more and more of the relatively nut-prone input: labor, in proportion to land. Input proportions in apple production become less "favorable". The opposite is true of the input proportions used in nuts as nut production declines. This phenomenon explains why with constant returns to scale in both functions the production possibility curve shows concave-to-the-origin curvature. Only if F-F in Fig. 1 coincides with the diagonal: i.e. if the intrinsic "usefulness" of L and D is the same in apple production as in nut production, will F'-F' for constant returns to scale be a straight line.

The above argument by proportions remains valid if we now introduce a little increasing returns to scale in both functions by "telescoping" each isoquant successively further towards the origin. In fact, as long as the F-F curve has shape and curvature as in Fig. 1, the production possibility curve, F'-F' in Fig. 2, will retain its convexity.

In this "mild" case of increasing returns to scale, with a still "convex" production possibility curve, the previous maximizing rules give the correct result for a maximum-of-W. Further, the constants embedded in the maximum problem retain their meaning.

¹ Only if F-F should coincide with the diagonal of the box, will proportions not change. Then increasing returns to scale would necessarily imply an inward bending production possibility curve.

This is true in two senses. One, they still reflect marginal rates of substitution and transformation. Any package of L, D, A and N worth one dollar will, at the margin, be just convertible by production and exchange into any other package worth a dollar, no more, no less: a dollar is a dollar is a dollar...¹

Two, the total value of maximum-welfare "national" output:

$p_A A + p_N N$, valued at these shadow-price constants, will itself be at a maximum. A glance at Fig. 2-b makes this clear: at the price-ratio denoted by the line $P'_A - P'_N$, Ω' is the point of highest output-value. As we shall see, this correspondence between the maximum welfare and "maximum national product" solutions is an accident of convexity.

(ii) It is, of course, entirely possible that both production functions exhibit sufficiently increasing returns to scale to give, for specified totals of L and D, a production possibility curve such as $F''-F''$ in Fig. 10.² This exhibits non-convexity in output

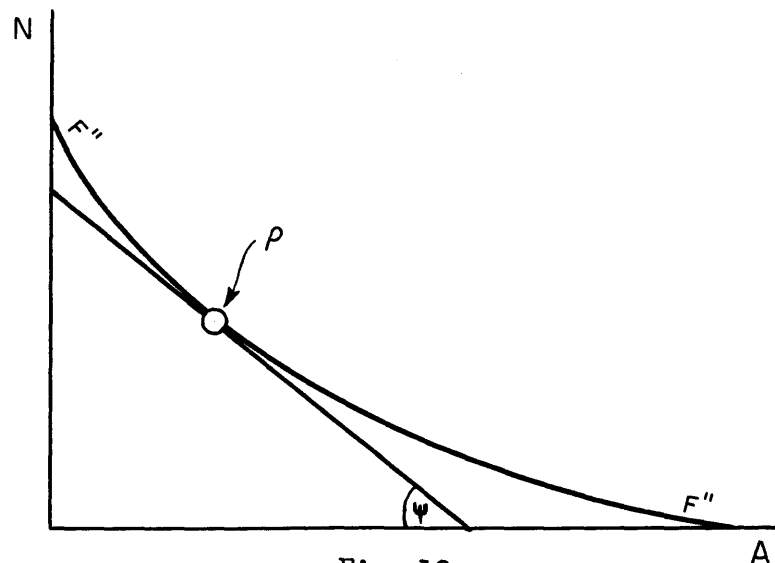


Fig. 10

¹ For the infinitesimal movements of the calculus.

² Try two functions which are not too dissimilar in "factor intensity".

space. (Try connecting any two feasible points). What now happens to the results?

If the curvature of $F''-F''$ is not "too sharp", the constants given out by the maximum-of- W problem retain their "dollar is a dollar" meaning. They still reflect marginal rates of substitution in all directions. But maximum W is no longer associated with maximum shadow-value of output. A glance at Fig. 10 confirms our geometric intuition that in situations of non-convex production-possibilities the bliss point coincides with a minimized value-of-output. At the prices implied, as denoted by $|\tan \Psi|$, the assumed Ω -point ρ is a point of minimum $p_A A + p_N N$.¹

But with non-convexity in output space, matters could get much more complicated. If the production possibility curve is sharply concave outward, relative to the indifference curves, it may be that the "minimize profits" rule would badly mislead, even if both industries show declining M.C.-s. Take a one person situation such as in Fig. 11. The production possibility curve $F'''-F'''$ is more inward-bending than the indifference curves (U), and the point of tangency Δ is a point of minimum satisfaction. Here, unlike above, you should rush away from Δ . The maximum welfare position is at Δ' -- a "corner tangency" is involved. The point is that in non-convex situations relative curvatures

¹ For $\frac{p_A}{p_N} = |\text{tangent } \Psi|$, $(p_A A + p_N N)$ is at its maximum at the intersection of $F''-F''$ with the A-axis. Recall, incidentally, that in situations of falling M.C. producers were required to minimize profits.

curvatures are crucial: tangency points may as well be minima as maxima. (1) (2)

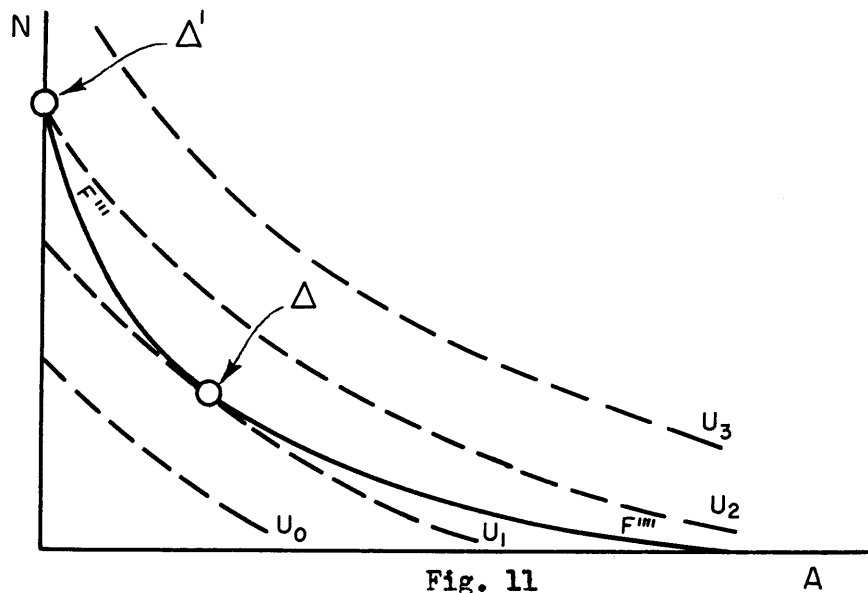


Fig. 11

A

¹ Recall that in our discussion of Part IV corner-tangencies were important in situations where no feasible internal tangencies existed. Here there exist perfectly good and feasible internal tangencies -- but they are loci of minima rather than maxima. The second order conditions, expressed as inequalities, constitute the crucial test of optimal allocation.

² It is tempting, but a mistake, to think that there is a unique correspondence between the curvature of the production possibility curve, and the relative slopes of the nut and apple M.C. curves. It is true that the $\left[\frac{M.C.A}{M.C.N} \right]$ ratio associated with a point such as Ω' in Fig. 2-b must be greater than $\left[\frac{M.C.A}{M.C.N} \right]$ at any point of more A and less N on $F'-F'$ (e.g. δ): the absolute slope of $F'-F'$ has been shown to equal

$\frac{PA}{PN} = \left[\frac{M.C.A}{M.C.N} \right]$, and at Ω' the slope is less steep than at δ . It is also true that along a non-convex production possibility curve, such

as that of Fig. 10, $\left\{ \begin{array}{l} \text{plus A} \\ \text{minus N} \end{array} \right\}$ is associated with a decline in $\left[\frac{M.C.A}{M.C.N} \right]$.

But it does not follow, e.g. in the first case of Fig. 2-b, that at Ω' $M.C.A$ must be rising for plus-A sufficiently to offset a possibly falling $M.C.N$ (Remember in moving from Ω' to δ we move to the right on $M.C.A$ but to the left on $M.C.N$). For any departure from Ω' will, in general, involve a change in input shadow-prices, hence shifts in the M.C. curves, while the slopes of the curves at Ω' were derived from a total cost curve calculated on the basis of the given, constant, Ω -values of w and r . The point is that cost curves are partial equilibrium creatures, evaluated at fixed prices, while movement along a production possibility curve involves a general equilibrium adjustment that will change input prices. Hence it is entirely possible that at say Ω' , in Fig. 2-b, both $M.C.N$ and $M.C.A$ are falling, though $F'-F'$ is convex.

So much for non-convexity. In its mildest form, if isoquants and indifference curves retain their normal curvature and only returns to scale "increase", non-convexity need not violate the qualitative characteristics of the maximum-of-W problem. The marginal-rate-of-substitution conditions may well retain their validity, and the solution still could give out a set of shadow prices, decentralized responses to which result in the maximal configuration of inputs, outputs and commodity-distribution. But certain non-marginal total conditions for effective real-life market functioning : e.g. that all producers have at least to break even, are necessarily violated. The shortcoming is in market-institutions: the maximum-of-W solution requires such "losses". The important moral is that where increasing returns to scale obtain, an idealized price system is not an effective way to raise money to cover costs. It may, however, still be an effective device for the rationing of scarcities.¹

VI

We have examined, in some detail, what conditions on the allocation and distribution of inputs and outputs can be derived from the maximization of a Social Welfare Function which obeys certain restrictions.² We have

¹ No mention has been made of the case that is perhaps most interesting from an institutional point of view: production functions that show increasing returns to scale initially, then decreasing returns as output expands further. No profit seeking firm will produce in the first stage, where A.C. is falling, and A_{Ω} and N_{Ω} may only require one or a few firms producing in the second stage. If so, the institutional conditions for perfect competition: very many firms, will not exist. One or a few firms of "efficient" scale will exhaust the market. This phenomenon lies at the heart of the monopoly-oligopoly problem.

² See the last note on p. 10

done so, however, using a statical mode of analysis and have ignored all the "dynamical" aspects of the problem. To charge that such statical treatment is "unrealistic" is to miss, I think, the essential meaning and uses of theorizing. It is true, however, that such treatment buries many interesting problems -- problems, moreover, some of which yield illuminating insight when subjected to rigorous analysis. Full dynamical extension is not for here, but a short catalogue of what such extension would consist in is perhaps warranted.

What about allocation and distribution have we ignored?

(1) The perceptive reader will have noticed that very little was said about the dimensions of A , N , L_A , D_A , L_N and D_N . The static theory of production treats outputs and inputs as instantaneous time rates, "flows" -- apples per day, labor-hours per week, etc. This ignores the elementary fact that in most production processes outputs and the associated inputs, and the various inputs themselves, are not simultaneous. Coffee plants take five years to grow, ten year old brandy has to age ten years, inputs in automobile manufacture have to follow a certain sequence, it takes time to build a power station and a refinery (no matter how abundantly "labor and land" are applied.) One dynamical refinement of the analysis, then, consists in "dating" the inputs and resultant outputs of the production functions, relative to each other. In some instances only the ordinal sequence is of interest; in others absolute elapsed time, too, matters -- plaster has to dry seven hours before the first coat of paint is applied.

(2) Another characteristic of production, on this planet at least, is that service flows are generated by stocks of physical things which yield their services only through time. Turret-lathe operations can be generated only by turret-lathes and these have congealed in them service flows which can not be exhausted instantaneously but only over time. In a descriptive sense, turret-lathe's services of today are "joint" and indivisible from some turret lathe's services of tomorrow. Strictly speaking, this is true of most service flows. But some things, like food, or coal for heating, or gasoline, exhaust their services much faster than e.g. steam-rollers, drill-presses, buildings, etc. The stock dimension of the former can be ignored in many problems; this is not true of the latter set of things, which are usually labeled as fixed capital.¹ A second dynamical extension, then, consists in introducing stock-flow relationships into the production functions.

(3) Lags and stock-flow relations are implied also by the goods-in-process phenomenon. Production takes place over space, and transport takes time, hence seed can not be produced at the instant at which it is planted, nor cylinder heads the moment they are required on the assembly line. They have to be in existence for some finite time before they are used.

(4) One of the crucial intertemporal interrelations in allocation and distribution in a world where stocks matter and where production takes time, is due to the unpleasant (or pleasant) fact that the inputs

¹ Much depends on arbitrary or special institutional assumptions about how much optimization we leave in the background, for the "engineer". For example, machines of widely varying design could very likely yield given kinds of service. "A lathe is not a lathe is ... " Further, no law of nature precludes the rather speedy using-up of a lathe - by using it e.g. as scrap metal. In some situations it could even be economic to do so. For more detailed examination of the various indivisibility, etc., aspects of capital as a factor of production, see Chapter Two .

of any instant are not manna from heaven. Their supply depends on past output decisions. Next month's production possibilities will depend, in part, on the supply of machine tools; this, in turn, depends on the resources devoted this month to the construction of new machine tools. This is the problem of investment. From today's vantage investment concerns choice of outputs; but choice of what kinds and amounts of machines to build, plants to construct, etc. today, makes sense only in terms of the input-uses of these things tomorrow. Input endowments, "L and D", become unknowns as well as data.

(5) Tomorrow's input availabilities are also affected by how inputs are used today. The nature and intensity of use to which machines are subjected, the way in which soil is used, oil wells operated, the rate at which inventories are run down, etc., partly determines what will be left tomorrow. This is the problem of physical capital consumption, wear and tear, etc. -- the problem of what to subtract from gross investment to get "net" capital formation, hence the net change in input supplies.

How do these five dynamical phenomena fit into the maximum-of-welfare problem? Recall that our W-function was assumed sensitive to, and only to X's and Y's consumption. Nothing was said, however, about the timing of such consumption. Surely not only consumption of this instant matters. In a dynamic context, meaningful welfare and preference functions have to provide a ranking not only with respect to all possible current consumption mixes but also for future time. They must provide some means

for weighing apples next week against nuts and apples today. Such functions will date each unit of A and N, and the choice to be made will be between alternative time-paths of consumption.¹

Given such a context, the above five dynamical phenomena are amenable to a formal maximizing treatment entirely akin to that of Parts I, II and III. They are, with one qualification,² consistent with the convexity assumptions required for solubility and duality. The results, which are the fruit of some very recent and pathbreaking work by Professors Solow and Samuelson (soon to be published), define inter-temporal production efficiency in terms of time-paths along which no increase in the consumption of any good of any period is possible without a decrease in some other consumption. Such paths are characterized by the superimposition, on top of the statical, one-period or instantaneous efficiency conditions, of certain inter-temporal marginal rate of substitution requirements. But the statical efficiency requirements retain their validity: for full-fledged dynamical Pareto-efficiency it is necessary that at any moment in time the system be on its one-period efficiency frontier.³

Incidentally, the geometric techniques of Part I are fully adequate to the task of handling a Solow-Samuelson dynamical set-up

¹ Note how little weight is likely to be given to current consumption relative to future consumption if we pick short unit periods. This year certainly matters, but what of this afternoon versus all future, or this second? Yet what of the man who knows he'll die tomorrow? Note also the intrinsic philosophical dilemmas: e.g. is John Jones today the "same" person he was yesterday?

² Capital is characterized not only by the fact of durability, but also by lumpiness or indivisibility "in scale". Such lumpiness results in non-convexity, hence causes serious analytical troubles. (cf. Chapter Two)

³ For possible exception to this, due to sensitivity of the volume of saving, hence of investment, to "as imputed" income distribution, cf. Part IV of Chapter Three

for a 2 by 2 by 2 world. Only now the dimensions of the production box hence the position of the production-possibility curve will keep shifting, and the solution gives values not only for inputs, outputs and prices but also for their period to period changes.

* * *

There are many dynamical phenomena less prone to analysis by a formal maximizing system than the five listed above. The qualitative and quantitative supply of labor-input in the future is influenced by the current use made of the services of people.¹ There are, also, important inter-temporal interdependences relating to the fact of space -- space matters because it takes time and resources to span it. Moreover, we have not even mentioned the really "difficult" phenomena of "grand dynamics". Production functions, preference functions, and even my or your Welfare function shift over time. Such shifts are compounded by what in a sense is the central problem of non-stationary dynamics: the intrinsic uncertainty that attaches to the notion of future.² Last, the very boundaries of "economics", as of any discipline, are intrinsically arbitrary. Allocation and distribution interact in countless ways with the politics and sociology of a society ... "everything depends on everything". But we are way beyond simple analytics.

¹ Although labor is in many respects analytically akin to other kinds of physical capital -- resources can and need be invested to expand the stock of engineers, as to expand that of cows and machines. Machines, however, are not subject to certain costless "learning" effects.

² While formal welfare-theory becomes very silent when uncertainty intrudes, much of economic analysis -- e.g. monetary theory, trade fluctuations -- would have little meaning except for the fact of uncertainty.

A Historical Note On The Literature

(For a short but substantive history of the development of thought in this field, the reader is referred to Samuelson's synthesis (non-mathematical), pp. 203-219 of Foundations. See also Bergson, "Socialist Economics", Survey of Contemporary Economics, Vol. I and Boulding, "Welfare Economics", Survey, Vol. III.)

The foundations of modern welfare theory are well embedded in the soil of classical economics, and the structure, too, bears the imprint of the line of thought represented by Smith, Ricardo, Mill, and Marshall. But in classical writing prescription and analysis are inseparably intertwined, the underlying philosophy is unabashedly utilitarian, and the central normative concern is with the efficacy of market institutions. In contrast, the development of modern welfare economics can best be understood as an attempt to sort out ethics from science, and allocative efficiency from particular modes of social organization.

The classical tradition reached its culmination in Professor Pigou's Wealth and Welfare. Pigou, the last of the great pre-moderns was also, as witness the Economics of Welfare, among the first of the moderns. But he was not the first. Pareto, writing in Italy during the first years of the century has a preeminent claim (1). It is his work, and Barone's after him (2) -- with their focus on the analytical implications of maximization -- that constitute the foundations of the modern structure. Many writers contributed to the construction, but A.P. Lerner, Abram Bergson, and Paul Samuelson come especially to mind (3). Bergson, in particular, in a single article in 1938, is the first to make us see the structure whole. More recently, Arrow has explored the logical underpinnings of the notion of a social welfare function in relation to social choice (4); Koopmans, Debreu and others are testing more complicated systems for duality (5); and Solow and Samuelson, in work soon to be published, have provided a dynamical extension (6) (7).

There is, also, an important modern literature devoted to the possible uses of the structure of analysis for policy prescription. Three separate sets of writings are more or less distinguishable. There was first, in the twenties and thirties, a prolonged controversy on markets versus government. Mises (8) and later Hayek (9)

were the principal proponents of unadulterated laissez-faire, while Dickinson, Lange, Lerner and Dobb (10) stand out on the other side. The decentralized socialist pricing idea, originally suggested by Barone and later Taylor, was elaborated by Lange to counter the Mises view that efficient allocation is impossible in a collectivized economy due simply to the sheer scale of the administrative burden of calculation and control.

Second, in the late 1930's, Kaldor (11) and Hicks (12) took up Robbins' (13) challenge to economists not to mix ethics and science and suggested a series of tests for choosing some input-output configurations over others independently of value.¹ Scitevsky (14) pointed out an important asymmetry in the Kaldor-Hicks test and Samuelson in the end demonstrated that a "welfare-function" denoting an ethic was needed after all (15). Little tried, but I think failed, to shake this conclusion (16). The Pareto conditions are necessary, but never sufficient.

Third, there is a body of writing, some of it in a partial-equilibrium mode, which is concerned with policy at a lower level of abstraction. Writings by Hotelling, Frisch, Meade, W.A. Lewis, are devoted to the question of optimal pricing, marginal-cost or otherwise, in public utility (M.C. < A.C.) situations (17). Hotelling, Wald, Mrs. Joseph, Little, and more recently Lionel McKenzie have, in turn, analyzed alternative fiscal devices for covering public deficits (18).

Bibliography

1. Manuel d'Economie Politique (1909)
2. "The Ministry of Production in the Collectivist State" (transl.) in Hayek, Collectivist Economic Planning (London, 1935).
3. See Lerner, The Economics of Control (Macmillan, 1944); Bergson "A Reformulation of Certain Aspects of Welfare Economics", Q.J.E., Feb. 1938, reprinted in Clemence ed., Readings in Economic Analysis Vol. 1; and above cited writings of Samuelson. For other works see references in Foundations (p. 219) and in Bergson's and Boulding's survey articles (cited above)
4. See Arrow, Social Choice and Individual Values (Cowles Commission #12, Chicago).

¹ The Hicks-Kaldor line of thought has some ties to an earlier literature by Marshall, Pigou, Fisher, etc. on "what is income".

5. Early focus on "duality" is due to Samuelson (Market Mechanisms and Maximization, Rand Corporation), and to a number of writers in the field of linear programming, notably Koopmans (see Activity Analysis, Cowles Commission #12, Chicago). For more recent work, reference is to be made to a yet unpublished Cowles Commission memorandum by Koopmans and Beckman.
6. Dorfman, Solow & Samuelson, Linear Programming and Economic Analysis (Rand Corp., forthcoming), esp. Chapters 11 and 12. Chapter 14 of this book contains a most elegant exposition, by R.M. Solow, of modern welfare theory in linear programming terms.
7. For a book length exposition of welfare theory see Reder, Studies in the Theory of Welfare Economics (New York, 1947).
8. For the translation of the original 1920 article by Mises which triggered the controversy, see Hayek, ed., Collectivist Economic Planning, London, 1935.
9. See especially "Socialist Calculation: The Competitive Solution" Economica, May 1940; for a broad front attack on deviations from laissez-faire see Hayek's polemic, The Road to Serfdom, Chicago, 1944.
10. Dickinson, "Price Formation in a Socialist Economy", Economic Journal, December 1933; Lange, On the Economic Theory of Socialism (ed. Lippincott) Minnesota, 1938); Lerner, op. cit.; Dobb "Economic Theory and the Problem of the Socialist Economy", Economic Journal, December 1933.
11. "Welfare Propositions in Economics and Interpersonal Comparisons of Utility" Economic Journal, LXIX.
12. "The Foundations of Welfare Economics", Economic Journal, LXIX and "The Valuation of the Social Income, Economica, 1940
13. The Nature and Significance of Economic Science (Macmillan, 1932).
14. "A Note on Welfare Propositions in Economics", Rev. of Econ. Studies, 1941, and "A Reconsideration of the Theory of Tariffs", Rev. of Econ. Studies, 1942
15. "Evaluation of Real National Income", Oxford Economic Papers, Jan. 1950
16. A Critique of Welfare Economics, Oxford.
17. Hotelling: "The General Welfare in Relation to Problems of Taxation and of Railway and Utility Rates" Econometrica, Vol. 6 (1938) is the first modern formulation of the problem. For a complete discussion of the issues and the literature, and for references, see Nancy Ruggles, "The Welfare Basis of the Marginal Cost Pricing Principle" and "Recent Developments in the Theory of Marginal Cost Pricing", Review of Economic Studies, XVII, #42 and #43.
18. See esp. Little, "Direct versus Indirect Taxes", Economic Journal September 1951.

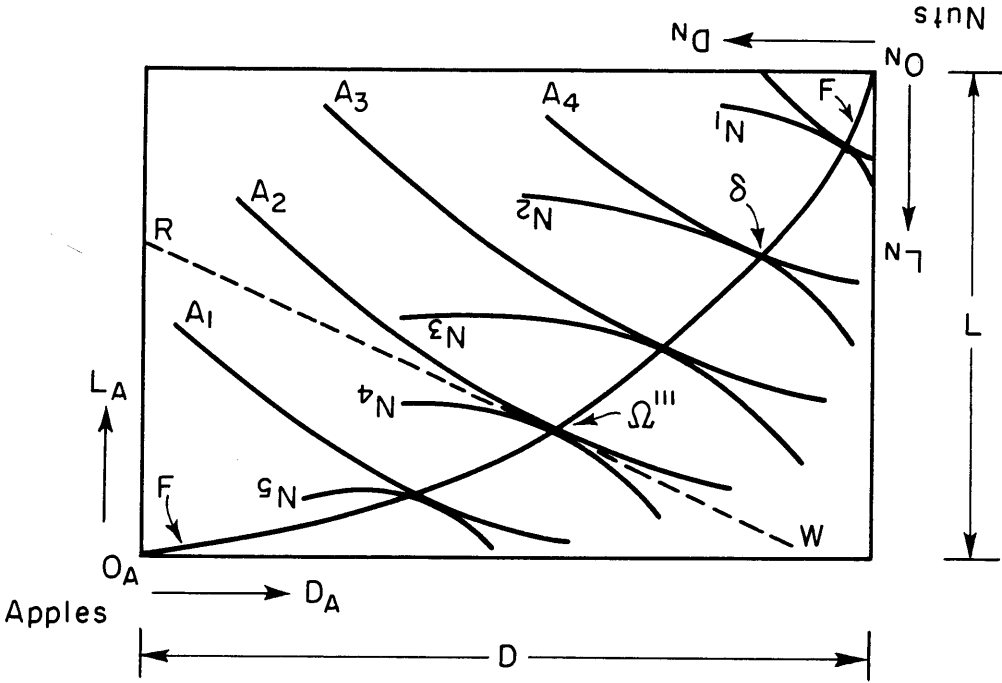


Fig. 1

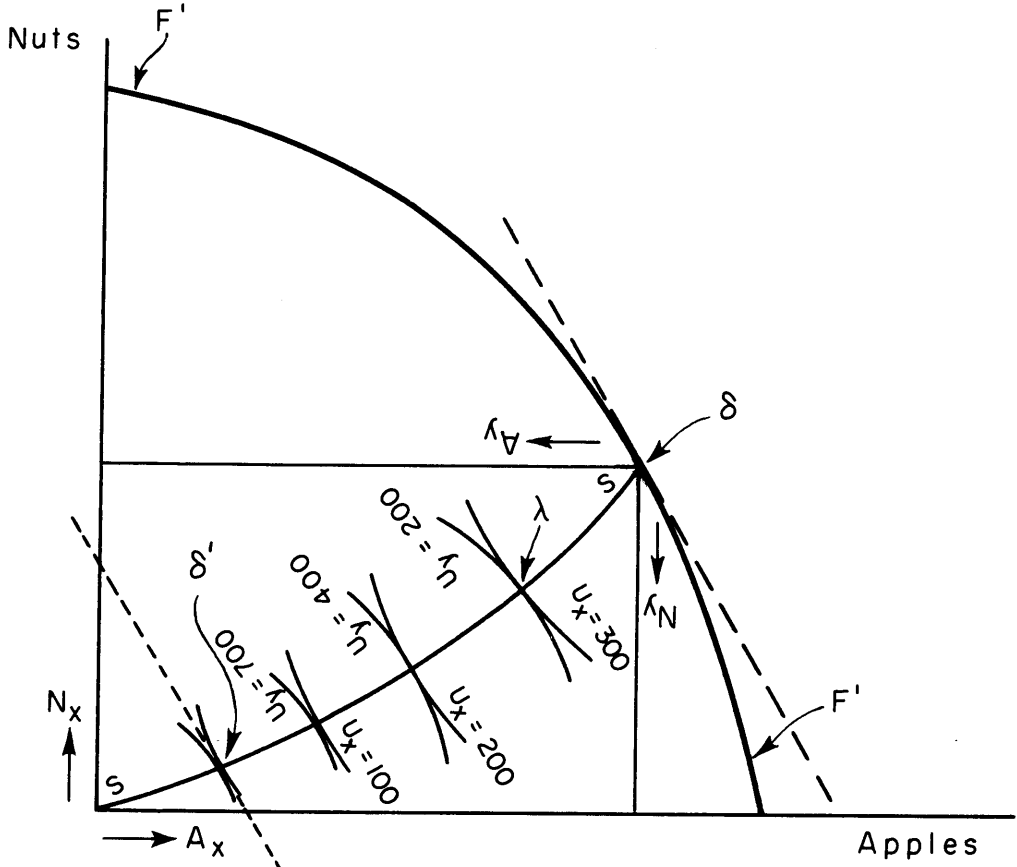


Fig 2-a

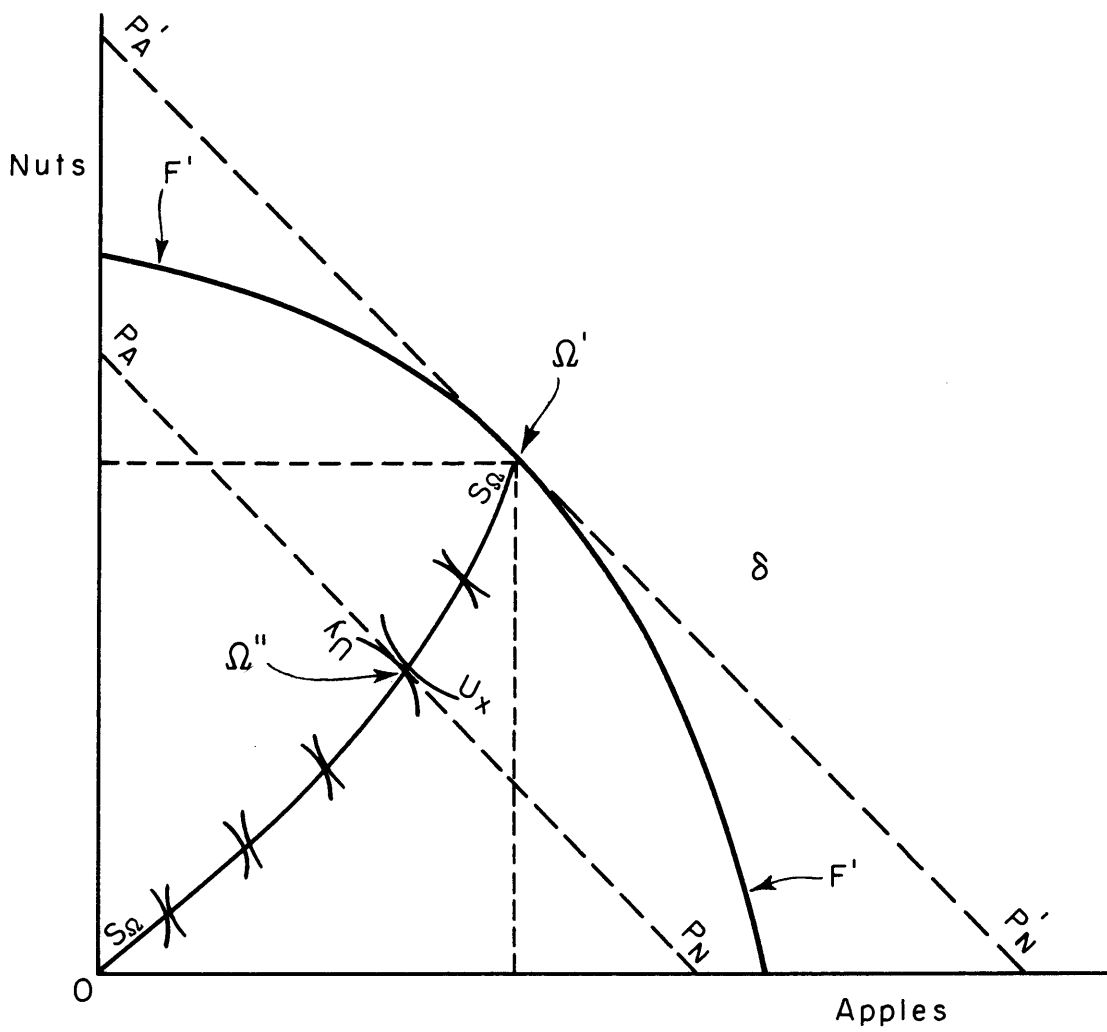


Fig. 2-b

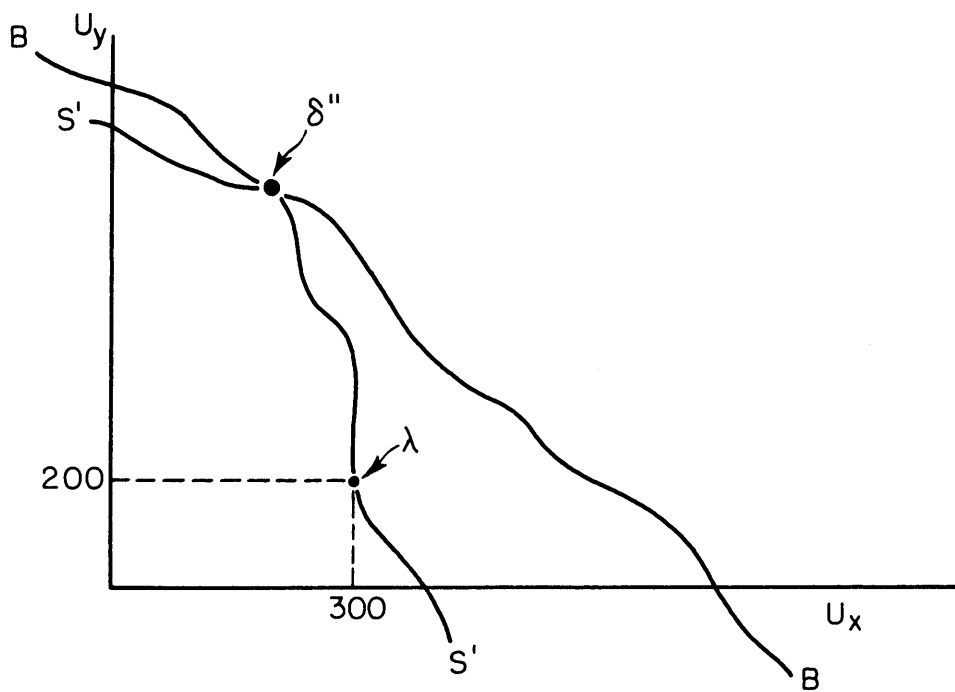


Fig. 3-a

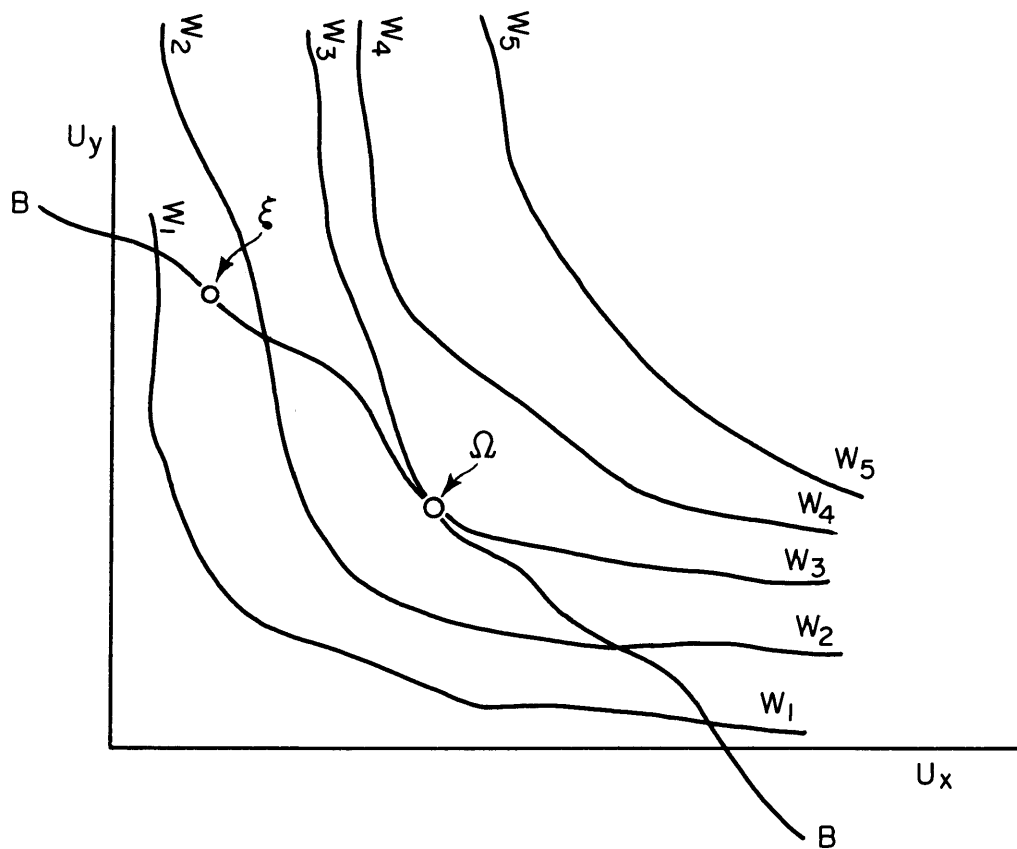


Fig. 3 - b

CHAPTER TWO

ELEMENTS

OF THE

PURE ECONOMICS OF "SOCIAL OVERHEAD CAPITAL"

ELEMENTS OF THE PURE ECONOMICS OF "SOCIAL OVERHEAD CAPITAL"

The literature on industrialization and economic development assigns great import to what Prof. Rosenstein-Rodan has labeled "social overhead capital". Yet there does not appear to exist in that literature, any systematic and comprehensive exploration of what it is -- if anything -- that makes such things as roads, dams, harbors, postal facilities, somehow different in kind from rolling mills, refineries, trucks, fertilizer plants and the like.¹ This chapter attempts to provide the theoretical foundations for such an exploration. Its purpose is to sort out the various mutually reinforcing and overlapping elements that underlie the notion of "social overhead": increasing returns to scale, indivisibility, durability, "external" interaction, non-appropriability, the "public", non-exhaustible quality of the output ... The mode of analysis is that of modern allocation theory: it provides the analytical techniques and fixes -- for better or worse -- the level of abstraction.

