THE SIMPLE ANALYTICS OF WELFARE MAXIMIZATION

By FRANCIS M. BATOR*

It appears, curiously enough, that there is nowhere in the literature a complete and concise nonmathematical 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 neoclassical generalized diminishing returns obtain in all but one dimension—returns to scale are assumed constant. 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 sustain the maximum-welfare position. 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; diminishing returns to scale; external interactions. The discussion is still stationary and neoclassical in spirit. Then, in Part V, the consequences of violating some of the neoclassical 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 an important variety of nonconvexity—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. Finally, Part VI contains some brief remarks on possible dynamical extensions. A note on the seminal literature concludes the paper.[1]

* The author, a member of the senior staff of the Center for International Studies, Massachusetts Institute of Technology, is indebted to R. S. Eckaus and R. M. Solow for suggestive comment.

[1] Anyone familiar with the modern literature will recognize my debt to the writings of Professor Samuelson. Reference is to be made, especially, to Chapter 8 of *Foundations of Economic*

I. *Inputs, Outputs and Commodity Distribution*

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 neoclassical model; elasticity in input supplies would make simple 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 indifference curves convex to the origin—one for X and one for Y. These reflect unambiguous and consistent preference orderings for each of the two individuals (X and 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 interpersonal implications whatever and for any one individual they only permit of statements to the effect that one situation is worse, indifferent or better than another. 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.

The problem is to determine the maximum-welfare values of labor input into apples (L_A), labor input into nuts (L_N), land input into apples (D_A), land input into nuts (D_N), of total production of apples (A) and nuts (N), and, last, of the distribution of apples and nuts between X and Y (A_X, N_X, A_Y, N_Y).

A. *From Endowments and Production Functions to the Production-Possibility Curve*

Construct an Edgeworth-Bowley box diagram, as in Figure 1, with horizontal and vertical dimensions just equal to the given supplies, re-

---

Analysis (Cambridge, 1947); to "Evaluation of Real National Income," *Oxford Econ. Papers*, Jan. 1950, II, 1–29; and to "Social Indifference Curves," *Quart. Jour. Econ.*, Feb. 1956, LXX, 1–22.

spectively, of D and L, and plot the isoquants for apples with the southwest corner as origin and those for nuts with origin at the northeast corner. 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 the 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 (FF).

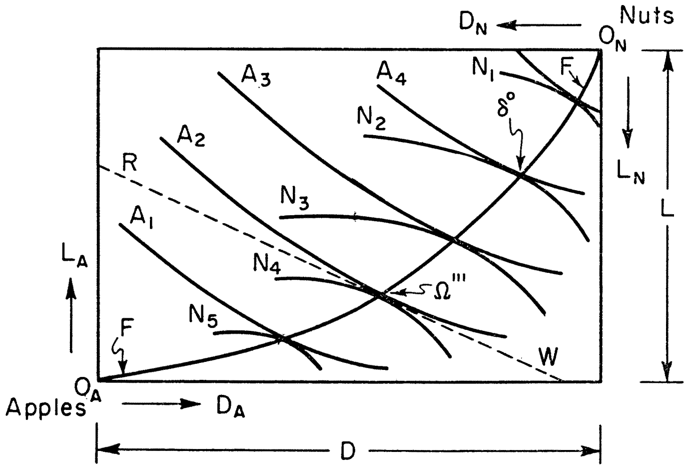

FIGURE 1

From this efficiency locus we can read off the maximal obtainable combinations of apples and nuts and plot these in the output (AN) space. Given our curvature assumptions we get the smooth concave-to-the-origin Pareto-efficient production-possibility curve $F'F'$ of Figure 2.[2] This curve, a consolidation of FF in Figure 1, represents input-output configurations such that the marginal rate of substitution (MRS) of labor for land in the production of any given quantity 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.[3]

[2] This presumes, also, that the intrinsic factor intensities of A and N differ. If they did not, $F'F'$ would be a straight line—a harmless special case. (See V-3-c below.)

[3] In marginal productivity terms, MRS, at any point, of labor for land in, *e.g.* apple production—the absolute value (drop all minus signs) of the slope of the apple isoquant (Figure 1)—is equal to

$$\left[\frac{\text{Marginal Physical Product of Land}}{\text{Marginal Physical Product of Labor}}\right]$$

in apple production at that point. In the symbolism of the calculus

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

The slope (again neglecting sign) at any point on the production-possibility curve of Figure 2, 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-