Specifically, Part I is an introductory examination of the various qualities by which "social overhead" capital has been identified and of the criteria against which the consequences of social overhead phenomena are to be tested. Part II explores the content of what is

¹ There is, however, a substantial literature on pricing of public utility services. While much of this takes a modern industrial context for granted, it is nevertheless of considerable relevance.

perhaps the most significant defining quality: increasing returns to scale and/or indivisibility, and traces its implications for efficient allocation in a stationary and statical context of perfect information. Part III consists, first, in an expository digression on the doctrine of external economies; second, it suggests a particular ordering of externality phenomena. This last is designed to clarify the ways in which externalities associate with "social overhead" and to identify the links between the externality and the indivisibility aspects. Use is made, in the ordering, of the recently revived "public good" notion of public expenditure theory.¹ Finally, Part IV suggests directions for future research.

The following is a skeletal table of contents:

- I On Definition and Identification
 - Test of Market Efficiency
 - Identification
 - The Efficiency Test Again
 - What about Growth Potential?

- II Indivisibilities and Increasing Returns to Scale
 - A. Rudiments
 - Durability
 - Lumpiness
 - More on Indivisibility
 - B. Indivisibility and Production Functions
 - C. Indivisibility, Efficiency and Markets
 - "Apples and Bridges"
 - Modifications
 - Institutional Implications
 - On Optimum Conditions Again
 - Durability
 - D. Indivisibility in Scale and Social Overhead

(CONT'D)

¹ By Prof. Samuelson, in the November 1954 and 1955 issues of the Review of Economics and Statistics.

- III External Economies and Social Overhead
 - A. A Digression on External Economies
 - By Way of Some History
 - The Modern Formulation: Apples and Honey
 - Comments
 - Some Queries
 - B. Externalities: An Ordering
 - Type 1: Organizational Externalities
 - Type 2: Technical Externalities
 - Type 3: Public Good Externalities
 - Comments
 - 1. On Type (1)
 - 2. Prof. Meade's "atmosphere"
 - 3. Prof. Samuelson on Types (2) and (3)
 - 4. Exclusion
 - 5. Arrangements to offset
 - 6. On Blends
 - C. Back to Social Overhead
- IV On the Economics of Social Overhead Capital: Agenda

* * *

It is evident from the above that the following is a statical analysis of the consequences for efficient allocation of certain defining qualities of social overhead phenomena, in a stationary context of perfect information. Needless to say, such analysis can provide only the elements of an economics of "social overhead". Even within the confines of allocation theory, narrowly conceived, complete treatment would require exploration, for instance, of problems of adjustment. This would involve, besides such statical questions of multiple "local" optima as are treated below, dynamical considerations of stability. The assumption of perfect information would require weakening. Further, if durability were to be given

its due, explicitly dynamical analysis would be required of efficient paths of outputs and inputs, with investment decisions treated as such. The grand dynamical fact of uncertainty, too, would have to be faced. Most important, the focus of attention would have to shift from "efficiency" to growth: What are the qualitative effects of "social overhead" phenomena on growth-potential? Do they give rise to or reinforce some of the vicious circles that enchain a low income economy?

But first things first: we must to the elements.¹

¹ It should be said here that since the purpose of the chapter is to bring into focus whatever available theory may be relevant to social overhead phenomena, much of what follows inevitably covers well-known ground. This is especially true of Part II on Indivisibilities. In order, however, not to lengthen what is already too long, no acknowledgment is made to the literature for propositions in general currency.

I ON DEFINITION AND IDENTIFICATION

What is it we mean by "social overhead capital"? Crudely put, the label is usually applied to some substantial physical installations, supposedly "required" for the production of certain goods or, more commonly, services whose production, moreover, is presumed to "require" the involvement -- as investor, or operator, or regulator, or subsidizer -- of government. More generally, it is useful to think not of the facilities as such, but of social overhead activities, such as "require" both large physical installations and government involvement. But whether we think of installations or activities -- the latter view will, in general, be adopted here -- this is both too broad and too crude a formulation. It is too broad because it encompasses all too much of government activity: it is surely not for the notion of social overhead to bear the full load of even an economic theory of government. And it is too crude, in that to have meaning, the word "require" has, in both instances, to be given content.

Since it is not the purpose of this paper to achieve an arbitrarily precise delineation of which of a government's activities are to be labeled "social overhead", no attempt is made formally and rigorously to reduce the breadth of coverage. The discussion has

most bearing on certain conventionally "economic" activities: transport, communications, power, and the like, but at least heuristic extension to national defense, police functions, etc. will be possible and perhaps even suggestive. But whatever ambitions one may entertain as to the range of relevance of the notion of "social overhead", it needs substantial refinement if it is to be of analytical use. We must, for one, invest with precise content the nature of the requirement for government involvement.

The Test of Market Efficiency

On the level of traditional market-economics, "requirement" for government involvement in an activity means only that a more or less idealized system of price-market institutions must be expected to fail in relation to that activity. Typically, such failure consists in "desirable" activities not being carried on at all, or not being carried on at the "correct" scale.¹ The "desirability" of an activity, or its "correct scale", in turn, are evaluated relative to the solution values of a maximum-of-welfare problem.

It is the central theorem of modern welfare economics that under certain strong convexity and independence assumptions about technology, tastes, etc., the equilibrium conditions that characterize a system of perfectly competitive market institutions (price equal to marginal cost everywhere, etc.) will correspond precisely to the conditions

¹ Failure could also consist in the flourishing of undesirable activities, but "negative" and preventive interferences are clearly of lesser relevance to our central theme -- though on application the distinction tends to blur.

implied by a Paretian maximum problem. Further, if competitively imputed incomes are continuously redistributed in costless lump-sum fashion so as to achieve the income distribution implied by a specified or implicit social welfare function, then the competitive market solution of inputs, outputs, and commodity distribution, as sustained by atomistic profit maximizing producers and preference maximizing consumers, will exactly correspond to the electronically calculated solution which maximizes, subject only to tastes, technology and initial endowments, that particular welfare function. This duality theorem¹ holds for the statical steady-state flow model of the Walrasian sort where the solution values are stationary time-rates; it holds, also, for dynamical systems involving capital formation (given, still, convexity throughout). For these last, the solution values are time paths of inputs, outputs, etc. One test, then, of social overhead is the failure of idealized market institutions to sustain (or stronger) to lead to Pareto efficiency.

There are two difficulties with this test. For one thing, it is not a test that permits of easy application even in concept.²

¹ For the relevant literature, see Appendix to Chapter One.

² While a relevant criterion, "easy" applicability, even in concept, does concern a different level of discourse. With perfect information and a Laplacian I.Q. the distinction between deduced consequence and definition vanishes. But much of economic theory is useful because, in a lesser state, and with feasibility limitations on experimentation, it is often easier to confront the implied consequences of a phenomenon, as deduced, with the facts of reality, than to so confront the defining quality itself.

How are we to "pick out" the activities that make the "invisible hand" fail, or fail worse than it does in guiding other activities? Moreover, it is much too broad a test. Many things in the "real world" will yield such market failure: uncertainty, imperfect information, lagging adjustment to changing "data", etc. and not all are central to the notion of "social overhead". By what additional tests are we to identify what kinds of activity are "overhead" to "society"?

Identification: Externality and Indivisibility

In a sense, such identification appears almost superfluous. If asked to make a list, most economists would more or less agree on at least a few basic items: schools, roads, irrigation dams, water mains, etc.¹ A few pedantic souls might quibble about the vagueness of the notion, but a measure of unanimity would likely prevail. What then are the qualities of such facilities that suggest the label "social overhead"?

One analytically ambitious notion is that we should search for social overhead activities among those which come at the "early" stages of "the" structure of production. Unfortunately, the distinction between "earlier" and "later" stages of production, though perhaps suggestive in a rough qualitative way, tends on examination to blur. If the pattern of interproducer flows is Leontief-like rather than triangular ("Austrian") -- i.e. if you can not so order industries in a Leontief transactions table as to fill most boxes above the diagonal

¹ What to do about various agreed items on the list would not be agreed so readily. But the political economy of public services is not here at issue.

with zeros -- most activities (in this case industries) will not be in a uniquely "earlier" or "later" relationship to each other. Coal is used not only to heat homes, but also to produce steel required in the production of mining machines designed for the production of coal. Further, even if one could in any concrete situation identify in a rough and ready way certain industries as producers primarily of consumers' goods (automobiles wouldn't quite do, except perhaps in Saudi Arabia, though where business expense accounts are limited *pâté-de-foie gras* might), it is not clear that these could safely be put aside in a search for social overheads. One has only to think of public parks, or museums, or "vacation land" parkways (no trucks allowed!) -- some economist's list is likely to contain these and similar "consumer" items. Conversely, though pig iron production comes "early", it need not, thereby, appear on any man's list. If transport and communication are overheads to society, it is not because they come at some "earlier" stage in a "vertical" sequence of production.

What then, are the attributes by which we might identify "social overhead" activities? Looking to the development literature, or the obviously germane writings on "public utilities", two dominant themes emerge: (1) external economies and (2) indivisibilities. The first, external economies, would, in its usual formulation, have us emphasize activities or industries, home production in which yields certain "external" benefits that are not "easily" or "costlessly" appropriated through market institutions, hence where decentralized choice, based

on price and profit considerations, will not lead to optimal results. This is an important theme -- but it raises a host of questions. How, for one, do the externality phenomena cited in the development literature in connection with social overheads relate to the neo-Marshallian construct of equilibrium theory? What is the meaning and import of non-appropriability, of non-exclusion? In what sense, if any, are the benefits associated with social overhead more "external" in nature than those yielded by standard manufacturing activity? Moreover, how do externalities link with the second theme, indivisibilities?

This last is of importance because of the virtually unanimous view, suggested, one supposes, by the emphasis on large physical installations, that a necessary though not sufficient attribute of activities that are to be listed as "social overhead" is that they involve substantial indivisibility in inputs, processes or outputs. In fact, lumpiness is often held to be the crucial identifying quality. But what makes for lumpiness and what are its special consequences? In what sense do they render an activity social and overhead?

Our procedure, in what follows, is tentatively to adopt these two themes, and explore in what sense they can be thought to constitute the elements of an economics of social overhead. Part II is devoted to indivisibility; externalities are examined in Part III.

The Efficiency Test Again

Our task, then, is to explore the content and consequences of indivisibilities and externalities, especially as these relate to our

original test: the Pareto-efficiency of markets. Yet it is unduly restrictive to base the "efficiency" test exclusively on market institutions. No matter how idealized such markets may be, analysis would be enmeshed, and needlessly, with particular forms of social organization. By taking a slightly different view of the fundamental "duality theorem", we can be rid of the institutional content and reintroduce it only as seems useful.

The analytical essence of the theorem lies in the by now well understood though still remarkable fact that with all-around convexity the technocratically formulated, institutionally neutral, Paretian maximum-of-welfare problem contains embedded within it a set of constants: "duals", Lagrangean multipliers, shadow-prices, which have all the analytical characteristics of prices, wage rates, rents, interest rates. Correspondence between Pareto efficiency and market performance implies, at the least, that decentralized decisions in response to these "prices" by atomistic profit and satisfaction maximizers result in just that constellation of inputs, outputs and commodity-distribution, that the maximum of the specified W-function calls for. Our initial dictum that social overhead phenomena "require" government involvement, means simply that by definition such phenomena destroy this correspondence. But note the various refinements the reformulation suggests:

- (1) Can the set of Pareto-efficient input-output-distribution points, in a world of social overhead phenomena, be characterized by some efficiency conditions that bear an analogy

to the marginal-rate-of-substitution conditions of a convex neo-classical universe?

- (2) What effect do social-overhead situations have on the existence, associated with each of the infinity of Pareto-efficient input-output-distribution combinations, of a set of shadow-price-like constants?
- (3) Should such associated sets of Lagrangean parameters exist, would decentralized atomistic equating of price to marginal cost everywhere in response to any one set of these yield a local profit maximum position for each producer, or would the implied inputs and outputs be points of minimum profit for some?
- (4) Even if every efficient point implies local profit maxima throughout, do such profit maximum positions give non-negative profits for all producers from whom positive production is required? Or do some efficient points imply some production at continuing losses?¹
- (5) Does each set of shadow-prices associate with just one efficient set of inputs and outputs, or do some production points share the same set of shadow prices? If there is such sharing, are all the production points associated with any one set of relative prices points of maximum and non-negative profit (as per (3) and (4)), or are they mixed?²

¹ For a more precise formulation, see Part V of Chapter One.

² Where indivisibilities obtain, this becomes a question of first order importance.

Each of the above questions could refer to a statical setting, where the solution values are unchanging instantaneous time rates, or to a dynamical though stationary system, where the answers are time-paths of inputs, outputs, prices, and where profit-maximizing gives way to maximizing of present discounted value, etc. But whether in a statical or dynamical context, the questions are all relevant to whether a decentralized price market "game" -- perhaps "for real" by genuine profit seekers, or by socialist civil servant plant managers following an injunction to "maximize profits", or, perhaps for no less "real", by technicians following out a computing routine -- will or will not sustain a Pareto-efficient configuration once the shadow-prices associated with that configuration are specified.¹

Questions (3) and (5) bear also, on a different and even more complicated set of problems:

(6) Will some "natural" price-market type computational game -- there are an unlimited variety of these, from Edgeworth's recontracting device to those of the Chicago commodity markets -- converge on the set of Pareto-efficient solutions (or the one solution implied by a specified welfare function, or that implied by a given initial ownership of productive resources)? This concerns movement to a solution from positions of disequilibrium. If the answer to (3) in a given situation is that the solution

¹ The fact that such a "game" may be efficient does not, of course, preclude the existence of other efficient "games". But none has been invented that has as relatively workable an institutional counterpart.

positions are all locally stable profit maxima, then there exists some routine of price and quantity responses that will lead to the solution, if, that is, the initial position is in the neighborhood of the solution.¹ But what if the initial configuration is random. What about stability in the large?

(7) Will some institutionally germane version of a decentralized market set-up tend to track the solution? Under what assumptions concerning the quality and extent of information available to each "player"?

This last set of issues -- "adjustment" dynamics -- have to do with the full dynamical stability of interconnected markets in systems where, in the general case, the equilibrium constellation is defined not by points but by time-paths. Here they will hardly be touched on. But it is important to keep in mind that much of the reality of social overhead phenomena lies in the realm of such "disequilibrium" dynamics.

So much for the various refinements of our original efficiency test. It will turn out that such refinement permits resolution of a number of confusions relating especially to the externality aspects of social overhead. On the other hand, all this has to do only with Paretian efficiency. What of growth? After all, the social overhead notion has played a significant part in, and only in, the development

¹ This puts the cart before the horse. The point is that we cannot assert statical stability in the small (3) without in effect prejudging dynamical stability (also in the small). The latter is a necessary condition for the former.

literature. What do overhead type phenomena imply for growth potential?

It is important, in this connection, to keep straight the relations between growth and Pareto-efficiency. Such efficiency is neither a necessary nor a sufficient condition for growth. In the context of the usual maximizing set-ups of economic theory, it is, in general, a necessary condition for non-wasteful growth.¹ From a non-efficient path, that is, the system can ex ante always improve its performance and get more current consumption or a faster growth-rate (more potential future consumption. But efficiency is never sufficient. It may be that with the specified consumption profiles or intertemporal preference patterns, no feasible time-path yields a positive rate of growth of whatever index of output (or output per head). Efficient paths will, in general, still do better than non-efficient paths, but none need yield expansion. Hence it must be asked: What qualitative effect on growth-potential can generally be attributed to social overhead phenomena other than their implications for market efficiency or for the solubility of a maximum problem? Formal analytical treatment becomes very difficult indeed; some useful generalizations may nevertheless be possible. But this takes us beyond the confines of this paper. Here we must to social overhead and indivisibility.

¹ Though for possible conflict, see Part IV of Chapter Three.

II INDIVISIBILITIES AND INCREASING RETURNS TO SCALE

A. Rudiments

Customarily, the "social-overhead" designation is applied to activities associated with sizeable fixed investment in physical installations. Much of public utility economics takes as its starting point a total cost curve with a barely rising variable component superimposed on a large initial "fixed" lump which is insensitive to output. But what makes for the "lump"?¹ More precisely, why can it not be made "very small"? Why can't we have a "little" bridge if only a few bridge-crossings are desired?

A superficial answer is that the equipment or facilities associated with the activity are themselves intrinsically lumpy. To provide bridge-crossings, you need a whole bridge: one-third of a bridge will not do. It would not be claimed, of course, that bridges, dams, water-mains and railroad-tracks are wholly immutable. Choice between bridges for pedestrians and bridges for ten-ton trucks, four-inch mains and twelve-inch mains, would be admitted. But all such facilities exhibit some indivisibility in one or another of their qualities.

Durability

Accepting this for the moment -- we shall probe deeper presently -- it is useful to distinguish between two crucial kinds of quantitative indivisibility associated with lumpy facilities. One, such facilities

¹ One can, of course, simply accept it as ascertainable fact, and use it as a datum. But much interesting insight lies behind cost curves.

represent, as it were, congealed services which have the peculiarity that they cannot be exhausted (used up) instantaneously or within just any arbitrary short period. More precisely, there is a more or less maximum time rate at which a lumpy facility, a human being for instance, will yield its potential services, and this maximum rate does not instantaneously exhaust such services.¹ A light-bulb is a stock of n hours of congealed services. But in any one hour (and as qualified below) the maximum yield of these services is $\frac{1}{n}$ th of the total. This is not true of a stock of gasoline or of chocolate bars: you "can" use all or any fraction within an hour.

This kind of indivisibility through time: durability, is the defining characteristic of what is generally called "fixed" capital.² The central quality of such capital can be re-stated with the aid of the usual capital theory diagram (Fig. 1). The vertical axis measures positive and negative outputs: the horizontal represents elapsed time. \dot{x} is the maximum instantaneous time rate of service yield. The meaning of indivisibility through time is simply that there are binding limits to shifting a slice of the far end of the output-tail towards the left. To a closed community, the far end of the tail is not fungible through time.

¹ But see qualifications below.

² To the extent that unfinished but divisible goods in process, or stocks of non-durable finished goods "can't" yield their services instantaneously it is because of a ceiling on the use-rate of complementary "fixed" capital inputs which have to be combined in finite quantity with the non-durable good: e.g. engines with gasoline, or human beings with chocolate bars. Note, incidentally, the arbitrariness of the "fixed" tag. Is an item "fixed" if it yields its services in a day, five years, an hour?

Of course, none of the above is strictly true. It is difficult to think of anything in the real world that can not be "used up" very fast indeed. You can quickly "exhaust" a new drill press by using it as scrap -- in certain situations it may even be economic to do so. And at worst, nothing prevents us from dumping trucks into the ocean. Again, putting a man to forced labor radically alters both the time-shape and the qualitative content of his service outputs.¹

These are extreme examples, but they suggest the slippery nature of the time-indivisibility notion. There are economic choices implied by using a piece of equipment in a way that sets a particular maximum to \dot{x} in Fig. 1. Further, there are economic choices in using inputs to construct such durable items to start off with. But in the end, some element of time-indivisibility remains that has its roots in the physics and topography of the universe.

Before turning to the second category, mention should be made of an other, related manifestation of indivisibility through time. In Fig. 1 - type representation of many situations -- and this is especially relevant to social overhead -- not only is the output tail prolonged; so is the input tail. It takes time to build a power station or to lay railroad tracks, no matter how abundantly labor, equipment and material inputs are applied. Choices are again implied,

¹ Such variability in the rate at which the "services" of even very fixed capital are exhausted lie at the heart of the reasons why a gross product notion in national accounting is necessarily arbitrary. Technology sets a limit to what is producible in a year only if inputs are specified. It is the essence of the pseudo-fixity of most kinds of capital stock that the service-flow inputs they provide are subject to choice.

but again only within limits.¹

Lumpiness

There exists, second, a different and in many respects more troublesome species of indivisibility: indivisibility-in-scale. In terms of Fig. 1, the meaning of such lumpiness in scale is simply that for some kinds of service (e.g. bridge crossings), \dot{x} , the maximum service yield per unit time, is not continuously variable. Many kinds of service-flows, e.g. many human functions and turret-lathe functions, require the existence of a physical stock of something, and the "size" of such stock, as measured by "capacity" output rate: \dot{x} , can not always be finely adjusted. For a given kind of service output, it may be that the choice is between, say, three sizes of a piece of equipment, with no gradations between them. If the service-streams of a series of e.g. machines are assumed additive, this means that there exists a finite sized "smallest" machine, whose maximum \dot{x} is the minimum of all maximum x -s for that kind of service.

It is true that in many cases of lumpy facilities there exists an option not always to make use of the obtainable output stream at the maximum time-rate. Bridges can be left uncrossed and railroad cars half-empty, etc... But the essential point is that if anybody is to get a bridge crossing, if anyone is to ride in the railroad car from Boston to New York, the bridge or railroad services yielded him are joint and indivisible from the service potential of the empty seat-space

¹ In some cases the limits are very inflexible: seven year old wine has to age seven years. Note, incidentally, that input or output streams are not ordained constant: Fig. 1 could show non-continuous and variable profiles.

and bridge space. You can make a bridge smaller and smaller, but in the end it has to cross the river, and if one man can pass, so can several just behind and just in front of him.¹

Note that the point is not that there is a fixed time-cost waste in not fully using an existing railroad car or a machine.² It is, rather, that even if you can precisely predict the traffic, you can not finely adjust your bridge or car design to match it.

Again, the arbitrariness of almost every particular manifestation of this kind of lumpiness-in-scale is plain. The bridge has to reach across the river, but availability of bridge surface for more than one car at a time does not mean that the structure could support the weight of more than one car. Anyway, why can't railroad cars and turret-lathes be made much smaller (as measured, let's say, by \dot{x})? It is true that much of man-made equipment is lumpy in scale, but present engineering designs are not necessarily immutable facts of nature. There is nothing intrinsically and a priori economic about large scale and lumpy modern plant and equipment. It all depends on the relative input scarcities, as well as on technology. Bridges are lumpy, but what about row boats, or a good swimmer with a bundle on his head? One has to go deeper than examples of lumpy pieces of equipment to see that indivisibility in scale, as indivisibility in time, while matters of degree, adjustment and, within limits, of choice, are nevertheless in the nature of things.

¹ Maybe ... see qualifications below.

² The interest cost imputable to congealed services that are not being used, or more even, any existence-cost that is a function of time alone (e.g. evaporated alcohol).

More on Indivisibility

This is hardly the place for exhaustive exploration of the deeper qualities of the universe that make for "indivisibility"¹. But a few points warrant mention.

(1) It is a mistake to think of indivisibilities as attributes only of inputs. For one thing, Crusoe's inputs are Friday's outputs (is the bridge an input or an output?) Further, in many cases indivisibility is best associated, at least heuristically, with a process and not any of the material components.² This is true, for example, of the lumpiness quality of a great many "once for all" functions required by production processes, such as starting up machines, setting books in type, etc.

(2) The above suggests the fundamental point that in the end indivisibility phenomena need not arise due to physical and chemical singularities, thresholds, etc. Non-linearities of a smooth sort could cause it all. The fact, for example, that the area of an r -inch sided square is r^2 , while its periphery is $4r$, suggests that a "perfect" division of a square acre is impossible. If only area matters, it is easy to do; the same, if only the periphery. But if both do (as they

¹ Some may well feel that even the above kind of classificatory exploration is a waste of time. Let the engineer give us production functions and we can try to live with its non-convexities. Tastes are various.

² In a note in the November 1955 issue of the Quarterly Journal of Economics, Professor Leibenstein has questioned whether there is any possible meaning in the association of the noun "process" with the adjective "indivisible". To him, it appears, saying indivisible process is like saying "female sky". I should not care to engage in scholastic argumentation but tend to the view that e.g. the processes of heating or cooling are divisible, in that they are subject to continuous variability along at least one of their relevant dimensions. Freezing, on the other hand, is a "discontinuous" process -- more precisely, there is a sharp "kink" at 32°F. But this kind of metaphysics is probably to be resisted.

do to a farmer who has to worry about fencing and plowing) perfect halving is intrinsically impossible. The two relevant dimensions do not relate linearly. This kind of difficulty occurs whenever more than one attribute of an input or an output matters, unless, of course, they are linearly linked.

(3) But we know that there do in fact exist singularities in the universe. Modern as classical physics have their constants: speed of light, the gravitational constant, etc. These, too, make for "indivisibility phenomena". And it is in large part arbitrary, and for many (not all) purposes irrelevant whether smoothly differentiable but non-linear functions or discrete step-functions with all sorts of gaps, best explain the very real phenomena which up till now we have called indivisibilities but which could equally well manifest themselves in smoothly increasing returns to scale.

But that such phenomena are real and important is hard to gainsay. As economists we can cajole or bully design engineers into designing processes, installations, etc. that save on congealed inputs and have smaller \dot{x} -s (Fig. 1), in particular when designing for low income countries.¹ Or we can turn into engineers and do our own designing so as to optimize in terms of the real resource scarcities of a given situation (programming techniques are obviously relevant). But the economically perhaps arbitrary, not completely nature-imposed quality of the indivisibilities associated with "standard" designs and ways of doing things, should not blind. Non-linearity and lumpiness are evident

¹ If this does not involve such a large increase in the input requirement for some other scarce resource, e.g. skilled manpower, that it turns out worse than the original set-up.

facts of nature.

So much for the qualitative nature of the indivisibility phenomena by which social overheads are often identified. To examine their consequences we must somehow introduce lumpiness elements into our formal analytical set up.

B. Indivisibilities, Increasing Returns to Scale and Production Functions.

Our task is to explore the consequences of generalized indivisibility phenomena in terms of the various efficiency criteria listed in Part I above, in context, as of the first instance, of a statical model of production. Use of a statical mode of analysis does, of course, preclude exploration of the consequences of indivisibility-through-time. Durability is a peculiarly dynamical phenomenon: it provides one important link between tomorrow and yesterday. There are many statical ways of begging its inter-temporal implications -- taking a long enough unit period to permit wearing out of all capital stock is a common one. Here we shall adopt the usual statical device of periodization by depreciation. This involves the arbitrary association of a particular fraction of what are intrinsically joint inputs with the output of a particular period, thus getting a per-period flow-cost of capital (stock) input. Since this is equivalent to assuming that producers rent all their capital equipment on life-time contract, we shall speak of the per-period price of capital as a gross rental rate made up of a depreciation and an interest charge.

Attention, then, will be exclusively on indivisibility in scale. There is a justification for this beyond analytical convenience. It happens that the crucial analytical consequence of indivisibility: non-convexity, is a consequence of scale-indivisibility alone. As will be shown below, durability, as such, does not violate the convexity requirements to which the efficiency tests are so sensitive.

How, then, are we to introduce scale-indivisibility? This could be done by injecting a large constant component into some total cost curves (long run).¹ It is of interest, however, to start further back, with production functions. What are the qualitative effects on a production function of scale-indivisibility in an input, or in the process, or in the output itself?

Exploration is easiest in context of the simplest possible one input-one output world and the qualitative results apply equally well to the many input-one output case.² In Figs. 2-a, b, c, are plotted three production functions which exhibit explicit discontinuities in input, output, and process. The time rate of output flow of a single homogeneous commodity, \dot{Q} , is made a function of the stock input K . In 2-a the indivisibility is in K : there are horizontal gaps because fractional units of K (e.g. of bridges) have no meaning. In 2-b the gaps are along the \dot{Q} -axis. This means that fractional units of output (e.g. of books) are ruled out. In 2-c, finally, there are no gaps, only sharp kinks due to what "must" be process indivisibilities.

¹ This would not be a true fixed cost: at zero output it would not be incurred.

² The many-input case does exhibit difficulties, however, that do not appear in a one-input set-up: e.g. isoquants with the wrong curvature. There will be more on this below.

The above implies that it is obvious in every case what about an activity is indivisible. A little thought suggests that in fact the input-output-process distinction involves a good deal of arbitrariness. Is the final product a bridge or bridge crossings? How are inputs defined: by some physical function in a particular process (elevator operator), or by some quality of the entity that serves the particular function (semi-skilled male)? Since there is no general rule about which one of the many relevant dimensions of an input or an output are to go on the axis, and since the same object can have both some lumpy and some finely divisible qualities, arbitrariness is intrinsic.

Matters are even more confused. There are no obvious and generally valid rules about which of the circumstances influencing production are to be treated explicitly as inputs. Are various qualities e.g. of the atmosphere, to be given axes, or are their effects to be built into the curvature properties of the function? There are many strategic rules which help sensible decision in particular and concrete cases: the stability of the particular circumstance, the ability of the decision maker in question to influence it, the general invariance of the function, etc. Further, we have a professional stake in putting identifiable and exchangeable (though not necessarily only purchased and sold) inputs and outputs on the axes, since it is in their combinations and allocations and scarcity values that we are interested. But whatever alternative is adopted, it is virtually impossible, and anyway pointless, to disassociate production-functions from particular decision-making institutions.

But if production functions, and perhaps choices between various types of indivisibility, are to be sensitive to arbitrary institutional considerations, will our results also be so sensitized? It happens that none of the essential general consequences of scale-indivisibility are affected by what in an activity is indivisible.¹ The essential quality which is shared by all the functions of Fig. 2 is that the set of feasible points in production (input-output) space is non-convex. The connecting straight line between any numbers of pairs of feasible points passes outside the feasible set. This is the one common quality that is necessarily implied by indivisibility-in-scale, and it is from it that flow all the subsequent results.²

This suggests a final point. What if a function shows smooth, non-kinky non-convexity of the increasing returns to scale variety, as in Fig. 2-d? This is clearly non-convex, yet there are no corners. It characterizes situations such as where output varies as the volume of a regular container, while e.g. material input varies as the surface area. Output then grows as the cube of the linear dimension, while input varies as the square. But there are no discontinuities. If non-convexity is the crucial factor, then such smooth functions belong in our generalized indivisibility-in-scale-category. And in fact, for

¹ With more than one input, pure output indivisibility leaves intact some efficiency conditions on inputs. For more on this, see the section on "Modifications" in Part II-c below.

² On the definition of convexity, see Chapter One, Part V.

purposes of the statical analysis, there is no significant difference between kinky and smooth non-convexity.¹ The correspondence is useful to us since it assures applicability of the analysis of increasing returns to scale in Part V of Chapter One.

C. Indivisibility, Efficiency and Markets

The consequences of non-convexity in input-output space for a community's production possibility curve and for individual producers' cost curves were discussed in some detail in Part V of Chapter One. It was there shown:

- 1) That a production point lying in a non-convex region of a production function necessarily implies an associated average cost curve (A.C.) that is downward sloping at that level of output;
- 2) That the associated marginal cost curve (M.C.) while of course below A.C., could at that particular point be falling or rising;
- 3) That the production possibility curve of the community could be convex or not depending on the extent of technological non-convexity throughout the system, on input endowments, etc.

The implied conclusions, as developed in Chapter One, were as follows:

¹ For purposes of growth-analysis, a function which has zero-slope in the output dimension at some initial point (the origin) may have special significance, particularly if there is no output yield to finite quantities of inputs. Indivisibility may then imply something beyond smooth increasing returns to scale.

1) If the maximum-of-welfare configuration (for any specified welfare function) calls for any output points which lie in non-convex regions of production functions, decentralized profit seeking producers, responding competitively to the centrally quoted and correctly calculated set of shadow prices, will not sustain the required levels of output. At the quoted shadow prices such output points imply $M.C. = p < A.C.$, hence perpetual losses, and the profit motivated producer will go out of business.

2) If shadow prices are not centrally administered, profit maximizing behavior will destroy the "many-firms" requirement for competitive behavior and may leave one producer exploiting whatever tilt there may be in the demand curve for the product by producing at $M.C. = \text{Marginal Revenue}$, both less than price. Again, markets will not do.

3) A Lange-Lerner system of bureaucrats, profit-maximizing in the small with respect to centrally quoted prices but ignoring losses, may or may not produce the right outputs, depending on whether these outputs occur at rising or falling $M.C.$ -s. At any rate, if the injunction: "maximize profits (in the small)" works, it is only because the Minister knows the answers already or because he is lucky and by sheer coincidence all $M.C.$ -s are rising.

In the Chapter One discussion we assumed a production function exhibiting smoothly increasing returns to scale along its full indicated range. Its one input counterpart would be as in Fig. 2-d.

But since the conclusions depend only on local non-convexity in the immediate neighborhood of the relevant solution points, they are clearly of application where non-convexity of the production functions, while not negligible, is local and sequentially kinky, as in Figures 2-a, b & c. Yet the kinky type function, with its explicit indivisibilities, is of sufficient interest to warrant fuller exploration. In particular, it illustrates the importance of demand considerations in determining what indivisibilities are of quantitative significance.

Apples and Bridges¹

We posit:

- (a) One inelastically supplied, homogeneous and perfectly divisible input, labor services (L)
- (b) Two homogeneous and perfectly divisible goods, apples (A) and bridge-crossings (B)
- (c) A Samuelson type properly convex social indifference function which provides a ranking for the community of all conceivable output combinations. Such a function presumes that "incomes" are continuously redistributed so as to maximize, in utility space, over the Social Welfare Function implied by the political consensus.²
- (d) That the production function for apples is of the form $\dot{A} = cL_A$, with units so defined as to make $c = 1$; further, that the production function for bridge crossings, $\dot{B} = f(L_B)$, though of constant returns to scale "in the large",³ shows substantial indivisibilities, as shown in Fig. 3.

This is the simplest of all possible general-equilibrium set-ups, and has, like many polar cases, exaggerated properties, yet it will

¹ While in some respects "bridges" are misleading -- we want to avoid durability difficulties -- the least one can do to acknowledge Dupuit's extraordinary contribution of 1844 is to stick to his example.

² cf. P.A. Samuelson, "Social Indifference Curves", Q.J.E. February, 1956.

³ This is a crude notion, but its meaning is clear: lumps of bridges can be reproduced at constant cost.

serve to demonstrate the curious consequences of the gross indivisibility of a bridge. Further, its unduly angular assumptions can and will be qualified.

The first step is to derive the production possibility curve of the community, by means of a sort of poor man's Edgeworth box (Fig. 4). The length of the box represents \bar{L} , the left vertical axis B and the right vertical axis A. There is no need to have anything upside down: the south-west corner is origin for the B-function and the south-east for A. The functions are drawn as assumed: the A function is a 45° straight line. The various maximal "efficient" combinations of apples and crossings lie along any and all verticals: along any one such, L is uniquely divided between A and B, and the various verticals' intersections with the production functions denote how much A and B are producible. The consolidation of these combinations in output space gives the production possibility curve, F-F, as in Fig. 5.

Before exploring the implications of F-F it is useful, and in this context also possible, to set out the family of cost curves implied by the production functions. Cost curves are partial equilibrium entities and have to be evaluated at fixed prices, hence in general we could not infer them from the production functions alone. But in our simple model, all costs are real labor-costs pure and simple, hence a single-valued, technologically determined function of output. Since only relative prices matter, we can use apples as numeraire and set the price of apples, p_A , equal to one. The total cost curve (T.C.)

for B, then, as for A, consists in the mirror images of the respective production functions turned on their sides. (Figs. 6-a and 6-c: if p_A were other than unity and/or if labor hours did not convert into apples on a one to one basis, we should have to stretch these image-functions by a scale factor). It is clear that the T.C. curve for B bounds a non-convex set of feasible cost-output points.

The implied long run A.C. curves are as drawn in Figs. 6-b and 6-d. That for B shows alternatively falling and (vertically) rising segments. The corresponding M.C.B. curve is even "queerer". In ranges of output where the A.C. curve is falling, M.C. = 0. At the output points where the kink occurs -- where a lump is just exactly "exhausted" -- the M.C. curve becomes infinite.

What does this set-up imply by our various efficiency and market tests? First, we must establish the implications of profit maximizing behavior by the "producer" of bridge services, in response to various possible prices. In order to abstract from the monopoly problem -- notably that in a "free market" context no bridge operator would act as a perfect competitor and take prices as insensitive to his actions -- let's posit a system of government administered prices. A glance at Fig. 6-b suggests three possible cases. If the quoted price per crossing, p_B , is less than $(p_B)_0$, the bridge-man will lose money, no matter how he rations the number of crossings, unless he provides no crossing (i.e. no bridge) at all. Price is everywhere less than A.C.,

and no profits are possible. It follows, in our rational world of perfect information and certainty, that with $p_B < (p_B)_0$, no crossings will be provided, i.e. no bridge will be built, by profit seeking entrepreneurs.

If the quoted p_B just equals $(p_B)_0$, entrepreneurs will be perfectly indifferent between no B "production", 1000 crossings (one bridge), 2000 crossings (two bridges), etc. If, finally, p_B exceeds $(p_B)_0$, entrepreneurs will realize a positive profit margin on every crossing and will attempt, if they really take the price quotations seriously, to get as far "east" in Fig. 6-b as possible.

So far, our situation is akin to the classical Ricardian constant cost set-up where at $p = \overline{M.C.} = \overline{A.C.}$ the producer is wholly indifferent about what output to produce, while at $p < \overline{M.C.} = \overline{A.C.}$ profit maximizing output is zero, and at $p > \overline{M.C.} = \overline{A.C.}$ planned output will tend to "infinity".¹ (Remember, all this is in the long run: sunk costs are zero.) The reason for the similarity is that our B production function shows constant returns for scale expansion by "lumps". But while considerations of stability fully resolve the dilemma in context of Ricardo's world of all-around convexity,² our problem is not so easily disposed.

¹ This situation gave rise to the so-called London School controversy about the instability of the all-around constant cost competitive model. By explicitly considering the simple dynamics implicit in the stability of equilibrium the paradox is easily resolved, precisely because in that set-up all is nicely convex.

² Though there does remain a residual indeterminacy about how many firms will produce the determinate industry output. But the very indeterminacy indicates that the question is wholly inessential. (cf. pp. 78-80, in Samuelson, Foundations of Economic Analysis, Harvard, 1947).

To see the difficulty, it is best to turn to our general-equilibrium construct: the production possibility curve, F-F. It is reproduced, in Figs. 7-a, b, c and d, precisely as derived from the production functions (given total L). Recall that we assumed, as given, a family of conventionally convex social indifference contours of the Samuelson variety, which can be exploited as though a single mind were engaged in maximizing it. In fact, these contours are a determinate function of an implicit social welfare function of "regular" content and curvature, and of individual taste patterns of the usual ordinal variety. Further, they are, in concept, subject to empirical inference from observable price-market data.¹ Superimpose these contours on Fig. 7-a and select the one S_0 , which is just touched, but not crossed, by the production-possibility curve F-F. (This is the only contour actually plotted.) Its point of contact with F-F, ϕ , marks the B-A mix which maximizes welfare in terms of the W-function implicit in the S-function. The absolute slope at ϕ , in turn, denotes the marginal rate of substitution (MRS) of A for B, in terms of the community's "ethically-corrected" private valuations at that maximum-of-welfare point.

Now we know, from elementary value and welfare theory, that for individually preference-maximizing consumers to be in equilibrium relative prices must just match MRS-s. If prices are efficiently to ration the ϕ -values of A and B, then, $\frac{P_B}{P_A}$ must just equal $(MRS)_{S_0}$ at ϕ : $\frac{P_B}{P_A}$ must, in other words, equal the absolute slope of S_0 at ϕ ,

¹ Given the premise that the observed society does in fact redistribute income in lump sum fashion according to its implicit social welfare function. On the original development, see P.A. Samuelson, op. cit.

as does the absolute slope of the price-line marked P-P. P-P, then, defined by $|\tan a|$, depicts the $\frac{P_B}{P_A}$ ratio required for optimal exchange and distribution: the optimal exchange-price ratio. But, by assumption, $p_A = 1$. The exchange-required $(p_B)\phi$, therefore, is just equal to $|\tan a|$.

Will production of B at the level required by ϕ "pay", from the point of view of the private profit seeker, when price is set at $p_B = |\tan a|$? His total revenue (T.R.) from bridge tolls will be $(B_\phi)(p_B)\phi = (B_\phi) |\tan a|$. $|\tan a|$ in Fig. 7-a is given by the line segment $\overline{1}$ divided by the segment $\overline{2}$. $\overline{2}$, however, is just equal to B_ϕ . Therefore total bridge revenue at ϕ , as calculated at the correct exchange-price, is measured by the line segment $\overline{1}$. The question of profitability hinges on T.R. exceeding total cost (T.C.)

In our simple labor-theory-of-value world, where, furthermore, given our judicious choice of units, labor-hours are convertible into apples one to one irrespective of scale, $(T.C.B.)_\phi$ is immediately identifiable in Fig. 7-a as the line segment $\overline{3}$. This denotes the number of apples the community has to sacrifice to obtain the necessary labor for the construction of the bridge required by B_ϕ . With $p_A = 1$, $A = p_A A$; and with $A = L_A$ (via production) the wage rate in equilibrium must necessarily equal p_A . Ipsa, $A = L_A = wL_A$. It follows that the (opportunity) labor cost, hence the total cost, of producing B_ϕ , must

just equal the apple-cost of B_ϕ as denoted by $\overline{[3]}$.¹

$(T.R.B.)_\phi$ is indicated by $\overline{[1]}$; $(T.C.B.)_\phi$ by $\overline{[3]}$. The provision of bridge services, then, at the correct exchange-price ratio, is a highly lucrative operation. Can we, therefore, count on decentralized profit-motivated production decisions to sustain ϕ , the best of all feasible configurations? Careful examination of the relevant line segments in Fig. 7-a suggests the answer. While at $p_B = (p_B)_\phi$, ϕ is a point of positive profit, it is not the point of maximum profit.² This last is at the point v , where $F-F$ just touches the highest price line (P^*) .³ If bridge entrepreneurs have a working knowledge of their cost curves, they will rush toward v , despite the fact that "locally" ϕ is a relative profit-maximum position.⁴

This is as it should be. Our previous analysis in terms of the (A.C.B.) curve showed that at any profit-yielding p_B profit maximizing would call for more and more bridges. This conclusion is now confirmed.

¹ All this comes out so nicely because with one scarcity, labor, and with the labor content of apples and bridges, though not of bridge crossings, technologically fixed, clearly only one relative constellation of $\overline{[wage-(T.C.A.)-(Bridge-T.C.)]}$ is possible. (This is, incidentally, a basic characteristic of the Leontief system.)

² We did not have to worry about the relative profitability of various outputs in the apple industry. All outputs are equally "nil-profitable".

³ In our simple case, and largely because of convexity "in the lump", the point of maximum bridge profits corresponds, as it always does in a really convex system, with the $(p_B)_\phi$ -valued maximum value of "national" output.

⁴ If, that is, they really do take $(p_B)_\phi$ seriously as the relevant demand curve.

$(p_B)_\emptyset = |\tan a|$ must, by the existence of positive profits, exceed the knife-edge price $(p_B)_0$ of Fig. 6-b, hence the Min. A.C.-s of any number of bridge crossings. Therefore at $(p_B)_\emptyset$, profits will rise with crossings.

To argue that, in fact, $(p_B)_\emptyset$ would not be maintained if three bridges were actually built and no apples produced -- MRS at v calls for a very low if not zero p_B -- hence that market processes will correct mistakes, is to miss the central point. Whatever the built-in dynamics, there does not exist a single price for bridge crossings that will at the same time keep both consumers and producers from rushing away from \emptyset , where they both ought to be. This is, here, the essential consequence of non-convexity, and this is what differentiates this case from the Ricardo-London School situation.

What is the one p_B which would sustain production at B_\emptyset ? From our previous A.C.-curve analysis we know that at $p_B = (p_B)_0$ profits, while precisely zero, would be nevertheless at their attainable maximum, hence profit considerations would at least not drive bridge operators away from B_\emptyset . The zero-profit price $(p_B)_0$ of Fig. 6-b corresponds in Fig. 7-b to $p_B = |\tan a'|$, denoted by $P'-P'$. For at $p_B = |\tan a'|$, $(T.C.B.)_\emptyset = (T.R.B.)_\emptyset$; profit, while zero, is at its maximum and there is no profit urge to move away from \emptyset .¹

¹ It is a different question whether $p_B = |\tan a'|$ would ever lead producers to B_\emptyset .

But this is of no help. At ϕ , in Fig. 7-a, MRS along the social indifference curve, S_0 , is equal not to $p_B = |\tan a'|$, but rather to $|\tan a| > |\tan a'|$.¹ Hence the price which will sustain production at ϕ will, in turn, drive consumers away. The dilemma is symmetric. If, and only if, it should happen that by sheer coincidence the S-contour through ϕ has curvature as does S_1 in Fig. 7-b, will one and the same p_B efficiently mediate both production and exchange.

A quick glance at Fig. 7-c suggests the consequences of an even flatter S-slope at ϕ . The implied exchange-price is then $|\tan a''|$ and denoted by the price line $P''-P''$. In terms of the Fig. 6-b A.C.-curve formulation this exchange price is less than $(p_B)_0$, hence less than A.C. at any positive B. It follows, therefore, as can readily be seen in 7-c, that the profit maximum position is at $B = 0$. Analysis by the vertical-intercept-determined line segments shows $T.C. > T.R.$ for any $B > 0$. This case then is somewhat akin to the smooth increasing returns to scale situation where $A.C. > M.C.$ for any level of output, hence with $M.C. = p$ any positive output implies losses. It differs in that where smooth increasing returns to scale obtain throughout the feasible output range,² no price can render positive output profitable as long as output is adjusted to equate M.C. with price.

A last case warrants a few remarks. What if S were as in Fig. 7-d? This implies that with one bridge built satiation occurs short

¹ Not to confuse Fig. 7-a, angle a' is not there indicated. It is formed by the B-axis and a straight line connecting ϕ with v (cf. a' in Fig. 7-b).

² This range being determined by general-equilibrium derivation of the production possibility frontier.

of "capacity" use of the bridge. The correct exchange price of crossings is zero. Yet it is clear that one bridge should be built. Points of zero bridge crossings lie on a lower social indifference curve. This is in a sense the most "realistic" case and most sharply points up the fact that if prices are to be used efficiently to ration scarcities, they are not suitable, except if all is convex, to raising revenue to cover cost. It is of interest to note, incidentally, that in our angular "right-angled" model, the maximum welfare position, bar satiation, will always be at one of the full-capacity "vertex" points. If satiation does occur, as in Fig. 7-d, the actual number of crossings will, within a range, be inessentially indeterminate.¹

Some Modifications: A-goods and B-goods

A number of variants of the above model warrant brief comment:

1) It is of interest to note what happens, if, while maintaining the sharp right-angled step function quality of the production function for B, it were made to exhibit increasing or, especially, decreasing returns to scale to successive "lumps" of input (instead of constant returns to "lumps" as above). It is clear, for instance, that the "smaller" the lumps, and the greater the decreasing returns effect, the better an approximation is the smooth concave-to-the-origin curve of neo-classical theory. Such "smoothing" and "bending" (in the right direction) substantially mitigates, as one would expect, the dilemmas

¹ If there are "subjective" costs, e.g. tired legs, associated with crossings beyond the number desired, S will tend to turn north-east and we get determinacy at a point of "excess" capacity.

associated with the use of prices for decentralized "computation". On the other hand, bending the other way compounds troubles. With smoothing, it produces a curve that approximates the concave-outward non-convex $F-F$ of Part V, Chapter One.

2) What if we drop the assumption of rigid and explicit indivisibility and explore the consequences, rather, of a production function such as that for B in Fig. 8-a. Maintaining all our other assumptions about apples, etc., the implied production possibility curve is as $F'-F'$ in 8-b, and the cost curves as in 8-c and 8-d. With normally convex social indifference curves there are four possible cases:

(a) If the highest S touches as S_1 , the required maximum of welfare point, x_1 , evaluated at the exchange-efficient p_B , will be a point of (i) locally minimum total value of product; (ii) minimum as well as negative B-profits (M.C.B. necessarily cuts the horizontal p_B from above); hence (iii) a non-sustainable point irrespective of whether "profit seekers" are in earnest, or Lange-Lerner civil servants. Moreover, the fact of tangency does not guarantee x_1 to be a proper local maximum. If S_1 were much flatter it would intersect $F'-F'$ again at a point of less B and more A, thus suggesting an even better point closer to the A-axis. "Vision in the large" is required to find the best point.

(b) If, in turn, the best S -contour touches as S_2 at x_2 ,¹ the

¹ Note that S_2 belongs to a different family of S -contours!

optimum point implies a locally profit-maximum level of B (M.C. cuts p_B from below), but this maximum level is still negative ($p_B = M.C. < A.C.$). Businessmen go out of business, but Lange-Lerner managers, following instruction, will do the right thing. But note: the central authority must know ahead of time that x_2 lies in a range of rising M.C.: it must know the solution, otherwise it could not without grave risk issue the "maximize profits" injunction. And if it has to know the solution to start off with, the whole Lange-Lerner scheme loses much of its point.¹ ($p_B^B + p_A^A$) at x_2 , incidentally, is at a local maximum, but not at its maximum. The associated price line cuts $F'-F'$ at a point close to the A-axis. Further, the fact of a maximal tangency again does not assure maximum-of-W. A flat S_2 contour could again cause trouble.

(c) If S is as at x_3 and if we know it, almost all is well. Profits are positive and at a maximum, and the required B point can be sustained by (perfect) markets. Moreover, the value of output, NNP as it were, is also at a maximum (at the relevant prices). But can we be sure from the tangency and the existence of competitive profits alone that at x_3 the world is doing as best as it can? Again, the answer is no. As drawn, x_3 is in fact the highest contour touched by $F'-F'$. But it is obvious that with slight modification of the intersection of $F'-F'$ with the A-axis, x_3 could again become inferior to output points of much more A and much less B.

(d) Finally, x_4 represents the "corner tangency" possibility that was discussed in Chapter One. Its analysis contains nothing fundamentally

¹ On all this see Part V, Chapter One.

different from that of the first three cases.¹

3) It is important to stress some special attributes of the model which could trap the unwary into false generalization. In our case, convexity of $F-F$, in output space, is evidently sufficient to guarantee proper convexity in input-output space as well. This is due to the one scarce input assumption, which assures that none of the complications due to input substitutions, etc. can arise. As Part V of Chapter One demonstrates for the two-by-two case, in a world of two or more inputs a convex $F-F$ is entirely consistent with increasing returns to scale in the production of every good. Hence a system could give a nice neo-classical production-possibility curve yet have all its Pareto-efficient production points show all-negative or even minimum producers' profits. In general, it is not possible to infer anything about the slopes of the (partial-equilibrium) cost curves from the slope of the (general equilibrium) production possibility schedule. Movements along the latter imply shifts of the former.

4) But in our case, at least, all the difficulties are contained in the non-convexity of $F-F$. If $F-F$ is convex, as in Fig. 9, we know by the definition of convexity that a point of internal tangency between it and a convex set of S -contours must either be the maximum-in-the-large, or, at the least, a maximum-in-the-large. By the definition of tangency, the shared slope defines a (price) line. If both the S -set and the F -set are convex, then by definition this line does not touch either

¹ The student of the conventional theory of the firm will have noted the resemblance of our cost curves to the usual "U-shaped A.C." set of curves treated in that context. Only here, the jump to general equilibrium is immediate and obvious, because of our assumption about apple production.

set outside of the one point or region of tangency. The feasible points all lie "south-west" , while the points of higher welfare are all "north-east". The conclusion holds equally for corner tangencies.

Institutional Implications

In all the above attention was directed at whether, with indivisibility in the technology, the maximum-of-welfare output configuration would or would not coincide with points of maximum and/or non-negative producers' profits. Such correspondence is necessary if price guided, decentralized and profit-motivated production decisions are to sustain the optimal output configuration. But it is surely not sufficient, unless we assume a Lange-Lerner type set-up of centrally quoted shadow prices which have parametric fixity to each producer. If prices are to be determined by market forces, their equilibrium values will not correspond to the maximum-of-welfare solution values -- even abstracting from all dynamical adjustment problems and retaining a steady-state statical frame of reference -- unless self-policing perfect competition obtains in all markets.¹

Self-policing competition requires "very many" producers in every market.² We know that no profit-seeking and competitive firm facing a

¹ And, of course, unless incomes are redistributed as required by the specific welfare function in hand. Assume, in what follows, that income distribution is being correctly managed.

² Or at least the potentiality of very many producers, ready and able to "enter the fray" instantaneously. This is sufficient in the London School constant-cost case, where the equilibrium number of firms in the industry is indeterminate. See, Part IV, Chapter One.

fixed output price will produce in a stage of increasing returns to scale, i.e. in a stage of declining A.C., since this would imply $p = M.C. < A.C.$ hence losses. If, then, indivisibilities or smooth non-linearities give increasing returns to scale over an output range that constitutes a significant fraction of the total market, a few firms of "efficient" scale will exhaust that market and the institutional conditions for perfect competition do not exist.

Our Figure 8 set-up illustrates the limiting case of this. Assume the set S_3 to dominate: x_3 is the feasible bliss point. Recall that at x_3 the shadow prices implied by the welfare-maximum permit profitable production of B; in fact at these prices B is a point of maximum producers' profit -- for a single producer. One efficient producer with $M.C. \geq A.C.$ exhausts the market.

Now drop the assumption of centrally administered prices. The producer of B is confronted by a market demand curve. If this has any tilt in it -- if it is even slightly sensitive to price -- and the curvature of S_3 suggests that it is -- we know that the maximum profit level of output lies at M.C. equal to marginal revenue (MR) and that this gives less B, and at a higher price, than x_3 .

By directing attention to an administered price set-up, our discussion managed to skirt all the problems which concern the relations of an industry to its component firms. Yet these are clearly of importance: they determine in large part whether competition can be used efficiently to regulate production. As has been emphasized ever

since Marshall, in situations where the technology is characterized by significant non-convexities it can not. To be precise, in such situations the full welfare-maximum solution of inputs, outputs and prices will not be sustained by a market solution.

That profitability is neither a necessary nor a sufficient test of efficiency, even in a frictionless and stationary world of perfect certainty, is a very old conclusion.¹ It is on occasion suggested, however, that it might be efficient to make partial use of a market profitability test in situations where an output decision can be split more or less meaningfully between a lumpy all-or-none type "investment" decision (e.g. building a bridge) and a further decision about its appropriate scale and intensity of operation. The presumption is that the prospect of monopoly profit is necessary and sufficient to assure that the building of a bridge is socially desirable, though it may be that once built it ought to operate at a scale and price that imply a loss.

Quite apart from the feasibility or acceptability of the institutional connotation -- let monopolists do the building and then control their price and sales policy -- which is, for our purposes irrelevant, and apart, also, from questions of institutional strategy, this view is incorrect. It is easy to prove, and has often been, that simple monopoly profit is neither necessary nor sufficient. It is, for

¹ This statement cannot be taken to pre-judge the question of whether better or more efficient criteria are or are not feasible etc., nor the larger question of institutions.

instance, the principal implication of the famous Dupuit bridge-problem¹ that "the" bridge may be socially desirable -- i.e. required by the solution of a Paretian welfare-maximum problem -- even though it is impossible to cover its costs by means of non-discriminatory, single-price, rationing of crossings. In old-fashioned terms, no single price can capture all the consumer's surpluses² associated with an output. It takes a perfectly discriminating monopolist, meticulously exploiting the tilt in each separate person's demand curve, to do so. On the other hand, a monopolist may earn positive profits on a lump of activity which should not be. His market calculations do not take into account possible losses of "producer's surplus" by factors which exhibit some degree of specificity to a competing activity.

On Optimum Conditions Once Again

What then can be said, on the present level of generality, about the conditions which define the optimum? In a sense, it is the central moral of all the above that when the constraining set of feasible points over which we maximize -- in the above the set bounded by our (F-F)-s -- is non-convex, then marginal, local, im kleinen type conditions can never suffice.³

¹ A translation of the original 1844 article is to be found in International Economic Papers, Vol. 2.

² However weighted for summing over many people.

³ Recall that it is the convexity of F-F that is of import in this connection and that with many inputs F-F could be convex despite some non-convexity in production functions.

Even if the mistake of accepting a tangency as prima facie evidence of a maximum is avoided -- even if the proper second order inequalities which define a true feasible maximum, whether internal or at a corner, are made explicit -- non-convexity makes it mandatory that total conditions be examined. Choice may be required between several local maxima, and such choice can only be made by comparing their total desirabilities. Further, the very non-convexities which render such total conditions of interest, cause a divergence between the private total conditions by which producers in an idealized market would calculate: total revenue minus total cost, and the social total conditions which are implied by a family of social indifference curves.

More systematically, where technology is non-convex and under competitive conditions -- or, via administered prices, "as if" competitive conditions -- the maximum-of-welfare could occur where:

- (1) Private and social marginal conditions and total conditions both coincide.¹ This will happen in the fully convex case and could happen, as at x_3 in Fig. 8-b, in regions of local convexity, even if the whole feasible set is non-convex.
- (2) The marginal conditions coincide but the total conditions do not. This is the case, for instance, at x_2 in Fig. 8-b, where profits at the maximum-of-welfare

¹ Substitute "decentralized producers" for "private" for a Lange-Lerner set-up.

B-point are at a local maximum, but negative. Private total calculations (in the B industry) underestimate the net social benefit of bridge services and B-output will be less than desirable. As indicated above, it is also possible that while the maximum-welfare point gives positive and locally maximum profits, producers with vision a-far rush to points of greater output (of B) because of the greater profitability (at the solution prices) of such a point of lower welfare. It is possible, in other words, that at the correct shadow prices competitive producers over-invest. (Note that this is not a dynamical cobweb phenomenon due to mis-information about the equilibrium price. Rather, it is due to a non-coincidence of the equilibrium exchange price and the equilibrium producers price).

(3) Neither the marginal nor the total conditions coincide. This is the case at x_1 : production of B should occur at a point of minimum and negative profits (M.C. cuts p from above). The profit maximum level of output could be greater or less than optimal.

We can rule out, virtually by definition, the fourth combination: coinciding total conditions but divergence "in the small". The last implies a profit-minimum in the small, hence M.C. cutting p from above.

This requires a falling M.C. hence, at least with a "normally" shaped T.C., $A.C. > M.C. = p$, ipso losses. But it is worth noting that where prices are set by markets and monopoly is permitted to develop, something like this combination could, in a very rough sense, occur. Investment in e.g. one "lump" may be both socially desirable and profitable at a monopolist's price and an output that is locally different from the maximum welfare point.

We shall find, incidentally, that the importance, where indivisibilities obtain, of total conditions, implies an analytical tie between the indivisibility aspects of social overhead phenomena and certain externality effects. The link is provided by the "public good" notion of public expenditure theory. But more on this in Part III. Here we had better turn first to a brief discussion of why it has not been wholly illegitimate to ignore indivisibility-through-time.

Durability

In discussing the statical implications of indivisibility-in-scale, we have entirely ignored the patent fact that the services of e.g. a bridge are also indivisible-through-time: they are not "exhaustible" instantaneously. The principal justification for such a procedure has already been suggested. The essential analytical consequence of lumpiness-in-scale is non-convexity: and most of the above concerned itself with the statical analysis of

such non-convexity. The interesting qualities of indivisibility-in-time are not amenable to the techniques of statics, and, more important, durability, unlike lumpiness in scale, does not in itself violate convexity.

This last is easy to see. Imagine that output, \dot{Q} , depends on, and only on, the input of a particular kind of capital stock: square feet of cleared land (A). It takes investment to clear land, but assume that once cleared it remains clear and suitable for cultivation for a fixed period n which is independent of the size of A . If output is only sensitive to land area there is no a priori reason whatever why $\dot{Q} = f(A)$ should show increasing returns to scale, whatever n may be. A could have infinite "life", yet the function could perfectly well be homogeneous and linear or give decreasing returns to scale.

What of the total cost curve? There is nothing about durability, per se, which is inconsistent with the existence, embedded in the maximum problem, of a parametric rental rate per square foot of cleared land -- it would reflect an interest charge and a depreciation charge. Such a solution-value rental rate, when multiplied into the A -s required for various \dot{Q} -s, will give T.C., hence the T.C. curve will have the precise curvature properties of the production function itself.

Why did we choose an input like A ? Because, while durable, its relevant dimension (area) is subject to smoothly continuous variation: it does not exhibit lumpiness-in-scale. It is true that many kinds of durable inputs, such as machines, do in fact give rise to non-convexity.

This, however, is not because of their durability, but because it happens that such durability is often associated with scale-indivisibility.

It is easy to become confused on this issue. After all, are not depreciation and interest charges fixed costs, hence do they not imply a declining A.C.? Once a producer has in fact signed a long-term lease, or himself invested in, e.g. A, it is of course true that his short-run unit cost curve will rise for outputs lower than what A can efficiently produce. But a falling short-run average total cost curve implies nothing about the long-run production function or cost curves. In a sense a fixed cost in any short-run gives rise to a falling A.T.C. just because it represents a scale-indivisibility -- due not to the technology but to contractual or equity commitment. But by the very fact that the fixed component reflects a "sunk" cost, it is irrelevant to the producer's short-run decisions, hence it does not give rise to non-convexity type difficulties.

It should be said again, however, that as a matter of fact, durability is often linked with scale-indivisibility. It is likely, for example, that a machine tool with a larger per period capacity will also be more durable.¹ Where so, non-convexity might occur even though it appears that maximum service flow capacity is subject to continuous variation. The time-scale distinction then rests on the fine point, that such a situation is due to scale-indivisibility of the more durable "big" machine itself, and further, that it is illegitimate to assign a

¹ This has to do with empirical fact. It is not a logically necessary relationship.

single axis to an input that is not homogeneous. At any rate the conclusion stands: "pure" durability does not give rise to non-convexity.