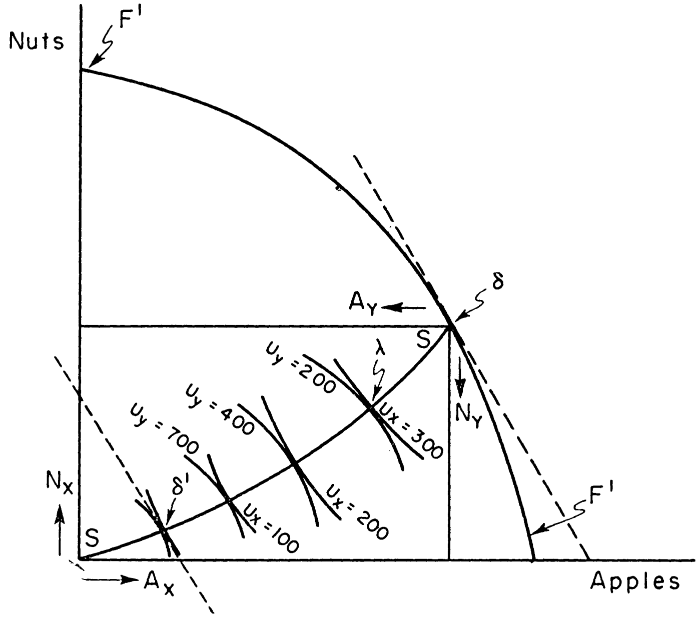

FIGURE 2

equality requirement of Figure 1. It is the marginal nut-cost of an "extra" apple—or the reciprocal of the marginal apple-cost of nuts.

B. *From the Production-Possibility Curve to the Utility-Possibility Frontier*

Pick any point, δ, on the production-possibility curve of Figure 2: it denotes a specific quantity of apples and nuts. Construct an Edgeworth-Bowley (trading) box with these precise dimensions by dropping from δ lines parallel to the axes as in Figure 2. Then draw in X's and Y's indifference maps, one with the southwest, the other with the northeast 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 Figure 2, $U_X = 300$, $U_Y = 200$. Note again, however, that this 200 is incommensurate with the 300: it does not imply that at λ X is in some sense better off than is 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 δ. Again, the locus turns out to consist of the points of tangency, SS, and for precisely the same analytical reasons. Only now it is the marginal subjective 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,[4] SS, which is associated with the single production point δ, we can now read off the maximal combinations of U_X and U_Y obtainable from δ and plot these in utility ($U_X U_Y$) space ($S'S'$, Figure 3). 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.[5]

There is a possible short-cut, however. Given our 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 SS associated with every output point δ. Go back to Figure 2. 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 SS, 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.

[4] 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.

[5] 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 Samuelson's proof of the nonexistence of community indifference curves such as would permit the derivation of demand curves for apples and nuts. The subjective "community" MRS between A and N for given fixed A and N, *e.g.*, at δ in Figure 2, would surely depend on how the A and N are distributed, *i.e.*, on which $U_X U_Y$ point on SS is chosen. Hence the slope of a "joint" XY indifference curve at δ is not uniquely fixed by AN. (See citation [11] in bibliography.)

If, for example, at δ one can get two apples by diverting resources and reducing nut-output by one, a point on SS 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 assumption that MRS = 1 both X and Y are left indifferent: U_X and U_Y remain unaltered. But we have an extra apple left over; since this permits raising U_X and/or U_Y, the initial situation was not on the U_XU_Y frontier.[6]

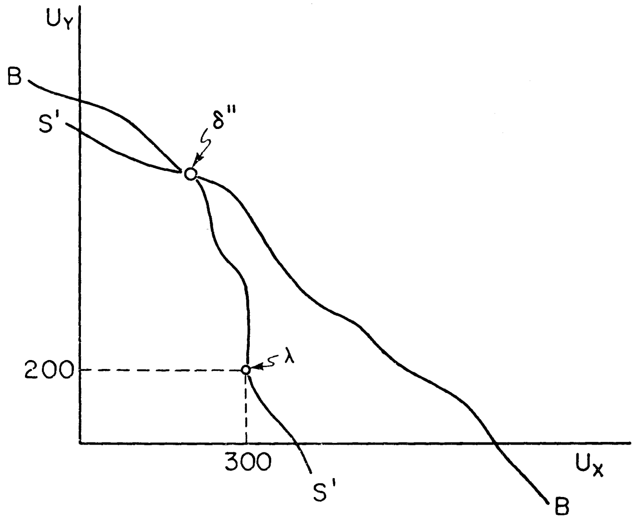

FIGURE 3

To be on the grand utility-possibility frontier (BB of Figure 3), then, MRT_δ must equal the (equalized) MRS of the indifference contours along the SS associated with δ. This requirement fixes the single U_XU_Y point on SS that lies on the "envelope" utility-possibility frontier, given the output point δ. Pick that point on SS, in fact, where the joint slope of the indifference curves is exactly parallel to the slope at δ of the production-possibility curve. In Figure 2 this point is at δ', which gives the one "efficient" U_XU_Y combination associated with the AN mix denoted by δ. This U_XU_Y combination can then be plotted as δ'' in Figure 3.[7]

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

[7] Never mind, here, about multiple optima. These could occur even with our special curvature assumptions. If, for example, both sets of indifference curves show paths of equal MRS

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, *BB*. Each point of this frontier gives the maximum of U_X for any given feasible level of U_Y and vice versa.