Yet it is only too clear, from our own examples, that ignoring durability gives rise to deep incongruities. Of what meaning is an F-F such as in Fig. 8-b? It is supposed to depict the alternative maximal combinations of bridge services and apples available to a society per unit time. But while there is certainly a lumpy cost in apples foregone during the period during which the bridge is being built, what about after it is built? Even in the simplest case¹ there will be a different F-F associated with (a) the period during which the bridge is being constructed and is not available for use; (b) a transitional period of perhaps limited use (c) a post-construction period. Much of our analysis above, as the F-F of Fig. 5, characterizes the middle type period, or at least a long enough period to include both construction and use. We could, of course, have selected examples which exhibit lumpiness without durability.² But when durability is involved, F-F's undergo irreversible shifts and undulations whose formal exploration demands dynamical modes of analysis. It happens that such analysis is entirely feasible for systems without lumpiness. With lumpiness, matters are worse in a dynamical situation even than in our statical set-up. The analytical machinery fails. But more on this in Part IV. Here we must back to Social Overhead.

¹ A continuous and constant input stream giving rise, at its termination, and without further inputs, to a constant output stream.

² e.g. haircuts (compliments of Prof. Solow).

D. Indivisibility in Scale and Social Overheads

We have explored, tentatively but at length, the statical consequences for Pareto-efficiency and market performance of indivisibility, of increasing returns to scale. Justification was based on the presumption that most of the so-called "social overhead" activities do in fact exhibit considerable indivisibility in scale. But so do other activities, which would not likely appear on many economists' lists: most of "heavy industry", for example, consists in big lumps. What then are the qualities which render lumpiness especially significant in e.g. power, transport, communications, etc.? In a sense, in a very low income community, most manufacturing activities are much akin to social overhead. They may be desirable, yet because of increasing returns to scale, they make, or would make, losses. But there do exist qualities which tend to differentiate social overheads from other initially lumpy activities.

Many of the most important of these qualities are tied up with durability and involve dynamical considerations. But a few fit even a more or less statical context. The following comments on these last are in the nature of casual conjectures:

1) Most of the activities that fall easily into the social overhead category appear to exhibit not only a sizeable initial "lump" but, further, low variable costs. Once beyond the "lump" the production function becomes very steep in the output-input(s) plane, hence T.C.-s

are very flat: M.C.-s very small. This causes A.C.-s to fall over a greater range than if the additional inputs required by additional output (from the initial "lump") were large. Examples are legion; the extra load-induced upkeep or operating cost on roads, bridges, telephone poles, switch boards, dams, water-reservoirs, etc. is probably small as a fraction of average total cost compared to that in steel, or machinery. This means that as compared to the latter, A.C. in social overhead activities is likely to decline over a greater range, and moreover, that a non-convex F-F is more likely.

2) But what does it mean "over a greater range"? This clearly implies some presumptions about demand, "size of the market" etc.? What are these presumptions? In the roughest and sketchiest of terms, it could be argued that:

Almost any feasible and Pareto-efficient output mix in a relatively low-income country¹ which, however, is moving beyond the wholly stationary subsistence stage, is likely to imply some price inelastic (in a range) derived demand for the services of certain kinds of social overhead. Almost any degree of commercialization of agriculture requires some transport and storage facilities -- even if only dirt roads. The same is true for at least some kinds of extractive activity. And, perhaps most important, even very small scale factory manufacturing requires some power, transport, etc. It is suggested that the derived demand curves for such social overhead services have the general

¹ And it is for such countries that the social overhead notion has especial significance, different from that of public utilities in wealthy countries.

character of d-d in Fig. 10.

The insensitivity to price increases (up to a cut-off point) would be due to a compounding of (i) strong technological limits to substitution of other inputs, e.g. of labor for electricity or even for a bridge; (ii) the relatively small share of a purchaser's T.C. represented by expenditures on social overhead service inputs (low wage rates do not imply low wage bill); more important, (iii) the virtually physical impossibility ("infinite" transport cost) of importing the social overhead service (how do you import the service of a bridge or a road between two points?)¹

On the other hand satiation effects -- how many times will one cross the bridge just to cross it -- compound the consequences of (i) and (ii) above, to make for inelasticity in the face of price declines. Lower price is not likely to induce enough demand pull to drag social overhead activities into a range of rising A.C. or M.C.²

¹ Components, of course, are importable. But much of the material that goes into dams, roads, etc. has itself high transport cost and much is unskilled labor that has to be "congealed". Incidentally, it may be that even the direct final consumption demand for social overhead services and for the products of social overhead users, tends to be price-inelastic. Precisely because they involve a very small fraction of a household's budget the income effects of price changes are negligible. But substitution effects could make this cut the other way.

² Use of the price elasticity notion in the context of a particular Pareto-efficient solution (which fixes all prices) is perfectly legitimate. The solution fixes demand and supply curves in the neighborhood of the solution and inelasticity suggests that the conclusions are not likely to change for small departures to neighboring efficient points. A better formulation would perhaps have rested on the assertion that finite but small demand for social overhead services characterizes most of the feasible efficient points.

In some respects the most interesting aspects of the "must have but cannot use fully" quality of social overheads relate to non-stationary situations and income-elasticities (Engel's curves). More is involved than simply low income effects: spatial considerations may be crucial. The services of a bridge, or a road, or a drain, are highly space-bound. Even if income expansion brings about more river-crossings (for direct consumption or as an input), unless the expansion occurs via increasing density within the "service area" of a bridge, the effect is as likely to be a demand for a second bridge as for an increased use of the first. And if so, the income effect for the services of a particular bridge could well be negative.¹ Ditto, for e.g. policemen. A lot about why social overheads are not pulled out of the stage of declining A.C. by rising incomes is explained by such localization (infinite transport cost) phenomena. To the degree that income pull with regard to social overhead services often works on the qualitative rather than quantitative margin, such spatial effects are further reinforced. But we must now turn to external economies.

¹ I.e., the services of a particular bridge could well be "inferior" in nature, and since at unchanging zero price (= M.C.) nothing will induce substitution effects, the income effect dominates to give a Giffen-situation.

III EXTERNAL ECONOMIES AND SOCIAL OVERHEAD

It is a frequent assertion in the economic development literature that certain kinds of activity: power, transport, etc., are in the nature of "social overhead", because they confer "external" benefits which do not enter private market calculations. The implication is that social overhead activities require government involvement because of associated external economies. The following consists in an exploration of the nature and content of this supposed association. Part A is in the nature of a digression on the development and present state of the external economy doctrine in context of neo-classical equilibrium theory. It is concluded that the usual emphasis on "divorce of scarcity from ownership" is misplaced and cloaks some more fundamental issues. Part B suggests an ordering that brings these deeper issues into the foreground. The organizing principle is contained in the hierarchy of "existence" and "efficiency" tests elaborated in Part I. Part C then identifies, in terms of the above ordering, what significant association may exist between social overheads and externalities. We are brought back, full circle, to indivisibility and increasing returns to scale -- though not without some added baggage in the form of "public goods".

A. A Digression on External Economies

By Way of Some History

Marshall, as has often been pointed out, proposed the external economy argument to explain, without invoking dynamical considerations,

the phenomenon of a negatively sloped ("forward falling") long-run industry supply curve in terms consistent with constant or rising M.C. in the "representative" firm. The device permits -- on the level of logic if not that of fact -- long-run competitive equilibrium of many firms within an industry, each producing at its profit maximum $p = \text{M.C.}$ position, without foreclosing the possibility of a falling supply price with rising industry output.¹

The mechanism is simple. It is postulated that there exist certain exploitable economies that come into play in response to an expansion in the output of the industry as a whole, economies whose effect is to shift down the cost curves of all the component firms. These economies, however, are not subject to exploitation by any one of the myriad of tiny atomized firms. Their own M.C. curves, at $p = \text{M.C.}$, rise both before and after the shift, due, presumably, to internal diseconomies associated with the entrepreneurial function which defines the firm. Even the modern formulation is not entirely without ambiguity -- institutional ambiguity is intrinsic to the device of parametrization: how many firms does it take for the demand curve of each to be perfectly horizontal? -- but it does provide a means for "saving" the competitive model, of ducking the monopoly problem.

¹ Remember, this refers to a so-called Marshallian supply curve. It has nothing to do with the Walrasian "maximum quantity supplied at given price" type schedule.

Marshall and then Pigou "preferred" as it were, the other horn of what they perhaps saw as a dilemma. The external economy device, while saving competition, implies a flaw in the efficacy of the "invisible hand" in guiding production.¹ "Price equal to M.C." is saved, but wrong. Market forces, they argued, will not give enough output by industries enjoying external economies and will cause industries with rising supply curves to over-expand. Hence the Marshall-Pigou prescription: to harmonize private production decisions with public welfare, tax the latter set of industries and subsidize the former.

It took the better part of thirty years, and the cumulative powers of Young, Robertson, Knight, Sraffa, and Viner,² to unravel the threads of truth and error which run through the Marshall-Pigou argument. The crucial distinction that provides the key to it all, is between what Viner labeled technological external economies, on the one hand, and pecuniary external economies on the other. The latter cause the long-run supply curve of an industry, say, A, to decline because the price of an input falls in response to an increase in A-s demand for it. The technological variety, on the other hand,

¹ Difficulties with income distribution were already by that time explicitly asserted.

² The strategic articles, with the exception of Young's ("Pigou's Wealth and Welfare", Q.J.E. XXVII, 1913, pp. 672-86), have all been reprinted in A.E.A. Readings in Price Theory, ed. Stigler & Boulding, Irwin, 1952.

has nothing to do with input prices. It involves certain organizational or other improvements in efficiency which are a function of increased industry output.

As regards pecuniary external economies, Robertson and Sraffa made it clear that in a sense both the Marshall-Pigou conclusions were wrong. For one thing, no subsidy is called for. The implied gains in efficiency are adequately signaled by the input price and profit maximizing output levels by the A-firms are socially efficient. Second, monopoly troubles may be with us, via, as it were, the back door. For what causes the price of B to drop in response to increased demand? We are back where we started: a declining long-run supply curve.

In the end, then, if internal technological economies of scale are ruled out, we are left only with technological externalities. All pecuniary external economies must be due to such technological economies somewhere in the system.¹ And it is true -- and this is what remains of the original Marshall-Pigou position -- that technological externalities are not correctly accounted for by prices, that they violate the efficiency of the price system as a computing device.

¹ Recall that only narrowly statical, reversible phenomena are admissible here. Needless to say, relevance to a world of disequilibrium dynamics is limited.

The Modern Formulation

In modern terms, the notion of external economies -- external economies proper that is: Viner's technological variety -- belongs to a more general doctrine of "direct interaction". Such interaction can be classified, as was done in Part IV - 7 of Chapter One, according to whether it involves producer-producer, consumer-consumer, producer-consumer, or employer-employee relations. It consists in interdependences that are external to the price system and hence unaccounted for by market valuations. Analytically, it implies the non-independence of various preference and production functions. The effect is to cause divergence between private and social cost-benefit calculations.

That this is so, is easily demonstrated by means of a simple production model.¹ Take as given:

- (a) One homogeneous and inelastically supplied input: labor services (L);
- (b) Two homogeneous and divisible goods, apples (A) and honey (H), produced by labor at non-increasing returns to scale. But while the output of A is dependent on L_A alone, we assume that honey production is sensitive to the level of apple output. (Prof. Meade makes pleasurable the thought of apple blossoms making for honey abundance.)²

¹ The specific example is Meade's (Economic Journal, LXII (1952), pp. 52-67). But his discussion is aimed at issues of income imputation: factors not getting the value of their marginal social product. He does not explicitly demonstrate that competitive producers will not give a Pareto-efficient output configuration.

² Not to be caught out, I had better make explicit the following further assumption: apple blossoms (or the nectar therein) are presumed to be exhaustible, rationable, private goods. More nectar to one bee means less to another. On the need for this assumption, see Part B below.

The two production functions, then, are as follows:

$$A = A(L_A) \quad (1)$$

$$H = H(L_H; A(L_A)) \quad (2)$$

By our labor supply restriction

$$\bar{L} \geq L_A + L_H \quad (3)$$

If we make the usual non-satiation, non-redundancy assumptions (in a one-input world this is hardly painful) we can drop the inequality.

The problem of Pareto-efficient production consists in determining the various maximal combinations of A and H: in deriving the production possibility curve. To do this, we maximize, subject to the production functions and the labor restraint, an arbitrarily weighted sum of A and H: $p_A A + p_H H$, where the "value" weights can be varied at will.¹

Following the usual Lagrangean procedure, we set up the expression

$$\Phi = p_A A(L_A) + p_H H(L_H; A(L_A)) + w(\bar{L} - L_A - L_H), \quad (4)$$

differentiate it with respect to the variables L_A and L_H ,

treating the weights: p_A , p_H , and the additional Lagrangean multiplier w , as constant, then set the resulting partial derivatives equal to zero.

¹ We could complete the system by specifying households with tastes, initial endowments, etc. (or, in place of the latter, a Welfare Function); or, we can simply imagine a small country of honey and apple producers facing perfectly elastic international demand curves.

$$\frac{\partial \Phi}{\partial L_H} = P_H \frac{\partial H}{\partial L_H} - w = 0 \quad (5)$$

$$\frac{\partial \Phi}{\partial L_A} = P_A \frac{dA}{dL_A} + P_H \frac{\partial H}{\partial A} \cdot \frac{dA}{dL_A} - w = 0 \quad (6)$$

For Pareto efficient production, then,

$$P_H \frac{\partial H}{\partial L_H} = w \quad (5')$$

$$P_A \frac{dA}{dL_A} + P_H \cdot \frac{\partial H}{\partial A} \cdot \frac{dA}{dL_A} = w \quad (6')$$

Equations (5') and (6'), together with the two production functions (1) and (2), and the labor balance equation (3), give five relationships in the seven unknowns, A , H , L_A , L_H , P_A , P_H , and w . Since only relative prices matter,¹ honey can be designated as numeraire and its price fixed at will. We are left with five equations in six unknowns and this is as it should be. Without explicit taste or final demand considerations it would be odd indeed if the system gave a determinate solution.

For our present purposes, however, it is useful to fix p_A also -- we can think of our community as embedded in a large world of competitive trade in apples. The equations will then give out a full Pareto efficient solution for all the variables, including w . But it can now be seen that profit-maximizing honey and apple producers,

¹ See Part III of Chapter One.

responding atomistically to the solution values of p_H , p_A and w , quoted, say, by a production board, would not sustain that solution.

Equation (5') is familiar enough and consistent with profit maximizing behavior. Honey producers will do for profit what they must do for efficiency: hire labor until the value of its marginal product just equals the wage rate.¹ But (6') gives trouble. Profit maximizing apple producers will not produce at the socially required level.

For maximum profit, apple producers will hire labor up to where

$$p_A \cdot \frac{dA}{dL_A} = w \quad (7)^2$$

But (7) is inconsistent with (6') unless $\frac{\partial H}{\partial A} = 0$. If apples have any cross effect on honey output -- e.g., if with other inputs held constant more apple blossoms make for more honey -- then the Pareto condition is not satisfied by atomistic profit maximizers.

¹ We assume internal tangencies and convexity throughout. The last is implicit in constant-returns to L: the A effect on H will reinforce convexity or tend to offset non-convexity. But more on this below.

² M.C. $\equiv \frac{d(T.C.)}{dA} \equiv \frac{d(wL_A)}{dA} = \frac{w}{1} \frac{dL_A}{dA} = p_A$. Therefore, if

we assume a smooth single-valued function, $p_A \cdot \frac{dA}{dL_A} = w$.

The direction and extent of the deviation is easy to see.

Rewrite (6'):

$$\frac{dA}{dL_A} = \frac{w}{P_A + P_H \frac{\partial H}{\partial A}} \quad (6'')$$

Comparing it with (7), it is clear that

$$\frac{\partial H}{\partial A} \leq 0 \Rightarrow \left(\frac{dA}{dL_A} \right)_{\text{Private}} \geq \left(\frac{dA}{dL_A} \right)_{\text{Social}} \quad (8)$$

If apples have a positive external effect on honey output, market determined L_A will be less than socially desirable. A-producers will stop hiring labor short of where the value of its marginal social product: the value of the extra A plus the value of the extra apple-induced H, matches its social cost as measured by w . The converse would be true if A were assumed to exert an external diseconomy in the production of H. Only in the case of neutrality: no external interaction, will the private criteria satisfy the social.

A different way to see this is to examine the relations of private to social marginal cost. The marginal (private) cost of apples to the apple producer is given by: $\frac{w}{dA/dL_A}$; that of the beekeeper

by $\frac{w}{\partial H / \partial L_H}$. The private marginal cost ratio, then, is

$$\frac{\frac{w}{dA/dL_A}}{\frac{w}{\partial H / \partial L_H}} = \frac{\frac{\partial H}{\partial L_H}}{\frac{dA}{dL_A}} \quad (9)$$

This is the ratio which market-mediation brings into equality with a given relative price constellation. And only if it should reflect the true marginal cost to society of an extra apple in terms of foregone honey -- the marginal rate of transformation, that is, between H and A -- can markets be judged efficient.

What is MRT in our model? Take the two production functions (1) & (2) and differentiate each totally:

$$dA = \frac{dA}{dL_A} \cdot dL_A \quad (1')$$

$$dH = \frac{\partial H}{\partial L_H} \cdot dL_H + \frac{\partial H}{\partial A} \cdot \frac{dA}{dL_A} \cdot dL_A \quad (2')$$

Divide (1') into (2') and cancel to get, in absolute (cost) terms:

$$\frac{dH}{dA} = \left| \frac{\frac{\partial H}{\partial L_H}}{\frac{dA}{dL_A}} \right| - \frac{\partial H}{\partial A} \quad (10)$$

If, then, $\frac{\partial H}{\partial A} > 0$, the true marginal social cost of an "extra" apple, in terms of honey foregone, is less than the market-indicated private cost, as given by (9). It is less precisely by the amount of positive "feed-back" on honey production due the "extra" apple.

Diagrammatically, the absolute slope, in Fig. 11, of the production possibility curve, F-F, reflects the $\left(\frac{dH}{dA}\right)$ of (10). In response to a fixed $\frac{P_A}{P_H}$ ratio, as, say, given out by the solution to a standard welfare-maximum problem, and denoted in Fig. 11, by $|\tan b|$, the socially desirable output mix is at z. But at z the private marginal

cost ratio: $\frac{\frac{\partial H}{\partial L_H}}{\frac{dA}{dL_A}}$, is greater than $\frac{dH}{dA}$: greater by $\frac{\partial H}{\partial A}$. Hence

private producers will not be satisfied to produce at z but will tend to a point of less A and more H . The converse would be true if $\frac{\partial H}{\partial A}$ were negative.

By combining (5') and (6'), eliminating w , and dividing through by $\frac{dA}{dL_A}$, we get the condition for Pareto efficiency in terms of private M.C.-s:

$$\frac{\frac{\partial H}{\partial L_H}}{\frac{dA}{dL_A}} = \frac{P_A}{P_H} + \frac{\partial H}{\partial A} \quad (11)$$

Clearly price equal to private marginal cost will not do. Further, if prices are market-determined, they will diverge from true, social marginal cost.

* * *

Note that:

(1) Throughout the elaboration of the model we assumed all around convexity, hence that internal tangencies necessarily represent points of maxima rather than minima. Could the feed-back effect: $\frac{\partial H}{\partial A}$, imply non-convexity in the production possibility curve?

It so happens that as long as both A and H are produced at non-increasing returns to variations in labor alone (as was originally assumed), and as long as the marginal rate of substitution (MRS)

between L_H and A as inputs in H-production decrease along normally convex isoquants, the positive external effect cannot cause non-convexity in the set of feasible output points. As a matter of fact, should the production functions show increasing returns to labor alone, the positive externality will (given our isoquant assumption) tend to offset the tendency to non-convexity. It acts precisely as do the efficiency-imposed variations in input proportions in the two input-increasing returns to scale case of Part V, Chapter One. There we found that where the MRS-equality requirement of efficient production gives rise to changes in input proportions for variations in the output mix, such changes in proportions could sufficiently offset the tendency to non-convexity of the output set -- due to the increasing returns to scale shown by the production functions -- to give a normally convex F-F despite such increasing returns to scale.

In the present setup, the feed-back effect brings about just such a change of input proportions in honey production, only more so. For in our Chapter One setup, more A and less N associated with both more L_A and more D_A ,¹ only their proportion in apple production became less and less favorable as more and more of the N-prone input: labor, had to be used on every apple-producing acre; conversely, for N. But here, more A and less H associate with an increased input in honey production of only one of the "factors". The other factor: A, is actually "withdrawn" from H as H expands. Hence the labor to apple-blossom ratio shifts very sharply in response to plus-H and minus-A;

¹ Land used in apple production.

and with decreasing MRS between L and A in apple production, more and more labor is needed to get the extra increment of H, hence paradoxically, even more A has to be given up. With some functions (e.g., if some A is required for plus H for whatever L_H), the production possibility curve can actually turn "south-west". This, incidentally, need not violate convexity, and anyway is of no significance, since if there is no disposal problem the positively sloped segments of an F-F are obviously irrelevant.¹

So much for convexity. Needless to say, the above does not imply that positive external effects need make for such curvature. One can invent all sorts of functions. The point is only that there is nothing inherent in external economies that necessarily makes for non-convexity.

(2) Any number of variations on the model suggest themselves. As Professor Meade pointed out,² the most general case can be denoted by production functions of the form

$$\left. \begin{aligned} A &= A(L_A, H(L_H, \dots), L_H) \\ H &= H(L_H, A(L_A, \dots), L_A) \end{aligned} \right\} \quad (12)$$

¹ It is of no interest here, but amusing, to work through the convexity implications of an external diseconomy. Unfortunately, it is not as easy to fix on an esthetically pleasing MRS assumption.

² Op. cit.

Interactions can be mutual, and as also emphasized by Meade, need not be associated with the outputs. Even in our case, it is perhaps more suggestive to think of L_A as producing some social value-product both in the A-industry and the H industry. In the most general formulation one can simply think of each production function as containing all the other variables of the system, some perhaps with zero weight.

(3) The question of whether technological external economies involve shifts of each other's production functions, or mutually induced movements along such functions, is purely definitional. If one chooses so to define each producer's function as to give axes only to inputs and outputs that are purchased and sold, or at least "controlled",¹ and the effects of everything else impinging on production (e.g., atmosphere, apple blossoms, etc.) are built into the curvature of the function, then it follows that all externalities proper will consist in shifts of some functions in response to movements along others. On the other hand, if, as in our apple-honey case, it seems useful to think of the production function for H as having an A-axis, then, clearly, induced movement along the function is a signal of externality.

(4) In our one scarce input model, the only possible distortions are in the allocation of the scarce L between the two industries, and in the final output mix. It would be simple to extend the analysis

1

Never mind the obvious ambiguities: costless control is implied, and nothing in the purchase-sale conception precludes inventories (though these introduce a dynamical element).

to two inputs: labor and land, and work out the consequences of externality for input-proportions and income distribution. Meade's previously cited Economic Journal article does much of this.¹

Some Queries

The modern formulation of the doctrine of external economies is not only internally consistent: it also yields insight. Yet one may well retain about it some residual unhappiness and dissatisfaction. There is no doubt that the Robertson-Sraffa-Viner distinction between the technological sort and the pecuniary gets to the nub of what is the matter with the original Marshallian analysis. It cuts right through the confusion which led Marshall and Professor Pigou to conclude that the price mechanism is faulty in situations where in truth it is at its best: in allocating less than infinitely elastic inputs between alternative productive uses. It also facilitates unambiguous formulation of the more difficult "falling supply price" case. But in a sense it only begs the fundamental question: what is it that gives rise to "technological" externalities, to the existence, in Pigou's terms, of "uncompensated services" and "incidental uncharged disservices"?²

1 Though his focus is so narrowly on distribution that some of the production-efficiency aspects of his model are not brought to a clear focus.

² The Economics of Welfare, (Fourth Ed., 1932), esp. Chapt. IX, Part II.

Most modern writers have let matters rest with more or less the Ellis-Fellner type explanation: "the divorce of scarcity from effective ownership"¹. Does non-appropriability, then, explain all externality? In a sense it does, yet because it tends to focus attention on institutional and organizational arrangements, whose invariance is of a different order from e.g., that of the laws of physics, the result is to distract attention from some deeper issues. Surely the word "ownership" serves to illuminate but poorly the phenomenon of a temperance leaguer's reaction to a hard-drinking neighbor's ("sound-insulated and solitary") Saturday night, or the reason why the price system, if efficient, will not permit full "compensation" for the "services" of a bridge?

There are other hints of difficulty and of problems unresolved. The allocation and production theory literature tends to dispose of externalities, especially inter-producer externalities, as of little empirical significance. The very examples cited to illustrate their existence -- apples, honey, birds, bees, marshes -- suggest the trivial, at least in an industrial context. Yet in the economic development literature notions of external economies are assigned a central role. Do they then matter after all?

Some aspects of the seeming inconsistency are easily resolved. The production and value theory discussions are explicitly stationary and statical, hence non-reversible externality phenomena are ruled out.

¹ "External Economies and Diseconomies", A.E.R., 1943, reprinted in A.E.A. Readings in Price Theory (supra).

The development literature is under no such restraint. Further, the only genuine production externality in the context of competitive equilibrium theory is the technological variety. Pecuniary externalities are disposed of as due either to technological externalities elsewhere in the system or to internal technological economies in supplier firms. While these last are of crucial importance to equilibrium theory -- they give rise to the monopoly or imperfect competition problem -- in the context of that theory they have nothing whatever to do with externality.

Yet the resolution is not complete. Pecuniary externalities are assigned a role in much of the writing about development that clearly cannot be justified on the level of statical, steady-state equilibrium theory. Some of the emphasis is to be explained, even on a sympathetic reading, by sheer confusion going back to that of Marshall and Pigou. But not all. We shall, in Part IV, briefly suggest the way in which pecuniary external economies might well play an independent role -- one distinct, that is, from being simply a signal of possible monopoly troubles.¹ The argument involves time-path problems of adjustment to equilibrium, and imperfect information, and is of particular relevance to "social overhead".

¹ An unreliable signal at that. A rightward shift of the demand curve facing a monopolist producing at a point of declining M.C. (both before and after) need not reduce his (M.R. = M.C.)-associated selling price. "Before and after" elasticities are involved.

But first, we must back to the first set of issues. If "non-appropriability" is, by itself, too flimsy a base for a doctrine of generalized externality, what sturdier foundation is there? Part B consists in an attempt to provide the design of one. Specifically, an ordering is suggested that brings the deeper issues associated with externality into the foreground. Part C, then, attempts the linkage with social overhead.

B. Externalities: an Ordering

The hierarchy of "existence" and "efficiency" tests elaborated in Part I suggests a consolidation of externality phenomena into three "polar" types: 1) Organizational Externalities, 2) Technical Externalities¹ and 3) Public-Good Externalities. These are not designed to be mutually exclusive: most externality phenomena are in fact blends. Yet there appears to exist a sufficient three-cornered clustering to warrant separation, both by cause and by consequence.

Type (1): Organizational Externalities

Imagine a world which exhibits generalized technological and taste convexity, where the electronically calculated solution of a Paretian maximum-of-welfare problem yields not only a unique set of

¹ I should much prefer "technological", but since this would necessarily confuse my Type (2) with Professor Viner's "technological" I tentatively fixed on "technical".

inputs, outputs and commodity-distribution, but where initial endowments plus lump-sum transfers render income distribution optimal in terms of the community's implicit but operative social welfare function. In this world, then, the marginal-productivity-imputed distribution of income would, at the "bliss point", be so corrected as to become consistent with the output-distribution implied by that point.¹ Further, the duality theorem and our convexity assumption assure that there exists, associated with the bliss point, a unique² set of prices, wages, rents, etc., in response to which decentralized profit and preference maximizers would sustain the production, exchange and distribution configuration required for bliss. All the production points constitute positions of maximum and non-negative producer's profits. Everything that matters is conventionally rationable, and either available in inelastic total supply,³ or producible at constant returns to scale. Assume, also, that tastes are entirely independent and that there is no uncertainty.

This is an Adam Smith dream world. Yet it is conceivable that due to more or less arbitrary and accidental circumstances of institutions, laws, customs, or technical convenience, competitive markets

¹ All this is needed to abstract from the income distribution problem and permit exclusive focus on Pareto efficiency.

² Or only inessentially indeterminate.

³ That supply need not, of course, remain constant: manna could fall (into the central storage bins) daily.

would not be fully Pareto-efficient.¹ Take, for instance, the Meade example of apples and honey. Apple blossoms are "produced" at non-increasing returns to scale and are (we assumed) an ordinary, private, exhaustible good: the more nectar for one bee, the less for another. It is easy to show that if apple blossoms have a positive effect on honey production, and abstracting from possible satiation and redundancy, a maximum-of-welfare solution, or any Pareto-efficient solution, will associate with apple blossoms a positive Lagrangean shadow price.² If, then, for reasons of "organizational" feasibility, markets do not impute to apple blossoms their correct shadow value, profit maximizing decisions will fail correctly to allocate resources (e.g., L) at the margin.

This is what I would call an organizational externality. It is essentially Meade's "unpaid factor" case. Non-appropriation, "divorce of scarcity from effective ownership", is the binding consideration. Certain "goods" (or "bads") with determinate shadow-values are simply not attributed. It is irrelevant here whether this is because the lake where people fish happens to be in the public domain, or because "keeping book" on who produces and who gets what may be clumsy (or costly in terms of resources)³. For whatever "organizational" reasons,

¹ Competitive behavior may have to be imposed in the allocation of non-produced scarcities which may be subject to monopolization.

² Set up a variant of our Apple-Honey model of Part III - A, introducing apple blossoms (B) explicitly. Add a production function $B = B(L_A)$, and substitute $B(L_A)$ for $A(L_A)$ as the second input in honey production. The solution will give out a positive Lagrangean shadow price for B, and profit maximizing producers of the joint products: A and B, will push L_A to the socially desirable margin.

³ Though on this last, see Comment #1, below.

certain variables which have positive or negative shadow value are not "assigned axes". The beekeeper thinks only in terms of labor.

The important point is that the difficulties reside in organizational arrangements, availability of information, etc. The scarcities at issue are rationable and finely divisible and there are no difficulties with total conditions. At the solution configuration every activity pays for itself. Apple blossoms have a positive and determinate shadow price, which, if it were but quoted by a central authority, would cause its production in precisely the right amount and even its distribution would be correctly rationed (though here feasibility may cause trouble).

Most of the few examples of inter-producer external economies of the reversible technological variety are of this type: "shared deposits" of fish, water, oil (if there are no pressure problems), etc.¹ Much more important, so are certain irreversible dynamical examples associated with investment. For instance, many of Pigou's first category of externalities: those that arise in connection with owner-tenant relationships where durable investments are involved, have a primarily "organizational" quality.² Perhaps the most important instance is the training of non-slave labor to skills -- as distinct from education more broadly (which belongs to Type (3)).

¹ Though indivisibility elements enter into some of these. Why can't somebody "own" part of an oil well?

² When not simply due, in a world of uncertainty, to inconsistent expectations.

In the end, however, and in particular if restricted to reversible statical cases, it is not easy to think of many significant "organizational externalities" pure and simple. Yet it will turn out that only this type of externality is fundamentally due to non-appropriation as such.¹

Type (2): Technical Externalities

Assume (a) that all goods and services are rationable, exhaustible, scarcities: more for Crusoe implies less feasibly available for Friday; (b) that individual ordinal indifference maps are convex and sensitive only to own-consumption (no taste externalities); and (c) that there exist no organizational "defects" of Type (1). If, then, the technology exhibits indivisibility or (smooth) increasing returns to scale in the relevant range of output, these give rise to what I call "technical externality".² Since indivisibility phenomena were examined, exhaustingly if not exhaustively, in Part II above, they require only brief comment here.

In situations of pure "technical externality" there does, of course, still exist a maximal production possibility frontier (F-F); and with a Samuelson-type social indifference map (S-S), i.e., a map "corrected" for income distribution, it is possible, in concept, to define a bliss point(s).³ Also, where indivisibility is exhibited by outputs and only

¹ There are, clearly, a number of loose ends associated with Type (1), e.g., what is "feasibility" in this context? Some of these are commented on below.

² Again, this is not the same as Viner's "technological".

³ This is saying very little, of course, except on the level of metaphysics and philosophy.

outputs, or, stronger, where smoothly increasing returns to scale is the only variety of non-convexity -- isoquants, for one, are properly convex -- the locus of efficient output combinations can be defined in terms of conditions on marginal-rates-of-input-substitution.¹ Moreover, bliss could possibly occur at a point where S-S is internally tangent to F-F, perhaps to a convex F-F.² But even in the least "pathological", most neo-classically well-behaved case, where there exists a meaningfully defined set of shadow prices associated with the bliss point, genuinely profit seeking competitive producers, responding to that set of prices, would fail to sustain optimal production.³ At best, even if at the bliss point configuration all M.C.-s are rising, some producers would have to make continuing losses, hence would go out of business. If, in turn, prices are not pre-calculated and/or centrally quoted but permitted to set themselves, monopoly behavior will result.

Further, bliss may require production at levels of output where losses are not only positive, but at a maximum (subject only to engineering efficiency). If so, the embedded Lagrangean constants may still retain meaning as marginal-rates-of transformation, but they will fail to sustain efficient production even by Lange-Lerner civil

¹ Inequalities due to kinks and corners are as good as equalities where all is smooth.

² Recall, however, that if F-F is not convex, a point of internal tangency between S-S and F-F, should one exist, could be a point of minimum S, or a point of "maximum non maximorum".

³ We again abstract, in all that follows, from the distribution problem. Posit "lump-sum" redistribution.

servants who care only about margins and not totals. $p = M.C.$ might still be correct; but it is insufficient. No decision about output-mix is possible without regard to total conditions. By themselves, neither internal tangencies nor corner tangencies, whether maxima or minima, can suffice.

On the other hand, given our assumptions, the Paretian contract locus of maximal (ordinal) utility combinations which is associated with any one particular output point is defined, as in the trouble-free neo-classical model, by the usual subjective, taste determined, marginal-rate-of-substitution equalities (or, at corners, inequalities). These MRS equalities, in turn, imply a set of shadow prices which if centrally quoted, would efficiently ration as among consumers the associated (fixed) totals of goods. Further, the particular exchange-prices implied by the bliss-point output mix would secure its optimal distribution. In the sphere of exchange, in other words, a decentralized price system works without flaw.

In what sense does this kind of situation exhibit "externality"? In the (generalized) sense that with regard to production decisions a price-market game fails as an analogue computer because, by its very nature, it cannot correctly account for the full social benefits associated with certain activities. With Type (1) externalities, though market institutions fail correctly to "solve" for the bliss values of all q -s and p -s, there exists embedded in the solution a set of p -s, which would result, if used for purposes of decentralized signaling, in the maximum welfare input-output and distribution

configuration.¹ This is not the case here. In Type (1) situations, at the bliss point there is complete correspondence between social and private pay-off, both at the margin and in totals. Here there is not: there is necessary divergence at least in the total conditions, and possibly at the margin. At the set of prices which will correctly ration the bliss-point bill of goods that bill of goods will not be produced by profit seekers, and may not be produced by Lange-Lerner civil servants.

There is one qualification to be made to the above. We have assumed that the bliss point requires production of at least some goods at points in the neighborhood of which the relevant production functions are non-convex. This was implied by our initial assumption that increasing returns and indivisibility obtain in the "relevant ranges". If this assumption is dropped, it may be that bliss occurs at a point where S-S is internally tangent to a locally convex F-F and where, further $M.C. \gg A.C.$ (evaluated at the solution prices). In this case no "externality" divergence of social and private calculation will occur. But (as emphasized in Part II) unless F-F is in fact convex throughout, the existence of such a locally stable tangency cannot be taken as evidence that that point is in fact the bliss point: the particular S-contour could cross F-F at a distance. Again, it is not possible to do without total conditions.

The crucial distinction between this kind of externality and that labeled "organizational", is that here the non-appropriability notion, as generally conceived, tends to miss the point. Strictly

¹

Remember, incomes are being optimally redistributed.

speaking, it is of course true of Type (2) situations that the price system does not "appropriate" the full social benefits of activities showing increasing returns to scale to those engaged in them. But the existence of such "uncompensated services" has in this case nothing whatever to do with "divorce of scarcity from ownership" or feasibility limitations on "exclusion". It is entirely feasible to own a bridge and profitably ration crossings; indeed, a private owner would do so. The point is, rather, that such profitable rationing, such "compensation" for services rendered, would, if bridge crossings were at less than capacity, inefficiently misallocate the "output" of bridge crossings. If in terms of scarce resource inputs the marginal cost of an additional crossing is zero, any positive toll will, in general, have the usual monopolistic effect: the resulting output configuration will be non-efficient. The essential characteristic of Type (2) externality is that the ration-price required by Paretian exchange efficiency is less than that required to cover cost.

This, incidentally, is where most pecuniary external economies lead: a supplier produces in a range of declining M.C. due to internal technological economies to scale, hence cannot make "ends meet" at the socially correct price and output level. The crucial associated difficulty at the level of social organization is monopoly. The only "externality" implied relates to total conditions: at the "solution" price total revenue less total cost will not correctly indicate the net social benefit that the activity yields.

All of this is, of course, well known. Marginal cost pricing of utility services has been, since Hotelling,¹ a constant source of articles. The only justification for the above discussion is that this is the one kind of externality that appears certainly and importantly associated with social overhead capital: with roads, canals, harbors, dams, etc. In a sense we are back where we started: indivisibility in scale.

Can we leave matters at that? Not quite. There is a third kind of externality, recently emphasized by Professor Samuelson, that is germane. We must briefly examine the relevance to "social overhead" activities of the notion of "public goods."

Type (3): Public Good Externalities

In some recent writings on public expenditure theory, Professor Samuelson has re-introduced into current discussion the notion of the collective or public good. The defining quality of a pure public good is that "each individual's consumption of such a good leads to no subtractions from any other individual's consumption of that good ...",² hence, "it differs from a private consumption good in that each man's consumption of it, X_2^1 and X_2^2 respectively, is related

¹ For references, see Appendix to Chapter One.

² P. A. Samuelson, Review of Economics and Statistics, Nov. 1954, p. 387.

to the total X_2 by a condition of equality rather than of summation. Thus, by definition, $X_2^1 = X_2$ and $X_2^2 = X_2$.¹

As Samuelson has shown, the form of the marginal rate of substitution conditions which define the Pareto-efficient utility possibility frontier in a world where such public goods exist, or at least where there are outputs with important "public" qualities, renders any kind of price-market routine virtually useless for the determination of the output-mix and of distribution, whether for purposes of computation or organizational decentralization. Where some restraints in the maximum problem take the form: total production of X equals consumption by Crusoe of X equals consumption of X by Friday, Pareto efficiency requires that the marginal rate of transformation in production between X and Y equal not the (equalized) MRS of each separate consumer, but of the algebraic sum of such MRS-s. This holds, of course, in what in other respects is a conventionally neo-classical world: preference and production functions are of well-behaved curvature, all is convex.

If then, at the bliss point, with Y as numeraire, p_X is equated to the marginal Y -cost of X in production (as is required to get optimal production), and X is offered for sale at that p_X , preference maximizing consumers adjusting their purchases so as to equate their individual MRS-s to p_X will necessarily "under use X ". A pricing game will not induce consumers truthfully to reveal their preferences. It pays each consumer to under-state his desire for X relative to Y ,

¹ P. A. Samuelson, Review of Economics and Statistics, Nov. 1955, p. 350.

since his enjoyment of X is a function only of total X, rather than, as is true of a pure private good, just of that fraction of X he pays for.

The two Samuelson articles¹ explore both the analytics and the general implications of "public goods". Here the notion is of relevance because much externality is due precisely to the "public" qualities of a great many activities. For example, the externality associated with the generation of ideas, knowledge, etc., is due to the public good character of these "commodities". Many inter-consumer externalities are of this sort: my party is my neighbor's disturbance, your nice garden is any passerby's nice view, my children's education is your children's good company, my Strategic Air Command is your Strategic Air Command, etc. The same consumption item enters, positively or negatively, both our preference functions. The consumptions involved are intrinsically and essentially joint.

This kind of externality is distinct from either of the other two pure types. Here technological non-convexities need in no way be involved. In fact the $MRT = \sum MRS$ condition is certain to hold true precisely where production takes place at constant or non-increasing returns to scale, hence where the production possibility set is necessarily convex. Further, there are no decentralized organizational rearrangements, no private bookkeeping devices, which would eliminate the difficulty. It is the central conclusion of the Samuelson model that where public good phenomena are present, there does not exist a

¹ And a third unpublished paper, which was read at the 1955 A.E.A. meetings and to a copy of which I came to have access while this chapter was being written.

set of prices associated with the (perfectly definable) bliss point, which would sustain, not to speak of lead-to, the bliss configuration. The set of prices which would induce profit-seeking competitors to produce the optimal bill of goods, would be necessarily inefficient in allocating that bill of goods. Moreover, even abstracting from production, no single set of relative prices will efficiently ration any fixed bill of goods so as to place the system on its contract locus, except in the singular case where at that output and income distribution point MRS-s of every individual are identically the same or zero.¹

* * *

The above ordering warrants a number of comments:

1) In a sense, Type (1) is not symmetrical with the other two categories. One can think of some non-trivial instances where the organizational element does appear to be "binding"; skill-training of people, for example. But even there, it could be argued that the crucial elements are durability, uncertainty and the fact that slavery as a mode of organization is itself in the nature of a public good which enters people's preference functions, or the implicit social welfare function, non-separably from the narrowly "economic" variables. In those instances, in turn, where bookkeeping feasibility appears to be the cause of the trouble, the question arises why bookkeeping is

¹ On the other hand, a precalculated and centrally quoted price vector will sustain, where all is convex on the production side, the production of an associated bill of goods. Prices can work, in this limited sense, in efficiently guiding production. Note the symmetry with Type (2), where price calculations fail in production but work in exchange.

less feasible than where it is in fact being done. In the end, it may be that much of what appears to partake of Type (1), is really a compound of Types (2) and (3), with dynamical durability and uncertainty elements thrown in. At any rate, a deeper analysis of this category may cause it substantially to shrink or vanish.

2) The relation of my tri-cornered ordering to Professor Meade's polar categories is of interest.¹ His first category, "unpaid factors", is identical to my Type (1). But he lumps my Types (2) and (3) into a composite category labeled "atmosphere". Meade's qualitative characterization of "atmosphere": e.g., of afforestation-induced rainfall, comes very close to the Samuelson public good definition.² Unfortunately he links this as necessarily bound up with increasing returns to scale in production to society at large, hence an Euler-caused, J. B. Clark-like, over-exhaustion problem.

If one abstracts from shared water-table type problems (let rain-caused water input be rigidly proportional to area) then Farmer Jones' rain is Farmer Smith's rain and we have my Type (3). However, and this is where Meade appears to go wrong, there is no reason why either farmer's production function (with an axis for rain) need show

¹ Op. cit. I had fixed on my breakdown when on a re-reading of Meade's article I found that he was in search of the same kind of ordering.

² See esp. bottom of p. 61 and top of p. 62, op. cit.

increasing returns to scale. It may be that without rain returns to bundles of non-rain inputs diminish sharply. In fact, and as we have seen, the public good quality of rainfall will cause an independent difficulty, i.e., that it ought to be "produced" by timber growers until its M.C. is equal to the sum of all the affected farmers' MRS-s for rain as an input, whatever may be the curvature of the latter's production functions.¹

3) My own (probably transitory) subjective sense of relative order (or at least lesser chaos) in these complicated matters owes a great deal to the Samuelson public good model. In particular, the model helps to separate the phenomena Meade lumped as "atmosphere". Interestingly, Professor Samuelson himself emphasizes the analytical bond between indivisibility and public good situations. In both an explicit "summing in" is required of "all direct and indirect utilities and costs in all social decisions".² In the pure Type (3) situation only utilities and costs at the margin require adding up, and while price cannot be used to ration the so-calculated bliss point output, it (price) would induce producers to produce it. In Type (2), on the other hand, it is the intra-marginal consumer's and producer's surpluses associated with various all or nothing decisions "in-the-lump"

¹ It should be said that Professor Meade concludes his article: "But, in fact, of course, external economies or diseconomies may not fall into either of these precise divisions and may contain features of both of them." This, incidentally, is my after-thought also.

² P. 9, A.E.A. Draft (supra).

that have to be properly (interpersonally) weighted and summed. Once done, the bliss outputs can then be properly rationed by price. But at the correct ration-price producers will require subsidy. The crucial shared quality of the two categories lies in the need to sum utilities over many people.¹

4) One more comment may be warranted on the significance, in a public good situation, of non-appropriability. "Exclusion" is almost never impossible. A recluse can build a wall around his garden, Jones can keep away his (educated) children from those of Smith, etc. But if thereby some people (e.g., the recluse) are made happier and some (e.g., the passers-by) less happy, any decision as to whether to "exclude" or not implies an algebraic summing of the somehow-weighted utilities of the people involved. And if the wall requires scarce resources, the final utility sum must be matched against the cost of the wall. When Type (3) blends with indivisibility in production, as it does in the case of the wall, the comparison has to be made between intra-marginal totals. Where no lumpiness is involved (e.g., the decibels at which I play my radio) only MRS and perhaps M.C. calculations are called for. But the really crucial decision may well be about how much perfectly feasible appropriation and exclusion is desirable.

5) It is of interest to speculate what nature of arrangements may be required to offset the three categories of externality. In

¹ For a possible exception to the need to add up utilities in the case of a Type (2) situation, see below.

concept, Type (1) can be offset by rearrangements of ownership and by proper bookkeeping, such as need not violate the structural requirements of pure competition. Further, no resort to non-market tests would be required.¹

Types (2) and (3) are not so amenable to correction consistent with decentralized institutions. The easiest possible case occurs where increasing returns obtain on the level of single producers'-good plants, much of whose production can be absorbed by a single user firm. Here vertical integration takes care of the problem. Not every process inside a well running firm is expected to cover its cost in terms of the correct set of internal accounting (shadow) prices. Total profits are the only criterion and it may be worth it to a firm to build a private bridge between its two installations on opposite sides of a river yet charge a zero accounting price for its use by the various decentralized manufacturing and administrative divisions. The bridge will make accounting losses, yet total company profits will have increased.

Note, however, that the profit-maximum scale of operation of the integrated firm is likely to be greater (cet. par.) than if the river had not been there to span, or could be spanned by some means of lower fixed-cost-to-variable-cost ratio. Hence the monopoly problem may simply be "pushed forward" to consumer markets. But as long as such

¹ The Emancipation Proclamation could constitute, of course, a substantial barrier.

integration is consistent with the requirements of competitive markets, no extra-market tests are required. The private total conditions: T.R. less T.C., correctly account for the net social gain.

Where a producers'-good firm producing at a stage of falling A.C. sells to many customer firms and industries, an adding up of all the associated T.R.-s and T.C.-s at the pre-calculated "as if" competitive prices associated with the bliss point would again effectively "mop up" all the social costs and benefits.¹ But the institutional reorganization required to get correct "decentralized" calculations involves horizontal and vertical integration, and the monopoly or oligopoly problem looms large indeed. The Type (3) case of a pure producers' public good (are there any?) belongs here: only input MRS-s along production functions require summing.

In the general case of a mixed producer-consumer good (or of a pure consumer good) that is "public" or exhibits increasing returns to scale over the relevant range, it is impossible to avoid comparison of inter-personal utility-totals. Explicit administrative consideration must be given, if you like, to consumer's and producer's surpluses for which no market-institution tests exist short of that provided by a perfectly discriminating monopolist. But to invoke perfect discrimination is to beg the question. It implies knowledge of all preference functions, while as Professor Samuelson has pointed out,² the crucial

¹ Assuming that all consumer goods are finely divisible and require no lumpy decisions by consumers.

² Cf. any of the three "Public Expenditure" articles (supra).

game-theoretical quality of the situation is that consumers will not correctly reveal their preferences: it will pay them to "cheat".

6) Examination is needed of various blends of Types (2) and (3), or for that matter, and as suggested by Samuelson, of blends of public and private goods even without increasing returns to scale. There are many puzzling cases. Do bridge crossings differ in kind from radio programs? Both involve an indivisibility and, where V.C. is zero for the bridge, M.C. is zero. The correct price for an extra stroller, as for an extra listener, is clearly zero. Yet bridge crossings have a distinctly private quality: bridges get congested, physical capacity is finite. This is not true of a broadcast. There is no finite limit to the number of sets that can costlessly tune in.¹ Radio programs, then, have a public-good dimension. Yet, in a sense, so do bridges. While your bridge crossing is not my bridge crossing, in fact could limit my crossings, your bridge is my bridge. What is involved here is that most things are multi-dimensional and more than one dimension may matter ... but we must back to "social overhead".

C. Back to Social Overhead

On the level of pure statics there is little more to say. What one may choose to label as a social overhead activity is, within wide

¹ My colleague, Professor R.S. Eckaus, suggests that it is possible to exhaust the space to which the broadcast is limited and that this makes the situation a little more like that of a bridge. Neither of us is entirely satisfied, however.

limits, a matter of taste. But if we accept the usual connotation, in all its imprecision, and as elaborated in Part I, it would appear that the most significant associated "externality", at least in a statical context, brings one right back to indivisibility-in-scale, compounded, in some instances, by the public good quality of the output. Pure organizational externality, while it suggests the need for certain kinds of government activity especially in a dynamical context -- e.g., extra-market encouragement of the training of labor -- does not, on the basis of tentative speculation, appear as a particularly significant quality of transport, communication, etc.

On this view, matters fall into place rather nicely. In a statical context organizational externalities have no special link with social overhead. Activities producing outputs with a marked public good quality: e.g., police services, education, etc., are almost necessarily a public function. Whether the "public" body is a government, an established church, etc., is immaterial here. The point is that as regards such activities there is bound to be gross divergence between private and social pay-off. Activities exhibiting externality of Type (2), however, are in an ambiguous position. For one thing, the quantitative significance of indivisibility and increasing returns to scale is a function of scale, of markets. This explains the particular emphasis on the role of social overheads in low income countries trying to generate growth. Secondly, profitable private operation is not quite as manifestly "inefficient" as e.g., in the case of a modern national defense establishment. It may be that on the

level of strategy in a given concrete situation, monopoly or oligopoly operation comes close enough to achieving the correct "all or none" decisions-in-the-lump, to recommend itself as a mode of organization. But statical analysis of stationary "bliss" positions is a grossly inadequate basis for even tentative conclusions in this regard. Part IV consists in a sketch of the directions in which the analysis has to be extended to permit a clearer view. In particular, it is essential to reintroduce that other dimension of indivisibility: durability.

IV ON THE ECONOMICS OF "SOCIAL OVERHEAD CAPITAL": AGENDA

It has been my purpose to set out the elements of pure theory on which an economics of social overhead would have to rest. Yet even of this limited task, much remains undone. The following is in the nature of speculation about an outline for further exploration.

Dynamical Efficiency (1): Imperfect Information and Adjustment to Equilibrium

A statical system, whose solution gives constant values for all the variables, is a special and polar case of a dynamical model whose solution consists in steady-state time paths, not necessarily constant. It is of interest, and of significance for social overhead phenomena, to explore dynamical systems exhibiting lumpiness which retain the characteristic of constant equilibrium solution values, but permit explicit consideration of the adjustment paths of the variables from initial states of disequilibrium -- after, say, an exogenous change in the data. In a sense, such a system is implied by all comparative statics.¹ But comparative statical analysis stops with considerations of stability "in the small", and with lumpiness this won't do. The

¹ Discovery of the implied dynamics of statical systems led to many of the major advances in the field of micro-dynamics, by Hicks, Metzler and especially Samuelson, Part II of whose Foundations of Economic Analysis is devoted to a detailed exploration of this type of stationary dynamics. These advances, in turn, permitted deeper understanding, for instance, of the formal identity between the existence of a solution for a statical system and the convergence of a dynamical iterative sequence. (Cf., in addition to Foundations, and the references therein, esp. R.M. Solow, "On the Structure of Linear Models", Econometrica Vol. 20, 1952.)

existence of locally stable and market-sustainable solution values (as, for example, x_3 in Figure 8), does not guarantee, in a lumpy world of less than perfect information-at-a-distance, that such values will ever be achieved by a price-market type administrative or computational routine.

Note, incidentally, that in a context of adjustment dynamics, Viner's pecuniary external economies, besides signaling what I have called technical externality of Type (2), acquire a significance of their own -- a significance, moreover, not unlike that of organizational and informational externalities of the statical variety (Type(1)). They imply the need for ex ante calculation of costs and benefits at alternative price constellations that may be markedly different from that of the initial state. Further they may require that producers make large, lumpy, irreversible commitments on the basis of their anticipation of such a very different, post adjustment, price vector. Moreover, even if the "new" equilibrium configuration gives non-negative profits all around once all decisions are correctly made and implemented (e.g., as at x_3 in Fig. 8), there is a transition during which things may have to get worse rather than better.

All this is well known, but its relevance to social overhead warrants emphasis. It is in this context that the stress of the development literature on pecuniary externalities (or on "balanced growth") is defensible. Most social overhead activities provide what are in part at least in the nature of producers' services. Prior to the availability of a given kind of "social overhead" service, e.g.,

without a transport net, or electricity, potential user-industries may consist of lots of small producers, each more or less profitably covering a small localized market and using highly labor intensive constant-or-decreasing-returns-to-scale processes.¹ The feasible "bliss point", however, may call for the building of a power plant, or a market uniting transport net, and for the consolidation of production by the user-industries into large (potentially profitable) factory operations using mechanical processes which, at anything much below the solution scales of output, show sharply increasing returns.²

By assumption, the society would be better off if the required changes, all feasible, took place. It is possible, though not implied, that at the solution values even the transport or power system would be profitable: one "lump" could just be exhausted; its services would then be correctly rationed by a price that exceeds average total cost. If so, production of the bliss point bill of goods would be sustained by genuine profit seekers. But to get to bliss, entrepreneurs would have to make large, irreversible commitments whose profitability is sharply sensitive to other entrepreneurs making their commitments. Long construction and gestation periods reinforce the need for "vision" at a distance. Further, there is no presumption that at bliss the social overhead facility will cover its costs -- if

¹ In the programming sense: a vector of inputs, proportions among which are fixed.

² We are postulating that this is the case, hence arguments that it is not "likely" to be because of factor endowments are out of order.

anything, the opposite is more "likely". Implied post-adjustment losses will of course compound the inertia of the system.

Such sequential increasing returns situations -- irrespective of whether they are only transitional -- are particularly significant in low income countries, where many activities have to start from scratch. Since many social-overhead service inputs are intrinsically non-importable and difficult to substitute against, the potential user industries, in considering whether to make the commitments required of them by "bliss", may have to count on becoming completely dependent on as yet non-existent activities. In effect, calculation is required in terms of "infinite" price changes.¹

The above set of issues, involving as they do a compounding of imperfect information with lumpy and -- via vertical complementarity in production -- mutually interdependent decisions, are badly in need of systematic analysis. Unfortunately the analytical and even purely mathematical difficulties are formidable. But at the least, a systematic and rigorous formulation of the various component-problems

¹ Never mind, here, the many ambiguities and loose ends in all the above. Metaphor is not meant to substitute for rigorous analysis.

appears to be a manageable task. Further, some simple yet meaningful models suggest themselves, which might illuminate much of what is at stake.¹

Dynamical Efficiency (2): Efficient Paths of Capital Expansion

To take account of durability, an explicitly dynamical system is required which permits of non-constant solution paths for all

1

On this range of issues, the development literature, while suggestive and full of insight, is in a state of considerable confusion. Much remains to be done to exhaust and elucidate the content of the major seminal ideas, due, among modern writers, to Allyn Young (1), Rosenstein-Rodan (2), Nurkse (3), and others. In the most recent literature Fleming (4) and Scitovsky (5), have attempted to create some order. The former, by sticking to a statical frame misses, I think, the point. The latter, however, is, as usual, definitely on the right track. Of unpublished writings on this range of issues, I have had access to papers by Fellner (6) and by Laursen (7).

(1) "Increasing Returns and Economic Progress", Economic Journal, Dec. December 1928

(2) "Problems of Industrialization ...", Economic Journal, June-Sept. 1943

(3) "Problems of Capital Formation in Underdeveloped Countries", Oxford, 1953

(4) "External Economies and the Doctrine of Balanced Growth", Economic Journal, June 1955

(5) "Two Concepts of External Economies", Journal of Political Economy", April 1954

(6) "Individual Investment Projects in Growing Economies: General Characteristics of Problem and Comments on the Conference Papers", Investment Criteria and Economic Growth, Center for International Studies, M.I.T., 1955

(7) "External Economies in Economic Development" (unpublished)

the variables. For convex and, in particular, linear situations, there exist in the literature a number of such systems which do or can take account of inter-industry, inter-sectoral flow and stock requirements: Ramsay (1), von Neumann (2), Leontief (3), Malinvaud (4), and finally Solow and Samuelson (5), are among the principal contributors. The last have developed an intertemporal extension for Pareto production efficiency and have shown the existence of an associated price-rent-interest constellation for efficient paths.

Some of the elements that characterize such dynamical models are discussed in Chapters One and Three. Here, let it suffice to assert that:

(a) With linear production functions, dynamical phenomena of durable-capital formation and consumption, lags, stock-flow relations, etc., are amenable to formal maximizing treatment and are consistent with solubility and duality. While society at large needs vision at a distance -- there are nasty difficulties with "horizon" and the composition of "terminal" stock -- a price-market "game", with decentralized responses to prices and their instantaneous rates of change,

(1) Economic Journal, 1927

(2) Review of Economic Studies, Vol. XIII, No. 1

(3) Chapter 3, Studies in the Structure of the American Economy (Oxford, 1953)

(4) Econometrica, Vol. 21, #2

(5) Chapters 11 and 12 of Dorfman, Solow and Samuelson, Linear Programming and Economic Analysis (Rand Corp., forthcoming)

will sustain Pareto efficiency.¹ Durability, as such, does not cause trouble in terms of any of the existence and efficiency tests, elaborated in Part I above.

(b) Unfortunately, indivisibility-in-scale spoils things rather badly. Non-convexity is as damaging to dynamical systems as to statical. For one thing, modern mathematics provides only laborious numerical methods for handling even convex non-linear systems. As a result, we know very little about the properties of efficient solutions, about duality, etc. It is a good guess, however, that with non-convexity matters are chaotic. At any rate, it may be worthwhile to try numerical exploration of some very simple dynamical models with "lumps".

Dynamical Efficiency (3): Uncertainty

The central fact of grand dynamics is uncertainty. To the extent that the decisions required by a problem involve commitments whose justification lies in the shape of the future, uncertainty plays a central role. The more irreversible the commitments and the longer-lasting their effects, the greater is apt to be its significance. Since one of the identifying qualities of social overhead activities is the use of durable, long-lasting facilities, uncertainty must be given explicit consideration in any economics of social overhead. This, of course, is easier said than done.

In dealing with uncertainty, the development literature tends to focus on presumed divergences between private as against social risk calculations. While this is clearly of importance in gauging the

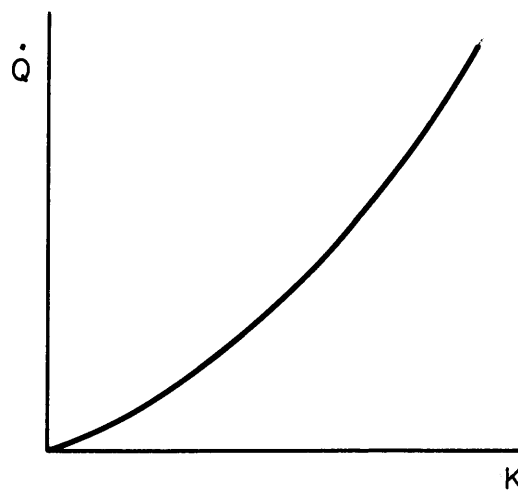
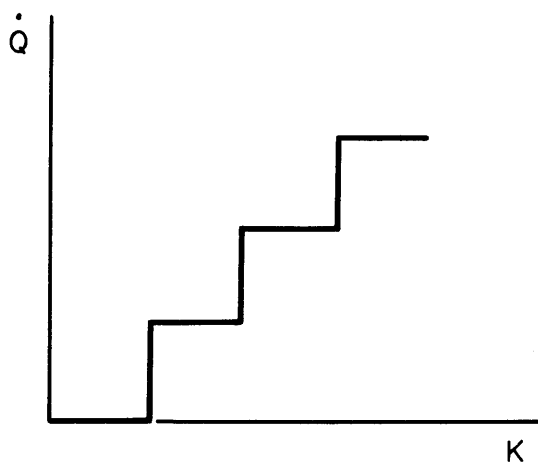
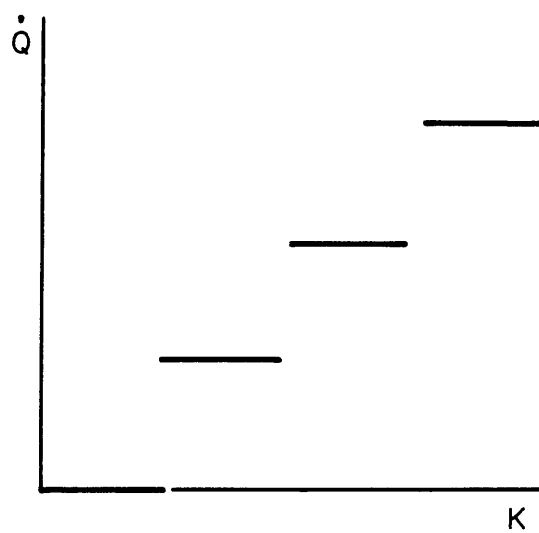
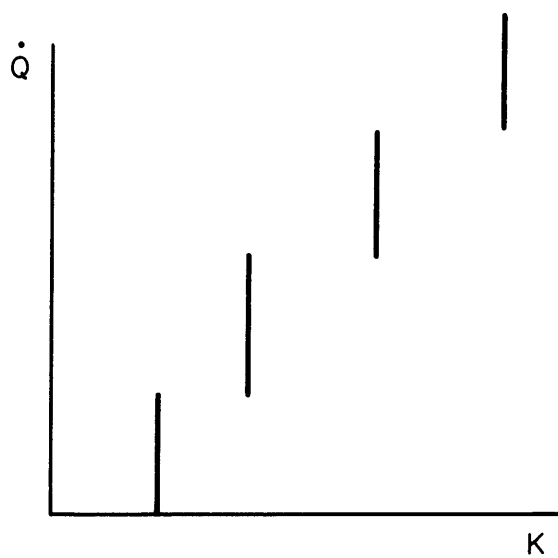
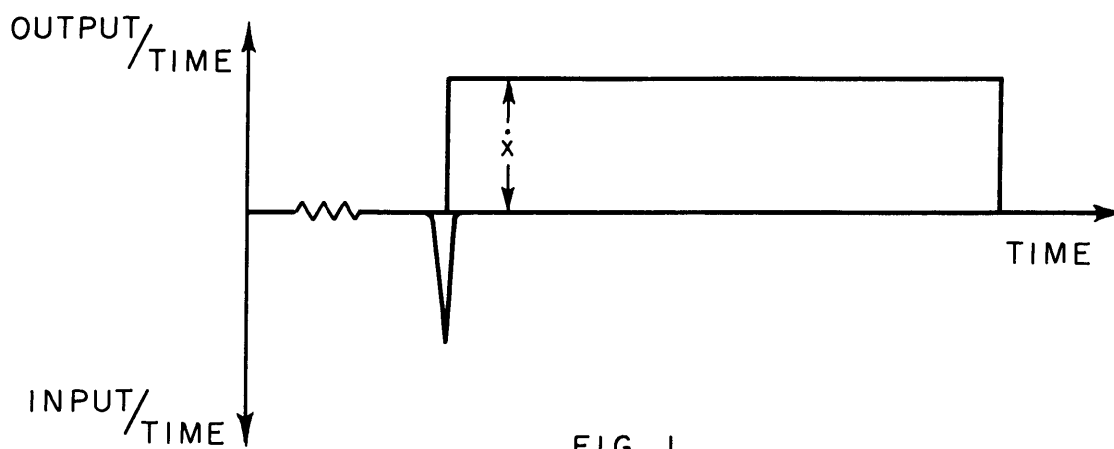
¹ Futuresmarkets may be required.

potential efficacy of market institutions, it tends to divert attention from a problem that in a sense is logically prior to the organizational issue. How should uncertainty influence decisions in regard to social overheads? The notion of flexibility, it seems, has been neglected by analytical economics, except on the purely statistical level (or in after-thought). Some effort to subject the notion to formal handling may pay considerable dividends.

On the level of pure economics, there is probably little to be done, however. It is likely that uncertainty will continue to gain the full recognition which is its due only on the level of planning strategy.

Growth

The influence of social overhead phenomena on growth potential constitutes too massive a subject matter to lend itself to brief speculative comment. It is evident, however, that our understanding of the role of indivisible and durable social overhead activities in the vicious circles which enchain a low income economy is rudimentary. But this is for another day.



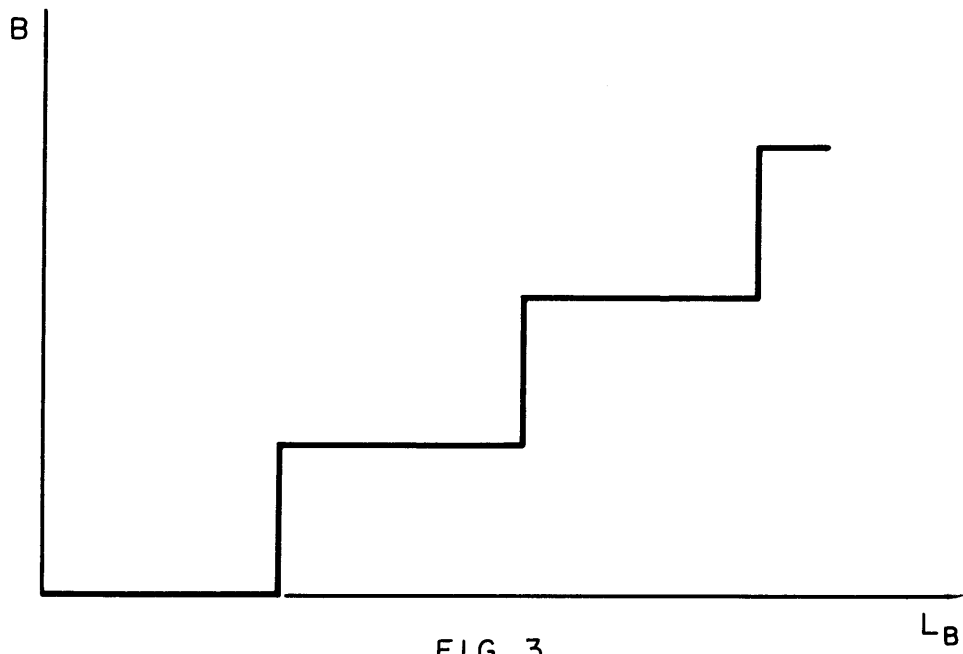


FIG. 3

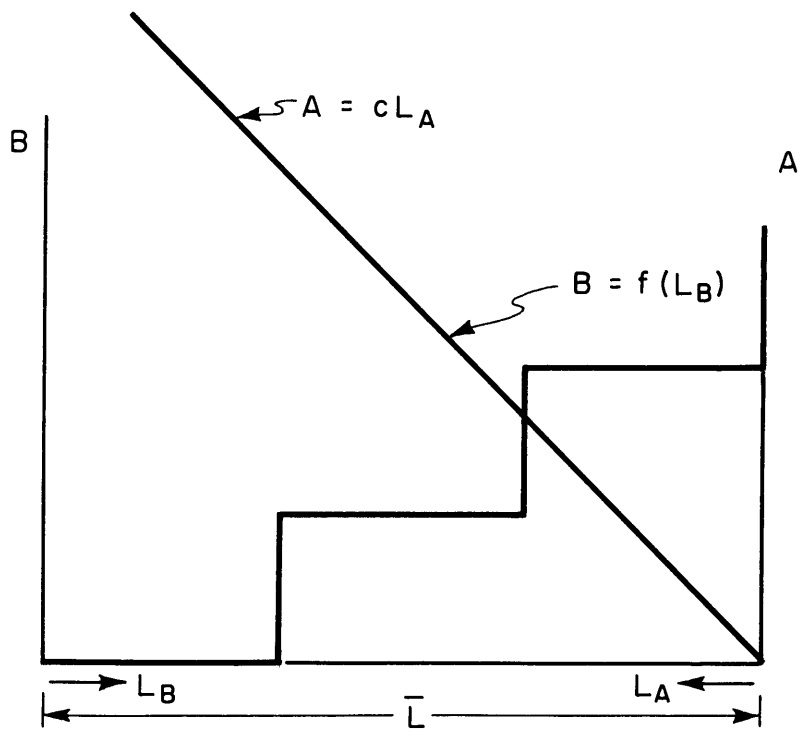


FIG. 4

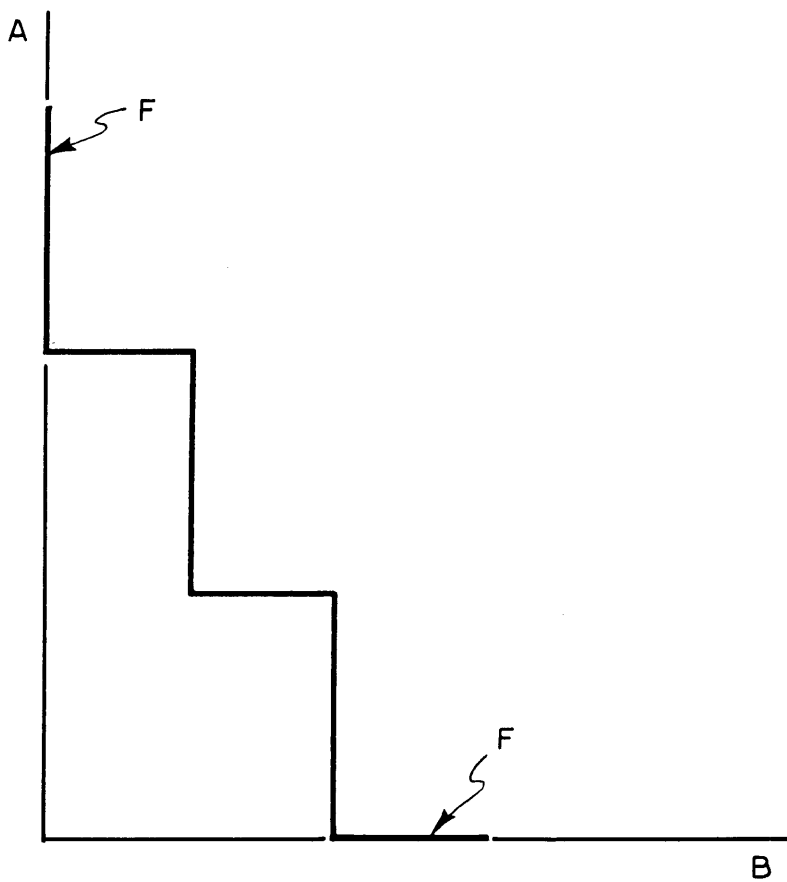
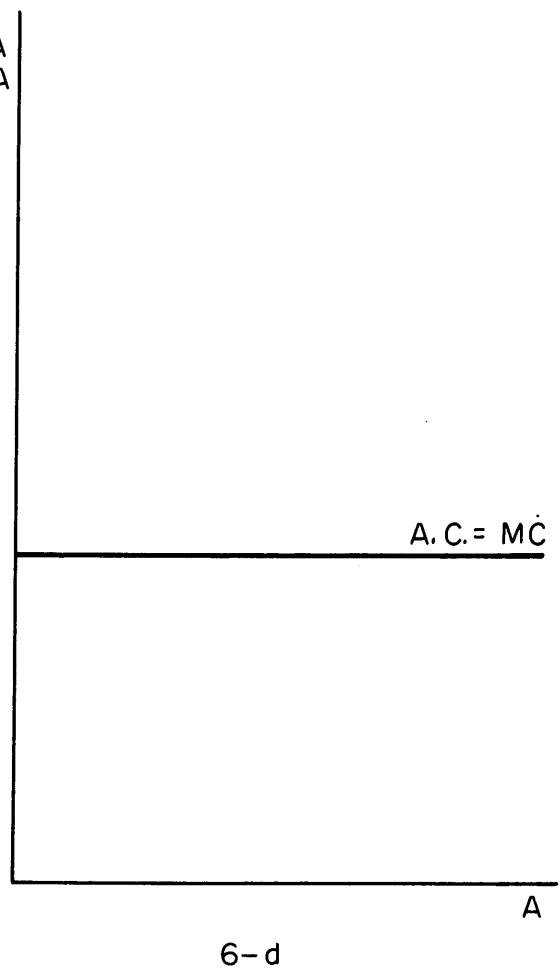
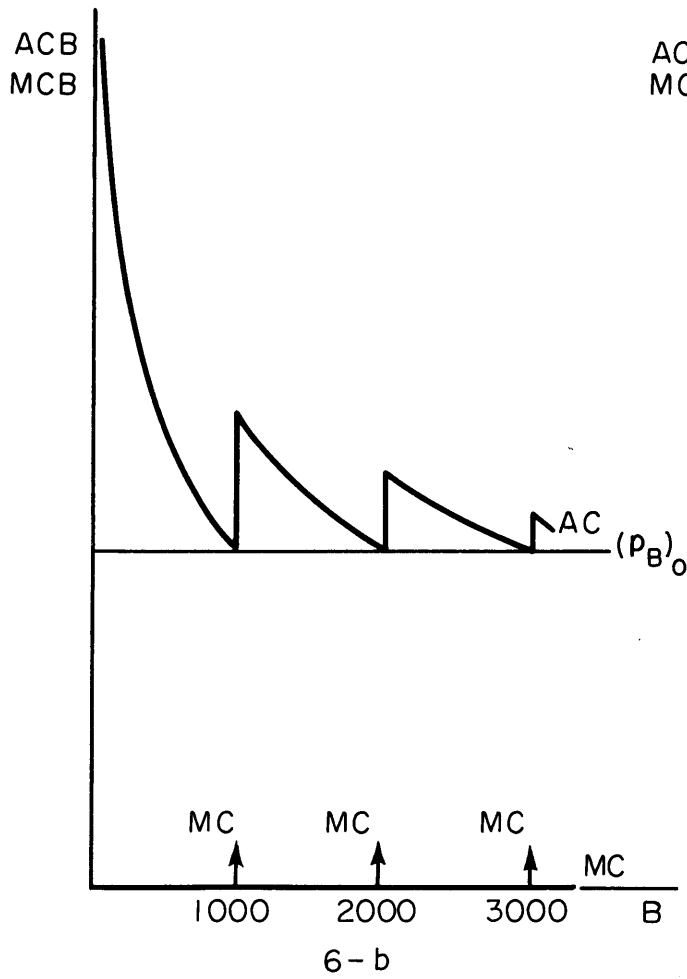
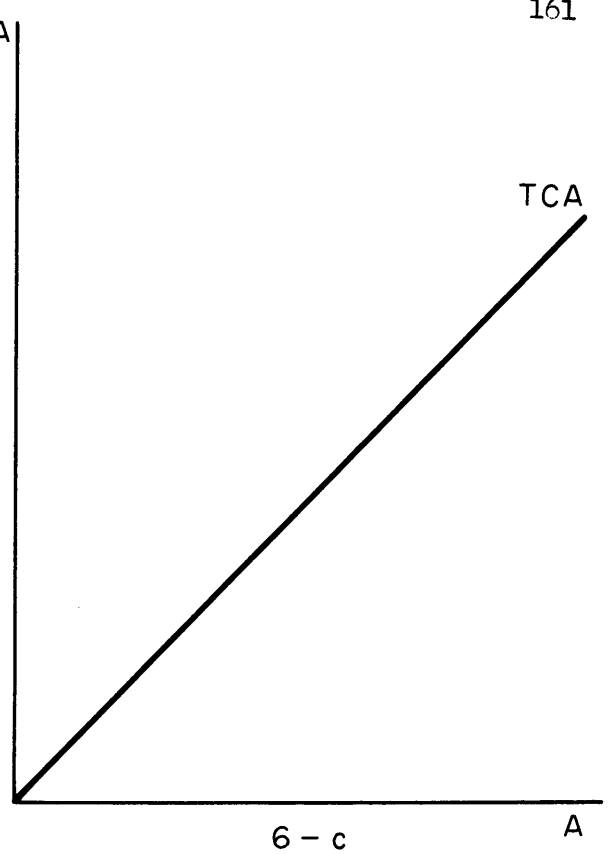
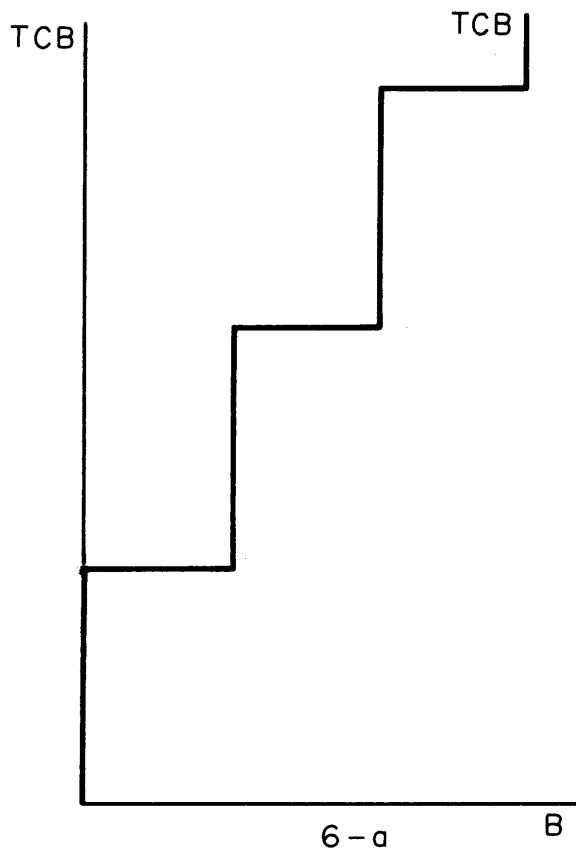
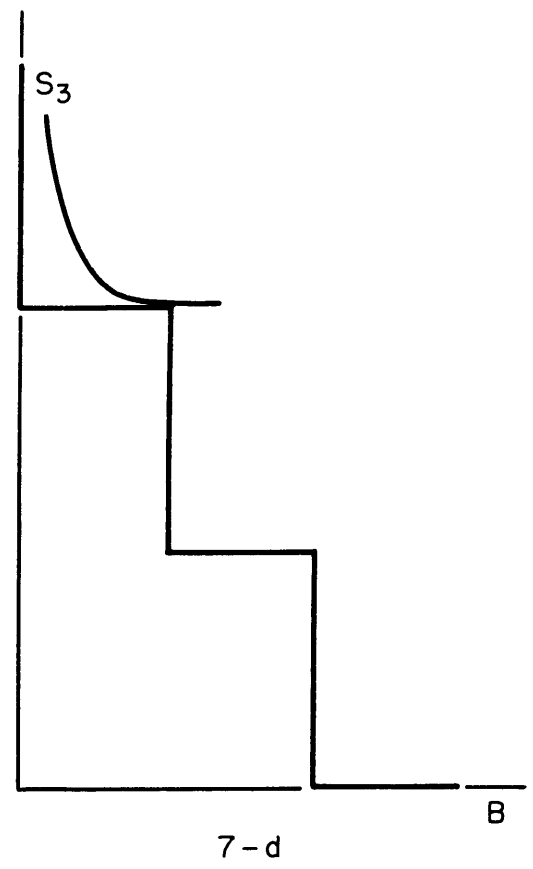
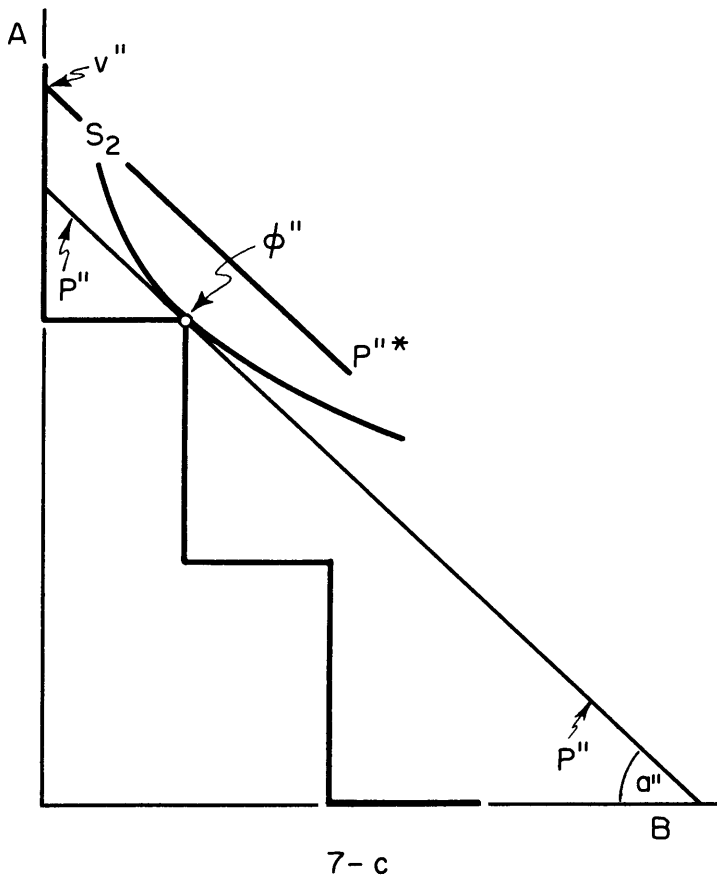
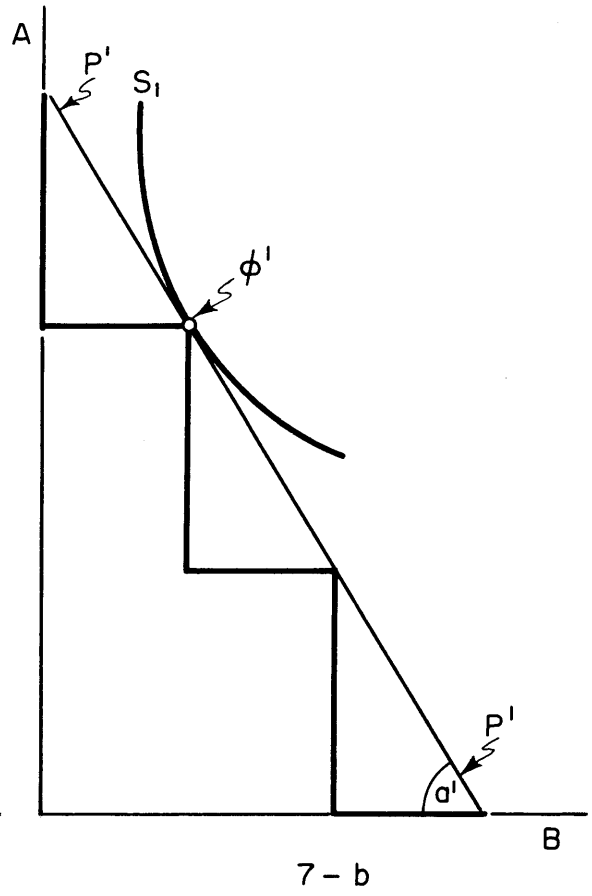
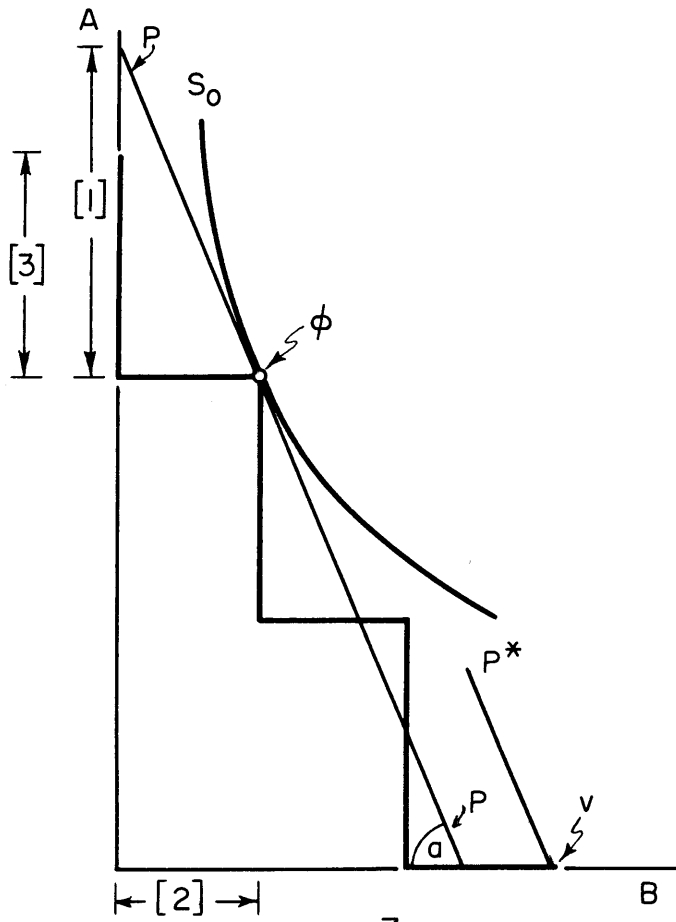


FIG. 5





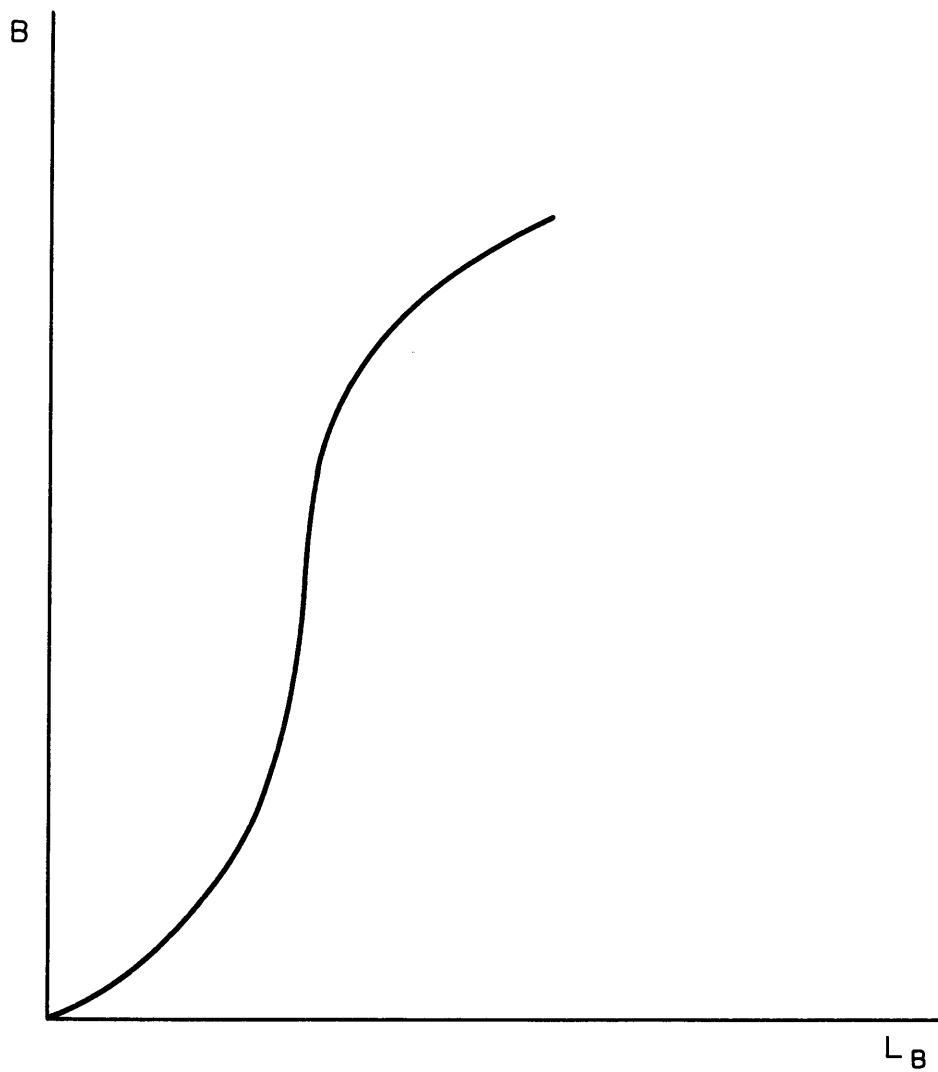
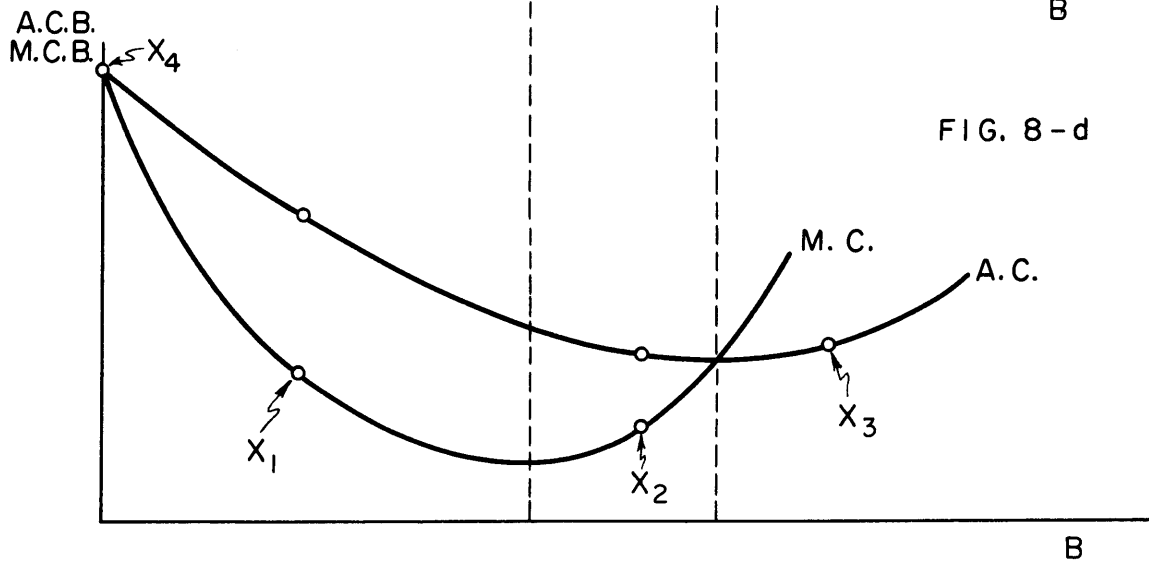
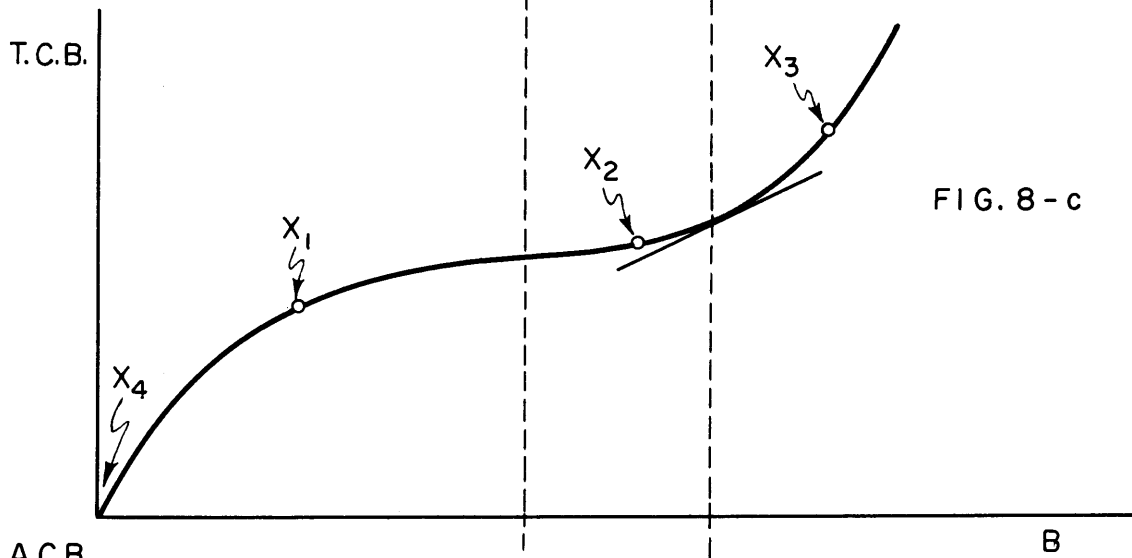
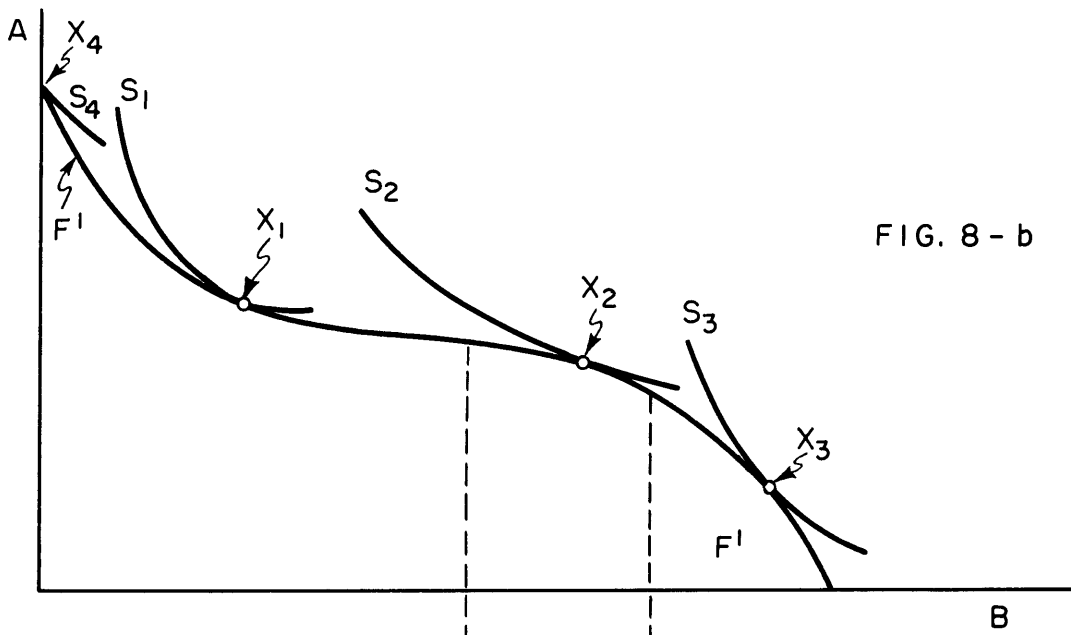


FIG. 8-a



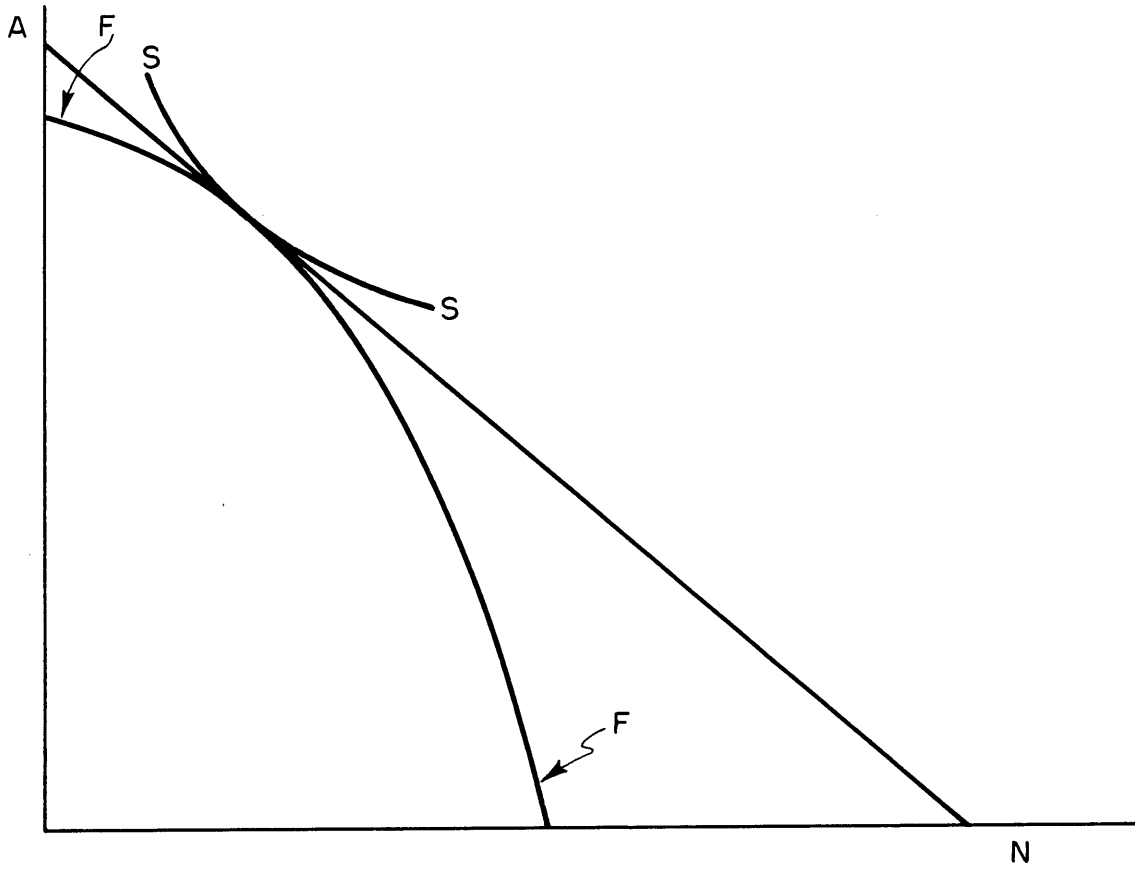


FIG. 9

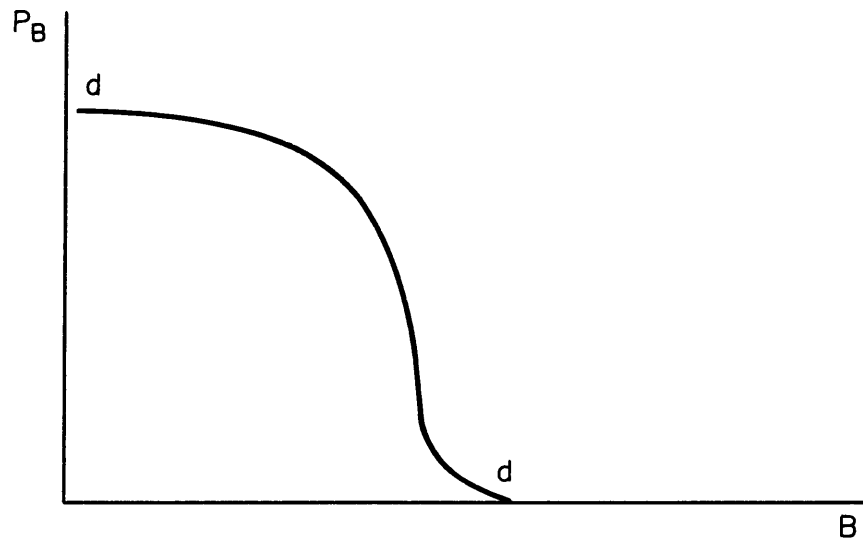


FIG. 10

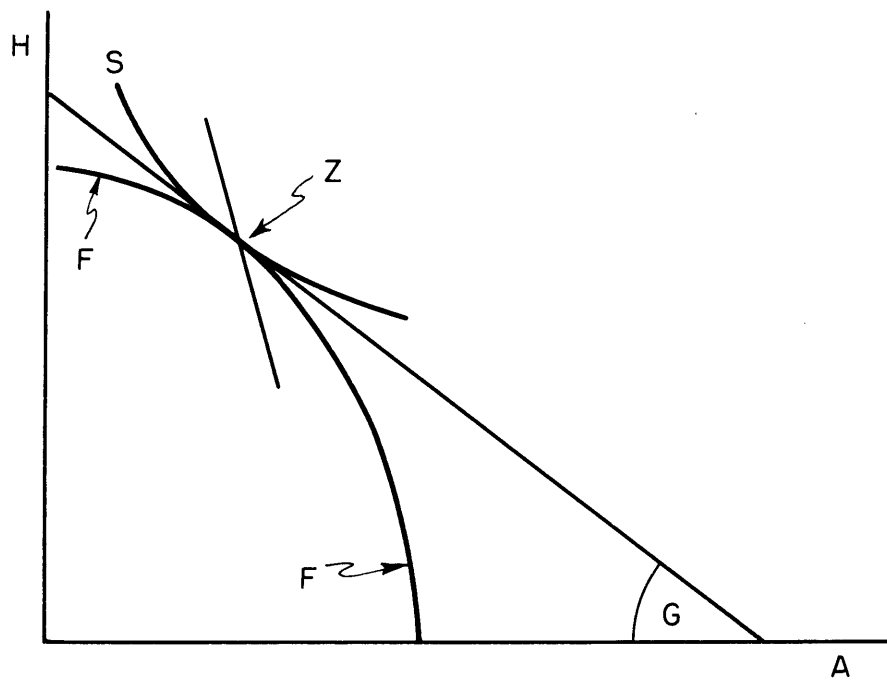


FIG. 11

CHAPTER THREE

ON CAPITAL PRODUCTIVITY, INPUT ALLOCATION AND GROWTH

ON CAPITAL PRODUCTIVITY, INPUT ALLOCATION AND GROWTH

Theorizing about economic growth is often based on concepts which received their development in the pure theory of production. This is right and fruitful. Modern allocation theory is a powerful engine of analysis and we have barely begun to harness it to gain insight into economic growth. But its use imposes high requirements of rigor and precision. Imprecise application can badly misguide the imaginative investigator, lead to flaws of analysis, and bad prescription.

As a case in point I should like, in Parts I and II, to examine some often misunderstood aspects of the relationship between the interest rate, the marginal productivity of capital and a more recently popular member of the economist's set of concepts, the capital-output ratio.¹ Then, in Part III, a frequent misconception about optimal input combinations is subjected to scrutiny. The specific issue concerns the desirability of capital intensive processes in countries whose capital is scarce relative to labor. Part IV consists in some general comments on input allocation as a maximizing problem and on the strategy of theorizing about economic growth.

I

I should like, first, to consider in context of a simple aggregate model the relationship between the aggregate "capital-output ratio", k

¹ As a matter of fact, it too has been around for some time in the guise of the average productivity of capital. But it took modern dynamics to turn it upside down and give it star billing.

and "the" interest rate, i . It is tempting to base historical analysis of economic growth on a presumption that, bar innovation, a declining interest rate must be associated with an upward tendency in the capital-output ratio. Such a presumed linkage between i and k provides the basis, for example, of a recent suggestion that during the past eighty years innovations in the United States and the United Kingdom must, on balance, have tended to reduce the capital-output ratio.¹ U.S. and U.K. time series covering these eighty years appear to indicate a secular decline in interest rates, both absolute and relative to real wages. Without innovation, so the reasoning goes, such a decline in interest rates would have been necessarily associated with a rising capital-output ratio. But as a matter of history, k during this period appears to have remained more or less stable. It follows, according to this line of thought, that in U.S.-U.K. experience innovations must have been "capital saving" in some net sense; otherwise falling interest rates would have resulted in a rising capital-output ratio.²

But would they, as a matter of logical necessity? Does a declining interest rate necessarily imply an upward tendency in the capital-output ratio?

I think not. Only in the "knife edge" case of perfectly constant returns to proportional variations in labor and capital -- in the case, that is, of homogeneous production functions of first degree -- will

¹ By Prof. Henry J. Bruton, in "Growth Models and Under-Developed Economies", Journal of Political Economy, August, 1955. Prof. W. Fellner also has done work along these lines. He, however, invokes income distributional considerations as well as variations in i and k (see Part II, below).

² Never mind, here, how much weight the evidence will bear. Incidentally, I am using the term "capital saving" to denote any innovation which will tend to reduce k . No particular relation to the more usual definitions found in the theory of production literature is implied.

variations in i and k be necessarily of opposite direction. Should there exist inherent non-linearities in the physics and topography of the universe, or some "hidden" third input in finite-elastic supply which does not directly enter the production function (e.g. "land"), returns to scale would not be constant and a falling i would not necessarily imply a rising k . It is, in fact, easy to construct production functions of "normal" curvature: non-increasing returns to scale and isoquants convex to the origin, which contain stages consistent with a falling level of i and a constant or even declining k .

The most widely used growth model, that of Harrod and Domar, rather begs the question: it assumes a constant k .¹ I shall here adopt a set-up more neo-classical in spirit; it can be viewed as a thawed-out version of Harrod-Domar. Assume one commodity, Q , and two factors of production K and L . It will simplify matters if we posit that K is the same commodity as Q , and that once congealed into stock form it will last for ever (capital consumption is zero). A physical unit of K rents for r dollars per period; its price, as that of a unit of Q , is p dollars. We are given a production function: $\dot{Q} = F(K, L)$;² it exhibits isoquants that are convex to the origin and non-increasing

¹ Constancy of coefficients is, moreover, a crucial assumption of the Harrod-Domar model -- in the sense that it is an assumption to which the characteristic results are especially sensitive. On this formulation, and for a most elegant development of a Harrod-Domar type model with variable proportions, see R. M. Solow, "A Contribution to the Theory of Economic Growth", Quarterly Journal of Economics, February 1956.

² A dot over a variable indicates that it is a flow magnitude, measured as a time-rate.

returns to scale. There is no need to introduce the time dimension directly -- we can simply posit that

$$\frac{d\dot{Q}}{dt} > 0, \text{ hence that } t_1 > t_0 \Leftrightarrow \dot{Q}_1 > \dot{Q}_0$$

Where does the interest rate come in? Assuming competitive "taking prices as given" behavior, atomistic profit-maximizing producers will bring the system into equilibrium at

$$p \frac{\partial \dot{Q}}{\partial K} = r \quad (1)$$

Capital, that is, will be used as an input up to the point where the value of its marginal product (in each period) just equals its rental for that period. Dividing through by the price of a unit of K, we get

$$\frac{\partial \dot{Q}}{\partial K} = \frac{r}{p} = \frac{\$ \text{ per apple per year}}{\$ \text{ per apple}} = \text{pure number/year}$$

But $\frac{r}{p}$, the rental of a dollar's worth of capital, is the apple-rate of interest which, in equilibrium and with constant price level, will just equal the money rate of interest.¹ Our (1), then, can be rewritten

$$\frac{\partial \dot{Q}}{\partial K} = i \quad (2)$$

¹ The general (equilibrium) relationship between the commodity own-rate of interest and the money-rate is given by $i = \frac{r}{p} + \frac{dp}{dt} \cdot \frac{1}{p}$

(cf. P.A. Samuelson, "Some Aspects of the Pure Theory of Capital", Q.J.E., May 1937, p. 471). Note, incidentally, that the assumption of non-exhaustible capital voids the need to distinguish between gross and net rental.

In competitive equilibrium, with constant price-level, it is the marginal product of capital that is brought into equality with the interest rate.¹

What about k , the capital-output ratio? It is immaterial, here, whether by k we mean an average capital-output ratio: $\frac{K}{Q}$, or an incremental ratio $\frac{\Delta K}{\Delta Q}$.² Whether average or incremental, k is not in any sense equivalent to the reciprocal of the marginal physical productivity of capital, $\frac{\partial Q}{\partial K}$. It is related, rather, to the average productivity of capital, $\frac{Q}{K}$ -- one is reciprocal to the other.³

The distinction is crucial. $\frac{\partial Q}{\partial K}$ measures the response in Q due to a variation of K alone, with all other inputs held constant.

$\frac{\Delta K}{\Delta Q}$, on the other hand, denotes movement along an expansion path, along which all other inputs (L) are optimally adjusted.

¹ Remember, in our set-up "output" and "capital" are measured in the same physical units; that is why there is no difficulty with equating the marginal physical product of K with the interest rate. "Apples" cancel out, and $\partial Q / \partial K$ is an own-rate of interest.

² In the Harrod-Domar context the two are the same: isoquants are L-shaped and equidistant along a ridge line that connects the vertices and the origin. But it is the essence of our problem that no such strict technological constancy can be expected to hold. For most purposes $\Delta K / \Delta Q$ is more important: the k of the modified Harrod-Domar equilibrium growth condition is incremental, and most of the U.S.-U.K. evidence relates time series of capital formation to changes in the level of output.

³ Assume that all capital capacity is in use, hence that we need not distinguish between the average productivity of the whole capital stock in existence and that of the capital stock actually in use.

The problem of $i - k$ variation along a given production function can now be posed as follows: does a falling interest rate,

$$\frac{di}{dt} = \frac{d\left(\frac{\partial \dot{Q}}{\partial K}\right)}{dt} \equiv \frac{d\left(\frac{\partial \dot{Q}}{\partial K}\right)}{d\dot{Q}} \cdot \frac{d\dot{Q}}{dt} < 0$$

necessarily imply that $\frac{K}{\dot{Q}}$ (or $\frac{\Delta K}{\Delta \dot{Q}}$) is rising, i.e. that the average productivity of capital, $\frac{1}{k}$, is falling over time? The answer is, I think, clear: neither the average nor the incremental capital-output ratio need increase over time just because $\frac{\partial \dot{Q}}{\partial K} = i$ is falling.

For proof we require only that there exist a production function of "well-behaved" convexity that does permit both

$$\frac{d\left(\frac{\partial \dot{Q}}{\partial K}\right)}{d\dot{Q}} < 0^1 \quad \text{and} \quad \frac{d\left(\frac{K}{\dot{Q}}\right)}{d\dot{Q}} \neq 0$$

(I shall use a case where $\frac{K}{\dot{Q}}$ remains constant hence equal to $\frac{\Delta K}{\Delta \dot{Q}}$

thus voiding the average-incremental distinction.)

Take the following very simple function

$$Q = \sqrt{K} + \dot{L} \quad (3)$$

It implies a family of isoquants of the usual curvature (see Fig. 1) and decreasing returns to scale along all $\frac{K}{\dot{L}} = \text{constant}$ paths. (Other, that is, than that which hugs the \dot{L} -axis. This last shows constant

¹ Since if $\frac{d\left(\frac{\partial \dot{Q}}{\partial K}\right)}{d\dot{Q}} > 0$, it follows that $\frac{d\left(\frac{\partial \dot{Q}}{\partial K}\right)}{d\dot{Q}} \cdot \frac{d\dot{Q}}{dt} \equiv \frac{d\left(\frac{\partial \dot{Q}}{\partial K}\right)}{dt} > 0$, because $\frac{d\dot{Q}}{dt} > 0$.

returns: for mathematical convenience, and since it alters nothing, I made the exponent of L be one.)¹

It is required to show that $\frac{d(\frac{\partial \dot{Q}}{\partial K})}{d \dot{Q}} < 0$ for values of \dot{Q} and K such that $\frac{K}{\dot{Q}}$ is a constant.

Differentiate (3) partially with respect to K

$$\frac{\partial \dot{Q}}{\partial K} = \frac{\partial(\sqrt{K} + L)}{\partial K} = \frac{1}{2\sqrt{K}} \quad (4)$$

Let

$$\frac{K}{\dot{Q}} = c > 0 \quad (5)$$

and substitute for K from (5) into (4):

$$\frac{\partial \dot{Q}}{\partial K} = \frac{1}{2\sqrt{c\dot{Q}}} \quad (6)$$

Now differentiate (6) with respect to \dot{Q} ,

$$\frac{d(\frac{\partial \dot{Q}}{\partial K})}{d \dot{Q}} = -\frac{1}{4\sqrt{c}\dot{Q}^{3/2}} < 0, \text{ for } \frac{K}{\dot{Q}} = c > 0 \quad (7)$$

This essentially completes the proof. We have in hand a production function of normal curvature such that along any path of capital accumulation characterized by a constant capital - output ratio the marginal productivity of capital -- the shadow rate of interest -- will fall.²

1 This makes $\frac{\partial \dot{Q}}{\partial L} = 1$ and invariant to K , though, of course, $\frac{\dot{Q}}{L}$ increases with K . If anyone is concerned, let him work out the proof for $\dot{Q} = \sqrt{K} + L$ 0.9999....; nothing of essence for my argument would change.

2 We have shown that a non-increasing $\frac{K}{\dot{Q}}$ is consistent with a falling i . Is it also consistent with a falling i (real wage rate, w)? $\frac{\partial \dot{Q}}{\partial L} = 1$, therefore $\frac{i}{w} = \frac{\partial \dot{Q}}{\partial K} / \frac{\partial \dot{Q}}{\partial L} = \frac{\partial \dot{Q}}{\partial K} = i$; $\frac{i}{w}$ and i will move the same way, hence $\frac{i}{w}$ is also falling.

Figures 1 and 2 "demonstrate" graphically the possibility of such a configuration. Figure 1 is a plot of the projections of $\dot{Q} = \sqrt{K} + \dot{L}$ on the K- \dot{L} plane; Figure 2 shows how \dot{Q} varies with K, at various fixed levels of \dot{L} . In Fig. 2 each of the K- \dot{Q} curves represents a vertical slice of the same production surface cut at right-angles to the \dot{L} -axis at $\dot{L} = 0, 1, 2, 3, \dots$ respectively. Innovation, i.e. a shift of the production function, is not implied.

Now draw any straight line through the origin in Fig. 2: any point on the line will have a constant capital-output ratio equal to the reciprocal of the slope of the line. A ray with slope of $\frac{1}{2}$, for example, is a locus of points along which $\frac{\dot{Q}}{K} = \frac{1}{2}$, hence $\frac{K}{\dot{Q}} = 2$. As we move along such a ray of constant $\frac{K}{\dot{Q}}$ in the direction of increasing \dot{Q} and K, the slopes of successive K- \dot{Q} curves decline. Again we have $\frac{K}{\dot{Q}} = \text{constant}$ with $\frac{\partial \dot{Q}}{\partial K}$ falling.

It is easy to work out some numerical examples. Taking the case of $\frac{K}{\dot{Q}} = 2$, we can substitute $\dot{Q} = \frac{1}{2}K$ into the production function (3) and get

$$\dot{L} = \frac{1}{2}K - \sqrt{K}, \quad (\text{for } \frac{K}{\dot{Q}} = 2) \quad (8)$$

This, together with the production function (3), $\dot{Q} = \sqrt{K} + \dot{L}$, and the expression for the marginal productivity of capital (4),

$\frac{\partial \dot{Q}}{\partial K} = \frac{1}{2\sqrt{K}}$, can be used to evaluate \dot{L} , \dot{Q} , and $\frac{\partial \dot{Q}}{\partial K}$ for various specified amounts of K. Columns 1-5 of Table 1 summarize the calculations for a capital-output ratio of two.

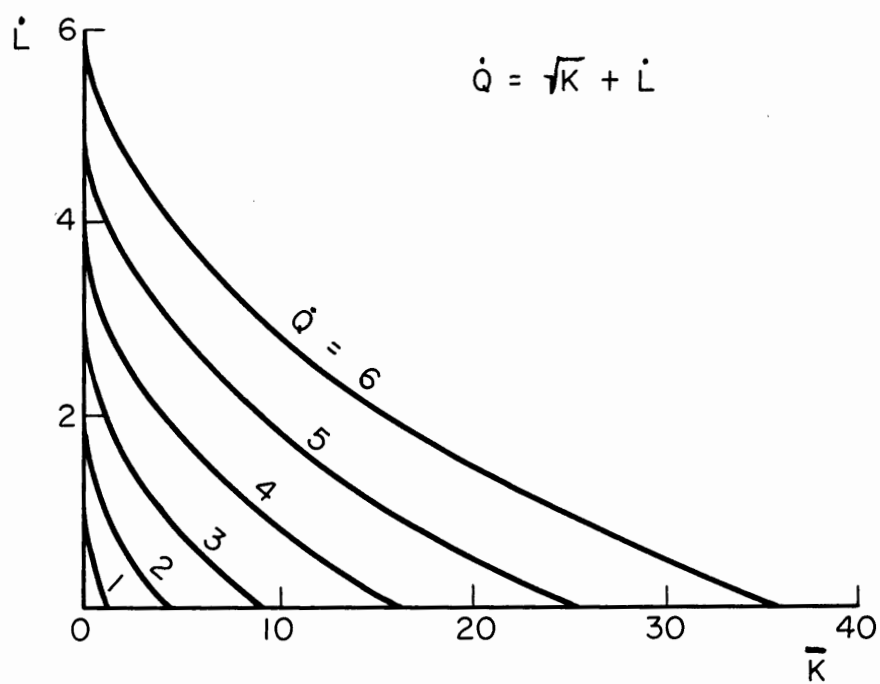


Fig. 1

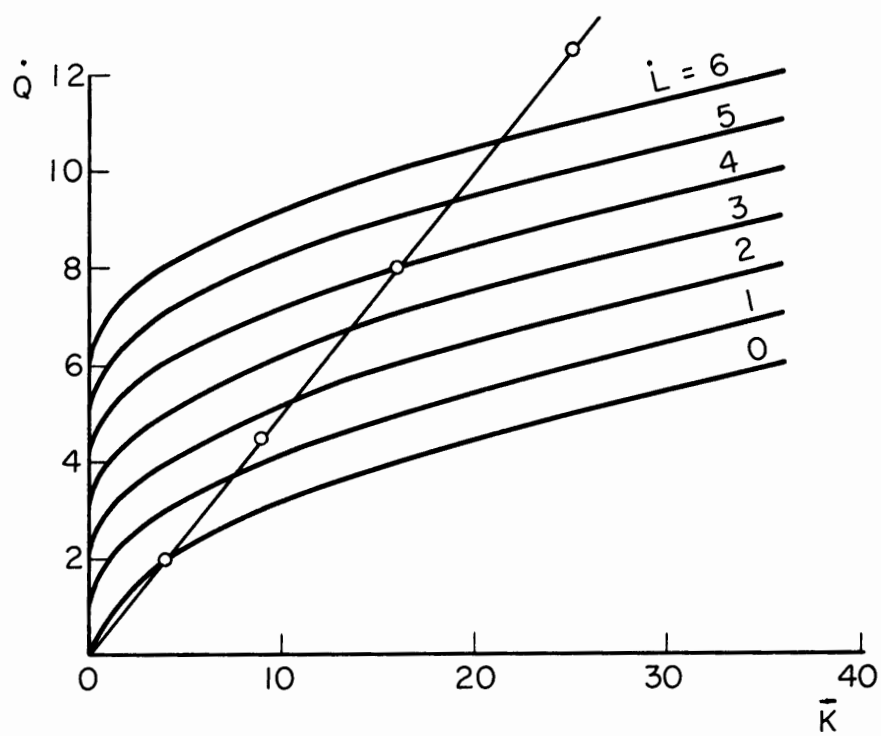


Fig. 2

1	2	3	4	5	6	7
\bar{K}	\dot{L}	\dot{Q}	$\partial \dot{Q} / \partial K$	K/\dot{Q}	$\frac{\partial \dot{Q}}{\partial K} \cdot \frac{K}{\dot{Q}}$	K/\dot{L}
0	0	0		2	-	-
4	0	2	1/4	2	1/2	-
9	1.5	4.5	1/6	2	1/3	6
16	4	8	1/8	2	1/4	4
25	7.5	12.5	1/10	2	1/5	3 ¹ / ₃
↓	↓	↓	↓	↓	↓	↓

Table 1

These points can be plotted in Figs. 1 and 2 (they are marked with \odot). They show, once again, a path of capital accumulation and labor force growth that results in a constant capital-output ratio and a declining (shadow) rate of interest.

To summarize: It has been shown that "well behaved" production functions do not require that a falling rate of interest be accompanied by a rise in the incremental or average capital-output ratio. This conclusion does depend, of course, on what one means by "well-behaved". If only homogeneous and first degree functions are admissible, then a constant $\frac{K}{\dot{Q}}$ does imply a constant i . Along such a function only proportions matter, and once you "freeze" one proportion: $\frac{K}{\dot{Q}}$, everything except scale is fixed. But if diminishing returns to proportional changes in \dot{L} and K are allowed, variation in i will not uniquely determine the behavior of k . Hence inferences based on a unique $\frac{di}{dt}$ versus $\frac{dk}{dt}$ relationship -- e.g., that in the U.S. and the U.K. innovations must have tended to reduce the capital-output

ratio rest on special assumptions of fact about the shape of "the" production function which require explicit articulation and defense.^{(1) (2)}

II

Is there any way to salvage the inference that in U.S.-U.K. experience innovations must have tended to reduce k ? Does a combination of a declining interest rate with a non-rising capital-output ratio imply anything about the world that could be tested against -- and perhaps refuted by -- observable fact?

The answer to the second question, at least, is yes. There is a definitional relationship between the marginal productivity of capital, its average productivity $\frac{1}{k}$, and the share of output imputed to the owners of capital, that is directly relevant. The theorem states that (1) a falling marginal productivity of capital is consistent with (2) a non-rising capital-output ratio only if (3) the income-share of

¹ Unfortunately for his argument, Bruton's paper (op. cit.) fails to posit constant returns to scale -- he explicitly introduces diminishing returns to scale based on finite-elastic third inputs. It has been suggested to me, on the basis of a footnote reference, that Bruton had in mind a Cobb-Douglas type function with diminishing returns against a hidden fixed input. Such a function, with its unitary elasticities of substitution, would also save the logic of the argument. But dependence on such a very special function does take some of the sweep out of the inference.

² "My" type of production function implies that the implicit "third" input -- whose finite-elasticity is responsible for the decreasing of returns to proportional increases of L and K -- becomes increasingly better suited for combination with L than with K as the scale of production expands. I should not care to defend the plausibility of the hypothesis that such phenomena are wide-spread, though they may not be wholly unreal: easily accessible deposits can be machine mined but the tough marginal shafts need labor. (What of the increase in goods-in-process and skill requirements though?) But my sense of what is "likely" for the economy as a whole does not function very briskly in context of a model based on a "global" production function.

capital is falling. There is a simple proof of this proposition in Professor Fellner's new book:¹

$$\frac{\partial \dot{Q}}{\partial K} = \left(\frac{\partial \dot{Q}}{\partial K} \right) \left(\frac{K}{K} \right) \left(\frac{\dot{Q}}{\dot{Q}} \right) = \left(\frac{\partial \dot{Q}}{\partial K} \cdot \frac{K}{\dot{Q}} \right) \left(\frac{\dot{Q}}{K} \right) \quad (9)$$

If $\frac{\partial \dot{Q}}{\partial K}$ is falling, by identity $\left(\frac{\partial \dot{Q}}{\partial K} \cdot \frac{K}{\dot{Q}} \right) \left(\frac{\dot{Q}}{K} \right)$ must also be falling. If then, $\frac{\dot{Q}}{K}$ is not falling, $\left(\frac{\partial \dot{Q}}{\partial K} \cdot \frac{K}{\dot{Q}} \right)$ must be getting smaller. But this last is the share of capital income $\left(\frac{\partial \dot{Q}}{\partial K} \cdot K \right)$ in total output, \dot{Q} .

Does this help the innovation-inference? It depends on one's interpretation of the facts of economic history.² If these facts show that the "secular" share of capital-owners in total income has not declined during a period of falling interest rates, and if one postulates income-imputation "as if" by perfect markets, then it follows that the non-rising of $\frac{K}{\dot{Q}}$ implies off-setting innovations of the "capital saving" sort.

This is not the occasion to examine the evidence. In recent U.S. experience the share of "labor" income (wages, salaries, plus some imputed fraction of the income of unincorporated enterprises) appears to have remained more or less stable. In a two factor world this would imply constancy in capital's share. But whatever the data may show, this is a direction of research that might validate some historical

¹ Trends and Cycles in Economic Activity, Holt, 1956, (p. 122). I had heard Professor Fellner use just this relationship in specifying the qualitative properties that must characterize innovations if these are to off-set diminishing marginal returns to capital in a world where K is growing faster than the cooperating factors. On leafing through his new book, I hit upon this proof of the proposition. Note, incidentally, that in Column 6 of Table 1, $\frac{\partial \dot{Q}}{\partial K} \cdot \frac{K}{\dot{Q}}$ is in fact falling.

² And one's acceptance of the model. For evidence, see some recent work by Phelps-Brown & Hart and Phelps-Brown & Weber in the June 1952 and June 1953 issues of the Economic Journal.

inference about innovations.¹

III

The crucial analytical distinction emphasized in all the above is that between a marginal notion and an average; specifically, that between the marginal productivity of capital and its "adjusted" average productivity.

Confusion about an analogous distinction between various labor-productivity notions and their relation to per capita output (income)² has marred some of the recent literature on input and investment allocation in economic growth. The result has been some very dubious advice about how resources should be used in countries eager to grow quickly.

As a case in point, I should like to examine some of the arguments that have recently been advanced in the literature in favor of capital-intensive (high $\frac{K}{L}$) processes of production in capital scarce - labor rich countries. The most comprehensive defense of this view -- a true

¹ Implications for $\frac{K}{L}$ variations might provide a related avenue of exploration (cf. Column 7 of Table 1). For more comprehensive analysis it would pay to make the model explicitly dynamic and solve for the time-paths of the variables under alternative assumptions about production functions. R.M. Solow's aforementioned essay provides the machinery. Unfortunately, the price-interest implications become problematical in non-constant returns to scale situations.

² In our one-commodity closed model consisting exclusively of real magnitudes income is identically equal to output.

defense in depth -- is to be found in a recent article by Professors Galenson and Leibenstein.¹ But since they erect much of their case on presumed income-distributional effects on the volume of saving and on neo-Malthusian sensitivity of population growth rates to the investment pattern, and relatively little hangs on confusions "in the small" between various productivity notions, I should prefer to examine a passage from the above cited Bruton article that highlights the difficulties.² Part IV is addressed to the larger issues raised by Galenson and Leibenstein.

The following implies both a paradox and its resolution:

"Now what is the relevance of all this for the underdeveloped countries? Since their output is very low relative to that of the more highly developed countries, their chief objective is to increase the capacity output of their economies as rapidly as possible. To do this, it is evidently desirable to have capital as productive as possible - that is, K/O as low as possible. [1] To the extent that the rate of interest reflects the shortage of capital relative to labor, the low K/O will obtain. [2] But this means low productivity per worker and hence low real per capita income. [3] Therefore, to increase per capita income, we must increase the proportion of capital in the input-mix; we must move around the equal-product curve, using more capital and less labor to achieve the same output" (p. 327)³

¹ "Investment Criteria, Productivity and Economic Development", Quarterly Journal of Economics, August, 1955.

² It should be said that the bulk of Bruton's article has little to do with input allocation and that the quoted passage, as I interpret it, is in the nature of a dictum. It is a convenient point of departure only because, unlike in many recent examples in the literature, difficulties are not embedded in impregnable rhetoric.

³ The sentence numbering is mine.

I believe that the "paradox" -- that it is both desirable to have a low $\frac{K}{Q}$ and undesirable because it implies low output per head -- is illusory; further, that the suggested resolution [3] is unacceptable.

The nub of the difficulty, one suspects, lies in a confusion between (1) the marginal productivity of labor, $\frac{\partial Q}{\partial L}$ (2) the average productivity of employed labor, $\frac{Q}{L_e}$ and (3) income per head of population, $\frac{Q}{P}$.¹ But whatever the cause, the resulting prescription is so very misleading, and in more or less diffuse form sufficiently common, to warrant examination of the critical numbered sentences.

[1] It is not true that a high rate of interest due to the shortage of K relative to \dot{L} necessarily implies a low $\frac{K}{Q}$. Take an extreme case of scarce K relative to \dot{L} : a situation such as in

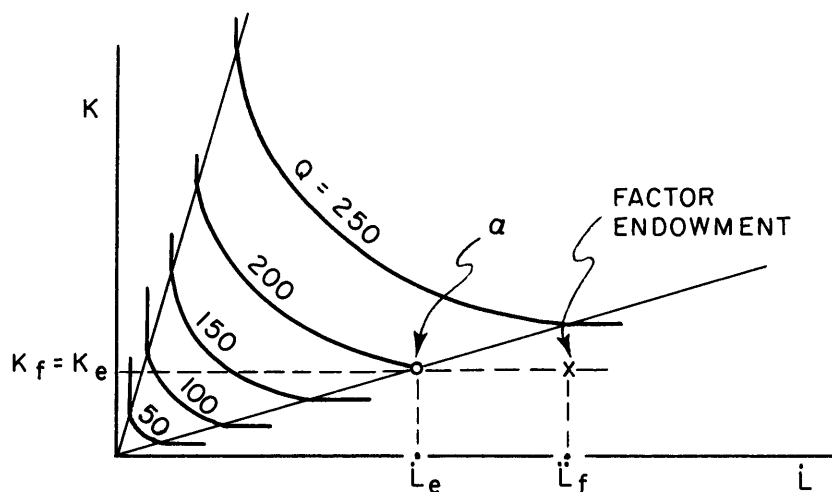


Fig. 3

¹ The "average productivity" of the potential, full employment, labor force may or may not be assumed proportional to $\frac{Q}{P}$.

Fig. 3, where labor is technically redundant.¹ In such a situation the maximization of output requires that production take place at point ∞ , where the marginal productivity of \bar{L} is zero and all output is imputed to the non-redundant input, capital.² But nothing in this set-up requires that this output be very large relative to K . In fact, in this polar situation, a "high" interest rate, e.g. 25%, implies a not so "very low" $\frac{K}{Q}$ of four.

[2] The notion that low $\frac{K}{Q}$ "means low productivity per worker and hence low real per capita income" is clearly in error. With given K , the lower $\frac{K}{Q}$ the greater will Q be, hence the greater $\frac{Q}{P}$. It is true that if, as in Fig. 3, maximization of output requires input proportions which reduce the marginal productivity of labor to zero, and if, also, income distribution is determined by, or "as if" by, a system of perfect markets, then all output will be imputed to the "owners" of capital.³ But just this may be required if per capita output is to be maximized.

¹ i.e. The factor endowment point in the \bar{L} - K space lies outside the labor intensive ridge line.

² At least in the neo-classical case where the isoquants flatten out smoothly. If in Fig. 3 the isoquants consisted of straight-line segments with sharp corners, the wage-rate \dagger capital-rent ratio at ∞ would, within a range, be indeterminate. This would provide some flexibility for non-distorting use of minimum wage laws, etc.

³ The words were designed to retain antiseptic independence of institutional context. The notion of this kind of imputation is neutral as between (1) genuine profit maximizers acting in "real" but perfectly competitive markets; (2) Lange-Lerner type bureaucrats ("take prices as given and maximize, or Siberia"); or, (3) technicians using electronic machines and trying to devise efficient computing routines.

That the resultant income distribution may be so "bad" in terms of some relevant values as to call for redistributive arrangements, even at the cost of some reduction in output,¹ is, in this context, irrelevant. Low $\frac{K}{Q}$ -- and to the extent that factor endowments and the production function require it, low $\frac{\partial Q}{\partial L}$ -- will make for larger rather than smaller Q or $\frac{Q}{P}$.

[3] It follows from the above that the prescription of sentence [3]: in order to raise per capita income increase "the proportion of capital in the input mix" and "move around the equal product curve, using more capital and less labor to achieve the same output" is very poor advice indeed. If the purpose is to maximize output (or output per given head of population) then the worst thing to do is to increase $\frac{K}{L}$ in the input mix, at least as long as L remains redundant. Until all L is productively employed you want to hug that labor intensive ridge line though it implies zero marginal productivity of labor. Even once out of the region of redundancy -- by way of capital accumulation sufficient to off-set population growth -- attempts to "pick" input proportions so as to achieve a higher marginal product of labor are not likely to put a system on a Paretian road to "bliss". If your interest is in output,² maximize it and let $\frac{\partial Q}{\partial L}$ turn out what it may.

¹ Due to the disincentive effects of tax plus transfer payment arrangements; or, if redistribution were achieved by say minimum wage legislation, to misallocation of inputs. (In Fig. 3, a positive wage rate would push profit maximizing producers "west", away from ∞ and to a position of reduced output.)

² The one-commodity assumption saves many difficulties: it makes solid observable fact of the notion of "real" income.

All this can be put more positively. Assuming:

- (a) That the rate of growth of population is insensitive to the choice of processes;
- (b) That the potential "full employment" labor force (measured as a time-flow: \dot{L}_f) is a function of population (P);
- (c) That the rate of saving is independent of the (as if) market imputed distribution of income; and

retaining the two input - one output zero capital-consumption model

to avoid output-mix problems, it follows that:

- (1) Maximization of \dot{Q} , or, equivalently of $\frac{\dot{Q}}{P}$ or $\frac{\dot{Q}}{\dot{L}_f}$, does not imply maximization of $\frac{\partial \dot{Q}}{\partial L}$, nor of the average productivity of the labor force actually employed: $\frac{\dot{Q}}{L_e}$.¹
- (2) There is no conflict between maximizing $\frac{\dot{Q}}{P}$ in any short period and maximizing the rate of growth of $\frac{\dot{Q}}{P}$ subject to the production function, input endowments and the rate of saving. On the contrary, current output maximization is a necessary condition for maximum output in any future period.
- (3) Further, if factor endowments lie outside the labor ridge-line, maximization of \dot{Q} , (\dot{Q}/P , \dot{Q}/L_f) does imply:²

¹ Never mind here the difficulties of deciding in a world of under-employment how many hours of whose time are or are not "employed". The important thing to note is that \dot{L}_e is a flow rate of labor services and not the "stock" of labor.

² My colleague, Prof. R.S. Eckaus, has made a searching analysis of this case: "Factor Proportions in Underdeveloped Areas", A.E.R. Sept. 1955. I happen to think that the redundant labor hypothesis in its boldest formulation is implausible: there are many "goods" in the world and international trade is not illegal. But with demand considerations thrown in, a strong case can be made that it contains an important germ of truth. Anyway, the empirical plausibility of the hypothesis is irrelevant to the issues here at stake.

- (i) Adoption of the most labor intensive processes and the spreading of K over as many laborers as give $\frac{\partial \dot{Q}}{\partial L} > 0$;¹
- (ii) $\frac{\partial \dot{Q}}{\partial L} = 0$, except for an indeterminacy if the production function is sharp-cornered (cf. footnote p. 183;)
- (iii) Maximization of the average productivity of K: $\frac{\dot{Q}}{K}$ and, if the production function shows constant returns to scale, maximization also of $\frac{\partial \dot{Q}}{\partial K}$;
- (iv) That $\frac{\dot{Q}}{L_e}$ may be equal, greater, or less than what it would be for sub-optimal techniques which combine the given "amount" of capital with less labor; moreover, this ratio is irrelevant to the choice of techniques.
- (4) Finally, if factor endowments lie inside the "cone" of ridge lines and there is no technological redundancy of labor, the maximization of \dot{Q} , $\frac{\dot{Q}}{P}$, or $\frac{\dot{Q}}{L_f}$ implies:

¹ Once we drop the one commodity assumption this has to be qualified: output mix considerations, hence demand, become relevant. If (1) the most labor intensive technique produces only apples, and (2) people want overcoats as well as apples no matter how expensive overcoats become relative to apples; further, if (3) for whatever reason, such as sharply decreasing returns to scale, comparative advantage does not permit complete specialization hence not all overcoat demand can be met by exporting apples and importing overcoats, then, and only then, some part of the capital stock will have to be used to produce overcoats. If so, fewer people will be employed than if only apples were produced. The point is that the production-possibility surface may well exhibit ranges of full employment of all factors as well as ranges of redundancy, and demand and trade conditions may place the system on a point in the latter. But the qualitative conclusion remains: for any given output-mix the more labor intensive and capital saving the technique the better.

- (i) That the maximum output configuration will require employment of the full labor force: $\dot{L}_e = \dot{L}_f$. In a multi-product world this again requires qualification due to demand-mix considerations (see footnote on p.186).
- (ii) That neither $\frac{\dot{Q}}{\dot{K}_e}$ nor $\frac{\dot{Q}}{\dot{L}_e}$ will be at a maximum -- but see qualification under (i) above -- though, of course, $\frac{\dot{Q}}{\dot{K}_f}$ and $\frac{\dot{Q}}{\dot{L}_f}$ will both be at their highest values. Decisions based on economizing one input alone will result in waste. Both inputs are scarce and process choice that ignores the opportunity cost of either the labor or the capital input will fall short of the best possible. Neither the labor theory of value nor the more recently fashionable "capital theory of value" ² are defensible.

¹ By the definition of this case, i.e. that $\frac{\partial \dot{Q}}{\partial K} > 0$, $\frac{\partial \dot{Q}}{\partial L} > 0$ -- and at least in a non-increasing returns to scale world -- neither $\frac{\dot{Q}}{\dot{K}_e}$ nor $\frac{\dot{Q}}{\dot{L}_e}$ can be at their maxima. If $\frac{\dot{Q}}{\dot{L}_e}$ were, it would be just equal to $\frac{\partial \dot{Q}}{\partial L}$, hence $\frac{\partial \dot{Q}}{\partial L} \cdot L_e$ would exhaust all \dot{Q} and $\frac{\partial \dot{Q}}{\partial K}$ would have to be zero (for $K > 0$).

² I owe the suggestion of this symmetry to my colleague, Prof. P.N. Rosenstein-Rodan.

IV

All the above rests on the premise that output is indeed the proper maximand.¹ This is not the occasion to explore what kinds of restrictions on Social Welfare Functions validate the proposition that maximizing "output" -- or, in an n-commodity world, being on a Pareto-efficient production-possibility surface -- is a necessary condition for a maximum of any such W-function. Suffice it to assert that the normative content of Pareto efficiency is loaded with value. More on this later -- here let me conclude with only a few comments.

1) In all the above I made uncomplaining use of the notion of a capital-output ratio. A warning to the innocent is in order -- and judging from the development literature innocence on this issue is a surprisingly prevalent state. There are few concepts in current use in analytical economics whose conceptual foundations are quite so

¹ As long as population is independent of input-mix, etc., maximization of output is equivalent to maximization of output per head. A more complete model, however, might well include feedback effects from the pattern of production, income distribution, and the like, on the rate of population growth. It is an important theme of the above cited Galenson-Leibenstein article that just such feedbacks are likely to render maximization of current output (via labor intensive processes) inconsistent with maximum future output per head. I find the empirical evidence on which the many links in their chain of reasoning rest a tenuous basis for giving up current output for fear of the birth rate, but certainly would not deny the logical possibility of their hypothesis. The remainder of my discussion assumes a rate of population growth that is independent of choice of production processes.

shaky.¹ A few difficulties can be dealt with by being explicit and rigorous: e.g. the question of how "net", how "gross". Then again, all the difficulties can be assumed away; the model used above does in effect do so. But it is important to realize quite how much is being assumed away. To most of the problems treated in the theory of capital the very notion of a capital output or capital-capacity ratio is completely inapplicable.² A central task of capital theory is to provide a device whereby we can make commensurate outputs and inputs that are not simultaneous in time. The capital-output ratio, which relates increments in "the" capital stock to increments in the flow

¹ Never mind here the often related difficulties of measurement. It might be asked why a capital-output, i.e. an average-product-of-capital notion should give more trouble than the notion of an average product of labor. Quite apart from the fact that it too gives trouble -- vide Part III of this paper and the chronic presumption in even the professional literature that a rise in what is an average productivity of labor somehow provides market-validation for a rise in wage rates -- the $\frac{\dot{Q}}{K}$ relationship involves some fundamental difficulties that $\frac{\dot{Q}}{L}$ avoids. In $\frac{\Delta \dot{Q}}{\Delta L}$, ΔL is a flow, calculated in labor-hours. In $\frac{\Delta \dot{Q}}{\Delta K}$, ΔK is an increment of stock whose contribution must be judged not only in terms of the \dot{Q} it permits today, but also in terms of its effect on the \dot{Q} -s of the future.

² The distinction between capital-output and capital-capacity ratios is crucial in a world that is more like Keynes' than J.B. Say's. With a two year lag, the incremental capital-output ratio between 1929 and 1931 was a negative number. The distinction is related, incidentally, to the $\frac{\dot{Q}}{L_e}$ vs $\frac{\dot{Q}}{L_f}$ distinction of Part III; for capital it is more useful to attach subscripts to \dot{Q} .

rate of "capacity" output, simply buries this problem.¹ In general, the discounting machinery of capital theory is indispensable for all problems where the time-shape and length of the "investment" input and the output streams are significantly variable.

Having said this, I would nevertheless defend the view that for some purposes a capital-output ratio notion can yield useful qualitative and "order of magnitude" insight. But it takes a master-surgeon to operate with faulty instruments.

2) The one-commodity assumption of the Harrod-Domar model avoids many of the nasty index number type problems that cause trouble in a many-good world. The assumption that there is a homogeneous something called "capital" skirts all the complex issues that arise from the fact that for purposes of evaluating the production-possibilities of a community its capital stock is a "who is who of all goods in being".² Equally convenient, here "real output" is not many numbers denoted by a single number; it really is a single number, at least from a statical, one-period point of view. The Pareto-efficient locus of instantaneous or one-period production possibilities with given inelastically supplied

¹ More strictly, to the flow rate of capacity "value-added". For an aggregate model of a closed economy the distinction does not matter. But if one considers parts of the economy, it surely does (e.g. consider the $\frac{K}{Q}$ ratio of the mail-order house industry). The notion of capacity is itself a very slippery one. There is no a priori presumption that cost curves turn sharp corners and head straight "north", and if they do not, "capacity" output becomes a matter of price and profit maximization.

² On the impossibility, except under very strong assumptions, of finding a single number for "capital" that will, given L , uniquely denote the productive potential of the system, see R. M. Solow, "The Production Function and the Theory of Capital", Rev. of Ec. Studies, Vol. 23 #2.

inputs is not a multi-dimensional surface but just a point. All the complexities due to the simultaneous interdependence of "best" output composition with income distribution with input mix are effectively begged.¹

3) Once we step out of the comfortable confines of stationary statics, however, an analogous problem arises, even with one physical commodity. An apple today is not the same as an apple next week, and if we permit some of the output of each period to be congealed into capital stock for use in producing more output in future periods, the choice is no longer simply one of more versus fewer apples. Rather, we have to choose between various time-paths of apple-consumption.²

Is it possible that the most desirable time-path -- in terms of my values, or Mr. Nehru's, or some political consensus -- might require violation of the one-period rule that in any short period apple production be maximized? The answer is yes, but only under very strong quantitative and qualitative assumptions.³

If different income distributions result in "widely" different rates of saving, hence widely different rates of possible capital

¹ The "best" blend of food and clothes depends on the relative distribution of "income" (abstract purchasing power) between the Gourmets and the Dandys; but this, in turn, depends on the relative productivity of the inputs "owned" by the G-s and the D-s in producing the particular mix of food and clothing chosen, etc.

² Nothing hangs on the "unit period" notion. We could equally well think in terms of instantaneous flow rates. Analytically the distinction is that between difference and differential equations.

³ Assume here that these value systems possess regularities that make Pareto-efficiency necessary for a maximum. None of the argument concerns "queer" welfare functions. Assume, further, that the rate of population growth is independently given.

formation; and if, also, there are "severe" limits¹ on interference with the distribution of income as imputed by markets, or by electronically computed partial derivatives, then, indeed, it is logically possible that a less than maximum output on Monday night permit a sufficiently higher volume of investment to justify itself. "Justify" in what sense? In the sense that Monday's sacrifice of apple consumption, due both to saving for investment and to "inefficient" production, might be more than compensated by the resulting higher rate of possible apple production on Tuesday, Wednesday, and all future. Compensate whom? That person or entity whose W-function is to "count" .

Much remains to be said on this range of issues. Suffice it here to assert that:

(a) Even strong assumptions about the sensitivity of the rate of saving to income distribution and about feasibility-imposed plus value-imposed limitations on redistribution are not a sufficient cause for picking input proportions in terms of their effect on the various marginal productivities of factors. The "Galenson-Leibenstein thesis"² that high $\frac{K}{L}$ processes are desirable even where labor is redundant because the rate of saving, hence of investment, will vary inversely with

¹ These may be due to a complex mix of considerations involving both values and feasibility. How much of that very scarce resource, the efficient civil servant-administrator, is to be allocated to operating an internal revenue service? What feasibility or value limits are there on coercive processes and institutional arrangements? etc.

² Op. cit.

the fraction of income going to wage-earners, requires for its validation the quantitative presumption that:

- (i) A larger $\frac{\partial \dot{Q}}{\partial L}$ times a smaller \dot{L}_e will necessarily result in a lower wage bill than a smaller $\frac{\partial \dot{Q}}{\partial L}$ times a larger \dot{L}_e . This certainly may be implied by "the" production surface -- or, dropping the one output assumption, by the full solution of a Walrasian system with prescribed non-optimality built-in -- but it is surely not so ordained. In the case of labor redundancy, at least, where at the maximum output point $\frac{\partial \dot{Q}}{\partial L} = 0$, the presumption that the wage bill will be smaller if less-than-efficient high $\frac{K}{L}$ processes are adopted is clearly wrong. Zero is a mighty small absolute number.
- (ii) That a larger rate of saving times a lower income will necessarily result in a larger volume of saving, hence investment, than a smaller saving rate times a larger output. The contention that it is not aggregate output but rather "the rate of investment ... that determines the extent of capital accumulation and as a consequence the capacity of the economy to produce goods and services in the future" is misleading. If "rate of investment" stands for $\frac{K}{Q}$ -- as is implied by much of the Galenson-Leibenstein text -- then it is the product of output with the investment rate that matters.¹

¹ Galenson-Leibenstein, Op. cit. p. 351, par. 1. It should be said that the possibility that (i) and (ii) may work against them is acknowledged in four lines of a footnote attached at the end of several pages of argument in favor of their thesis. (f.n. 6, p. 368).

Whatever one's view about the joint likelihood of the above combination of qualitative and quantitative circumstances, it is difficult to accept as conclusive the Galenson-Leibenstein prescription, to-wit, that underdeveloped countries should "alter conditions to conform with our criterion by making labor scarce artificially ... by legislation establishing relatively high minimum wages and working conditions; by direct governmental control of manpower; or, in the case of state industry, by imposing high labor productivity targets upon management", and that "the islands of favored employment will have to be protected by the government ... for individual entrepreneurs will find it difficult to resist the constant temptation of cheap labor".¹

(b) In general, and in context of the traditional confines of the economist's formal maximizing set-ups, instantaneous Pareto-efficiency is a necessary condition for full-fledged dynamical, inter-temporal efficiency. For a one commodity model this means that to be on an "efficient" growth path -- on a path, that is, where no increase in the consumption of any good of any period is possible without a decrease in some other consumption -- one must maximize current output

¹ Op. cit. p. 368. My skepticism is not inconsistent with an awareness of the need, if a poor country is to "grow", to increase the volume of saving; of the many serious reasons why this may not be possible without the perpetuation of sharp inequalities in income distribution, notably if it is impossible to get at other than urban or institutional income, etc. But perhaps there are better ways than by allocating resources inefficiently and sacrificing valuable goods and services. Nor do I rule out the many political and sociological reasons which may justify capital intensive "monuments" such as will give people a sense of progress, etc. ... But their benefits should be measured against the loss of output, lower growth-rate, etc. ...

no matter what the consequences for the marginal productivity partial derivatives.¹

(4) With strong enough assumptions about the non-stability or non-transitivity of value systems, with uncertainty and imperfect knowledge, with non-convexities in the technology, with various institutional limitations, one can prove anything or nothing.² The question of what kinds of very real complications should be introduced into a formal maximizing set-up has answers only in terms of the strategy of theorizing, or on the level of planning strategy. The two input - one output model which provides the frame on which much of the above hangs is, I think, useful in the only way that very abstract pure theory is ever useful in economics. By way of strong but flexible and easily variable assumptions about the forms of the functional relations, the parameters and the data (e.g. labor redundancy), it permits exploration of the limits of what is qualitatively possible and consistent with maximizing. But its insights are, at best, suggestive and can not provide a basis for quantitative prescription. They can, with caution and "good sense", yield qualitative or order-of-magnitude checks on either detailed quantitative planning of allocation -- that is itself based on a blend of much

¹ This assertion rests on some path-breaking work, soon to be published, by Professors P. A. Samuelson and R. M. Solow of M.I.T. I should like to acknowledge the good fortune of having had access to their manuscript.

² For that matter, as Professor Samuelson once remarked, nothing precludes us from repealing the law of gravity or solving constrained maximum problems for a world where force is not even nearly equal to mass times acceleration.

larger and less flexible models and the planners' horse sense -- or on "planning" in terms of institutions, the fiscal-monetary impact, etc.¹

But one thing is clear. If it makes sense to use a formal maximizing set-up at all, whether for qualitative insight or quantitative guidance, its unreality can never justify violation of its inner logic and consistency. Bad logic is a poor way to compensate for the fact that the world is complicated.

¹ The simple statical neo-Keynesian spending model that runs in terms of aggregate income, consumption, investment, government expenditure, etc. is the closest thing to an exception. But its forte, too, is that it yields quantitative answers that have qualitative, directional significance. The Leontief system and its programming variants constitute the best example of the more inflexible large system that is required for direct quantitative planning. Its glory, that it permits "identification" of "structural constants" from observable transaction flows, is based on very strong and inflexible assumptions (fixed coefficients and constant returns to scale).

BIBLIOGRAPHY

BIBLIOGRAPHY

Books

- Arrow, Kenneth J. Social Choice and Individual Values. Cowles Commission No. 12. Chicago:
- Baumol, William J. Welfare Economics and the Theory of the State. The London School of Economics and Political Science. London: Longmans, Green and Co., 1952.
- Bishop, R. L. "Economic Theory" (MS., forthcoming).
- Cairncross, A. K. Home and Foreign Investment, 1870-1913. Cambridge: The University Press, 1953.
- Committee on Price Determination. Cost Behavior and Price Policy. New York: National Bureau of Economic Research, 1943.
- Creamer, Daniel. Capital and Output Trends in Manufacturing Industries, 1880-1948. Occasional Paper 41. New York: National Bureau of Economic Research, Inc., 1954.
- Dorfman, Solow, and Samuelson. Linear Programming and Economic Analysis. RAND Corporation, forthcoming.
- Fellner, William. Trends and Cycles in Economic Activity. New York: Henry Holt and Co., 1956.
- Fisher, Irving. The Theory of Interest. New York: The Macmillan Co., 1930.
- Frankel, S. H. The Economic Impact on Under-Developed Societies. Oxford: Basil Blackwell, 1953.
- Hamkin, George, et al. Measures for the Economic Development of Under-developed Countries. New York: United Nations, Department of Economic Affairs, 1951.
- Harrod, R. F. Towards a Dynamic Economics. London: Macmillan and Co., Ltd., 1949.

- Hayek, Friedrich von. Collectivist Economic Planning. London: G. Routledge and Sons, Ltd., 1935.
- _____. The Road to Serfdom. Chicago: University of Chicago Press, 1944.
- Koopmans, Tjalling C. (ed.). Activity Analysis of Production and Allocation. Cowles Commission Monograph 13. New York: Wiley and Sons, 1951.
- Lange, Oscar. Price Flexibility and Employment. Cowles Commission for Research in Economics. Bloomington: The Principia Press, Inc., 1944.
- Lange, Oscar, and Taylor, Fred M. On the Economic Theory of Socialism. Edited by Benjamin E. Lippincott. Minneapolis: University of Minnesota Press, 1938.
- Leontief, W. The Structure of the American Economy, 1919-1929. 2nd ed. New York: Oxford University Press, 1951.
- Leontief, W., et al. Studies in the Structure of the American Economy. New York: Oxford University Press, 1953.
- Lerner, A. P. The Economics of Control. New York: The Macmillan Company, 1947.
- Lewis, W. Arthur. The Theory of Economic Growth. London: George Allen and Urwin, Ltd., 1955.
- Little, I. M. D. A Critique of Welfare Economics. Oxford: Clarendon Press, 1950.
- Marshall, Alfred. Principles of Economics. 8th ed. revised. London: Macmillan and Co., Ltd., 1947.
- Nurkse, Ragnor. Problems of Capital Formation in Underdeveloped Countries. Oxford: Basil Blackwell, 1953.
- Pareto, V. Cours d'Economie Politique, Vol. II. Lausanne, 1897.
- _____. Manuel d'Economie Politique. Paris: V. Giard and E. Briere, 1909.
- Pigou, A. C. The Economics of Welfare. 4th ed. London: Macmillan and Co., Ltd., 1952.

- _____. Wealth and Welfare. London, 1912.
- Reder, Melvin W. Studies in the Theory of Welfare Economics. New York: Columbia University Press, 1947.
- Robbins, Lionel H. An Essay on the Nature and Significance of Economic Science. London: Macmillan and Co., Ltd., 1948.
- Samuelson, Paul A. The Foundations of Economic Analysis. Cambridge, 1947.
- _____. Market Mechanisms and Maximization. RAND Corporation.
- Stigler, and Boulding (eds.). Readings in Price Theory. American Economic Association. Chicago: Richard D. Irwin, Inc., 1952.
- Wicksell, Knut. Lectures on Political Economy. 2 vols. London: Routledge and Kegan Paul, Ltd., 1934.

Articles

- Baran, P. A. "On the Political Economy of Backwardness," Manchester School of Economic and Social Studies, XX (January 1952), 66-84.
- Barone, Enrico. "The Ministry of Production in the Collectivist State" (transl.), in Hayek, F. A. v., Collectivist Economic Planning. London: G. Routledge and Sons, 1935.
- Bergson, Abram (Burk). "A Reformulation of Certain Aspects of Welfare Economics," Quarterly Journal of Economics, LII (February 1938), 310-334, reprinted in Clemence (ed.), Readings in Economic Analysis, Vol. I.
- _____. "Socialist Economics," A Survey of Contemporary Economics, Vol. I, Haley, Bernard F. (ed.). Homewood, Ill.: Richard D. Irwin, Inc., 1952.
- Boulding, Kenneth E. "Welfare Economics," A Survey of Contemporary Economics, Vol. II, Haley, Bernard F. (ed.). Homewood, Ill.: Richard D. Irwin, Inc., 1952.
- Bruton, Henry J. "Growth Models and Underdeveloped Economies," Journal of Political Economy, LXI (August 1955), 322-336.
- Champernowne, D. G. "The Production Function and the Theory of Capital: A Comment," The Review of Economic Studies, XXI (2) (1953-1954), 112-135.

- Champernowne, D. G., and Kahn, R. F. "The Value of Invested Capital," The Review of Economic Studies, XXI (2) (1953-1954), 107-111.
- Chenery, H. B. "The Application of Investment Criteria," Quarterly Journal of Economics, LXVII (February 1953), 76-96.
- Clapham, J. H. "Of Empty Economic Boxes," Economic Journal, XXXIV (1922), 305-314.
- Coase, R. H. "The Marginal Cost Controversy," Economica, N.S., XIII (1946), 169-182.
- Dickinson. "Price Formation in a Socialist Economy," Economic Journal, XLIII (December 1933), 237-250.
- Dobb, Maurice. "Economic Theory and the Problems of the Socialist Economy," Economic Journal, XLIII (December 1933), 588-598.
- _____. "Economic Theory and Socialist Economy: A Reply," Review of Economic Studies, II (1934-1935), 144-151.
- Dupuit, Jules. "On the Measurement of Utility of Public Works," International Economic Papers. No. 2. Edited by Alan T. Peacock and others.
- Durbin, E. F. M. "Economic Calculus in a Planned Economy," Economic Journal, XLVI (1936), 676-690.
- Eckaus, Richard S. "The Factor Proportions Problem in Underdeveloped Areas," American Economic Review, XLV (September 1955), 539-565.
- Ellis, Howard S., and Fellner, William. "External Economies and Diseconomies," American Economic Review, 1943, reprinted in Readings in Price Theory. American Economic Association.
- Fellner, William. "Individual Investment Projects in Growing Economies: General Characteristics of the Problem and Comments on the Conference Papers," Investment Criteria and Economic Growth. Cambridge: Center for International Studies, 1955.
- Fleming, Marcus. "External Economies and the Doctrine of Balanced Growth," Economic Journal, LXV (June 1955), 241-255.
- Frisch, Ragnar. "The Dupuit Taxation Theorem," Econometrica, VII (1939), 145-150.
- _____. "A Further Note on the Dupuit Taxation Theorem," Econometrica, VII (1939), 156-157.

- Galenson, W. and Leibenstein, H. "Investment Criteria, Productivity and Economic Development," Quarterly Journal of Economics, LXIX (August 1955), 343-370.
- Hayek, Friedrich von. "Socialist Calculation: The Competitive 'Solution,'" Economica, VII (May 1940), 125-149.
- Hicks, J. R. "The Foundations of Welfare Economics," Economic Journal, LXIX (1939), 696-712.
- _____. "The Valuation of the Social Income," Economica, VII (February 1940), 105-123.
- Hotelling, Harold. "The General Welfare in Relation to Problems of Taxation and of Railway and Utility Rates," Econometrica, VI (1938), 242-269.
- _____. "The Relation of Prices to Marginal Costs in an Optimum System," Econometrica, VII (1939), 151-155.
- _____. "A Final Note," Econometrica, VII (1939), 158-160.
- Kahn, R. F. "Some Notes on Ideal Output," Economic Journal, XLV (1935), 1-35.
- Kahn, A. E. "Investment Criteria in Development Programs," Quarterly Journal of Economics, LXV (February 1951), 38-61.
- Kaldor, Nicholas. "Welfare Propositions in Economics and Interpersonal Comparisons of Utility," Economic Journal, LXIX (1939), 549-552.
- Knight, F. H. "Fallacies in the Interpretation of Social Cost," Quarterly Journal of Economics, XXXVIII (1923), 582-606.
- _____. "On Decreasing Cost and Comparative Cost," Quarterly Journal of Economics, XXXIX (1925), 331-333.
- Lange, Oscar. "The Foundations of Welfare Economics," Econometrica, X (1942), 215-228.
- Laursen, Svend. "External Economies in Economic Development" (unpublished), 1956.
- Leibenstein, Harvey. "The Proportionality Controversy and the Theory of Production," Quarterly Journal of Economics, LXIX (November 1955), 619-625.

- Lerner, A. P. "Economic Theory and Socialist Economics," Review of Economic Studies, II (1934-1935), 51-61.
- _____. "A Reply to Mr. Dobb," Review of Economic Studies, II (1934-1935), 152-154.
- _____. "Statics and Dynamics in Socialist Economics," Economic Journal, XLVII (1937), 253-270.
- Lewis, W. A. "The Two-Part Tariff," Economica, N. S., VIII (1941), 249-270.
- _____. "The Two-Part Tariff: A Reply," Economica, N. S., VIII (1941), 399-408.
- _____. "Fixed Costs," Economica, XIII (November 1946), 231-258.
- Little, I. M. D. "Direct Versus Indirect Taxes," Economic Journal, LXI (September 1951), 577-584.
- Malinvaud, Edmond. "Capital Accumulation and Efficient Allocation of Resources," Econometrica, XXI (April 1953), 233-268.
- Meade, J. E. "External Economies and Diseconomies in a Competitive Situation," Economic Journal, LXII (1952), 52-67.
- Neumann, John von. "A Model of General Economic Equilibrium," Review of Economic Studies, XIII (1945-1946), 1-9.
- Pigou, A. C. "Empty Economic Boxes: A Reply," Economic Journal, XXXII (1922), 458-465.
- Ramsey, Frank P. "A Mathematical Theory of Savings," Economic Journal, XXXVIII (December 1928), 543-559.
- Robertson, D. H. "Those Empty Boxes," Economic Journal, XXXIV (1924), 16-31.
- Robinson, Joan. "The Production Function and the Theory of Capital," The Review of Economic Studies, XXI (2) (1953-1954), 81-106.
- Rosenstein-Rodan, Paul N. "Problems of Industrialization of Eastern and South-Eastern Europe," Economic Journal, LIII (June-September 1943), 202-211.
- _____. "Programming in Theory and in Italian Practice," Investment Criteria and Economic Growth. Cambridge: Center for International Studies, 1955.

- Ruggles, Nancy. "The Welfare Basis of the Marginal Cost Pricing Principle," Review of Economic Studies, XVII (1949-1950), 29-46.
- _____. "Recent Developments in the Theory of Marginal Cost Pricing," Review of Economic Studies, XVII (1949-1950), 107-126.
- Samuelson, Paul A. "Some Aspects of the Pure Theory of Capital," Quarterly Journal of Economics, LI (May 1937), 469-496.
- _____. "Welfare Economics and International Trade," American Economic Review, XXVIII (1938), 261-266.
- _____. "The Rate of Interest under Ideal Conditions," Quarterly Journal of Economics, LIII (February 1939), 286-297.
- _____. "The Evaluation of Real National Income," Oxford Economic Papers, II (January 1950), 1-29.
- _____. "Prices of Factors and Goods in General Equilibrium," Review of Economic Studies, XXI (1) (1953-1954), 1-20.
- _____. "The Pure Theory of Public Expenditure," Review of Economics and Statistics, XXXVI (November 1954), 387-389.
- _____. "Diagrammatic Exposition of a Theory of Public Expenditure," Review of Economics and Statistics, XXXVII (November 1955), 350-356.
- _____. "Social Indifference Curves," Quarterly Journal of Economics, LXX (February 1956), 1-22.
- Scitovsky, Tibor. "A Note on Welfare Propositions in Economics," Review of Economic Studies, IX (1941-1942), 77-88.
- _____. "A Reconsideration of the Theory of Tariffs," Review of Economic Studies, IX (1941-1942), 89-110.
- _____. "Two Concepts of External Economies," Journal of Political Economy, LXII (April 1954), 143-151.
- Solow, R. M. "On the Structure of Linear Models," Econometrica, XX (1952), 29-46.
- _____. "The Production Function and the Theory of Capital," Review of Economic Studies, XXIII (1955-1956), 101-109.
- _____. "A Contribution to the Theory of Economic Growth," Quarterly Journal of Economics, LXX (February 1956), 65-94.
- Sraffa, P. "The Laws of Returns Under Competitive Conditions," Economic Journal, XXXVI (December 1926), 535-550.

Vickrey, William. "Some Objections to Marginal Cost Pricing," Journal of Political Economy, LVI (1948), 218-238.

Viner, J. "Cost Curves and Supply Curves," Zeitschrift fur National-
okonomie, III (1932), 23-46.

Young, Allyn. "Pigou's Wealth and Welfare," Quarterly Journal of
Economics, XXVII (1913), 672-686.

_____. "Increasing Returns and Economic Progress," Economic
Journal, XXXVIII (December 1928), 527-542.

Biographical Sketch

Francis M. Bator

Education

Groton School, Groton, Massachusetts, 1939-1943
Massachusetts Institute of Technology
Undergraduate 1943, 1946-1949
Graduate 1949-1951, Summer 1956

Professional

1954- Economist, Senior Research Staff, Center for
International Studies, M.I.T.
1957 Faculty Lecturer, Salzburg Seminar in
American Studies
1954 Academic Consultant (Economics Lecturer in Basic
Strategic Studies, Captains' Course in Naval
Warfare), Naval War College, Newport, R.I.
1954 Consultant, Evaluation Staff, Air War College,
Maxwell Field, Alabama
1953-1954 Executive Assistant to the Director, Center for
International Studies, M.I.T.
1951-1952 Assistant to the Director, Center for International
Studies, M.I.T.
1951-1956 Member of the Board of Directors, Natvar Corporation,
Rahway, N.J.
1951 Consultant Economist, Project Charles (USAF Contract),
M.I.T.
1950-1951 Fellow in Economics, Department of Economics, M.I.T.

Military

1944-1946 1st Lieutenant, Infantry, A.U.S.

Professional Associations

American Economic Association
Econometric Society
Royal Economic Society

Publications

"Consumption and Investment Propensities: Comments",
American Economic Review, March 1953.

Analysis of Economic Capabilities for War (Twelve
Lectures delivered at Naval War College, October-
December, 1954).