C. *From the Utility-Possibility Frontier to the "Constrained Bliss Point"*

But *BB*, the grand utility-possibility function, is a curve and not a point. Even after eliminating all combinations of inputs and outputs

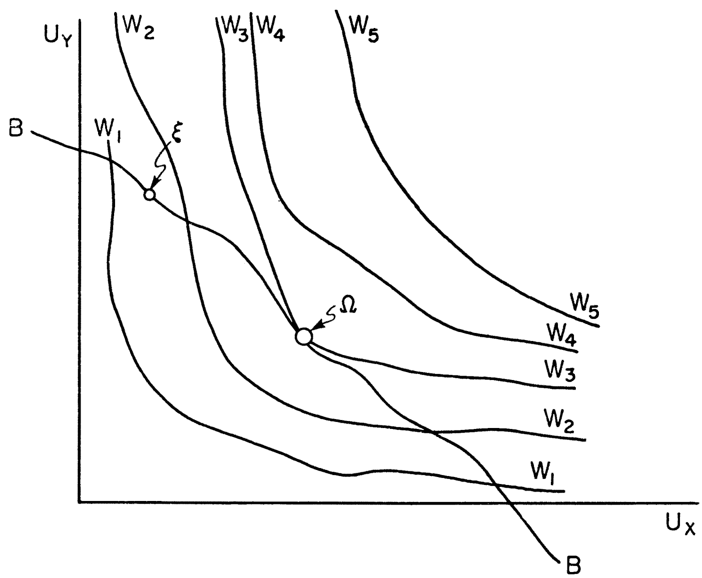

FIGURE 4

that are nonefficient in a Paretian sense, there remains a single-dimensional infinity of "efficient" combinations: one for every point on *BB*. 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 ascientific.[8] There are no considerations of

that coincide with straight lines from the origin and, further, if the two preference functions are so symmetrical as to give an SS_δ that hugs the diagonal of the trading box, then either every point on SS_δ will satisfy the MRS=MRT criterion, or none will. For discussion of these and related fine points see Parts IV and V.

[8] Though it may provide the anthropologist or psychologist with interesting material for scientific study.

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 indifference contours in utility space, as in Figure 4, the problem becomes fully determinate.[9] "Welfare" is at a maximum where the utility-possibility envelope frontier BB touches the highest contour of the W-function.[10] In Figure 4, this occurs at Ω.

Note the unique quality of that point Ω. It is the only point, of all the points on the utility frontier BB, that has unambiguous normative or prescriptive significance. Pareto-efficient production and commodity-distribution—being on $F'F'$ and also on BB—is a necessary condition for a maximum of our kind of welfare function, but is not a sufficient condition.[11] The claim that any "efficient" point is better than "inefficient" configurations that lie inside BB is indefensible. It is true that given an "inefficient" point, there will exist *some* point or points on BB that represent an improvement; but there may well be many points on BB that would be worse rather than better. For example, in terms of the ethic denoted by the specific W-function of Figure 4, Ω on BB is better than any other feasible point. But the efficient point ξ is distinctly inferior to any inefficient point on or northeast of W_2. If I am X, and if my W-function, which reflects the usual dose of self-interest, is the test, "efficient" BB points that give a high U_Y and a very low U_X are clearly less desirable than lots of inefficient points of higher U_X.[12]

[9] In the absence of implicit income redistribution these curves cannot be transposed into output-space. They are not community indifference curves which would permit the derivation of demand schedules. See fn. 5 and 12, also IV-3.

[10] 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.

[11] 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.

[12] 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 behavior, unless he be dictator, subjected by "nature" to binding constraints. In a sense only a society, considered as exhibiting a political consensus, has a W-function subject to empirical inference (*cf.* IV-3). The distinction—it has a Rousseauvian flavor—while useful, is of course arbitrary. Try it for a masochist; a Puritan. . . .

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

We can now retrace our steps. To Ω on BB in Figure 4, there corresponds just one point, Ω', on the production-possibility curve $F'F'$ in Figure 5. (We derived BB, point by point, from $F'F'$ of Figure 2: and the $F'F'$ of Figure 5 is copied from that of Figure 2.) Ω' fixes the output

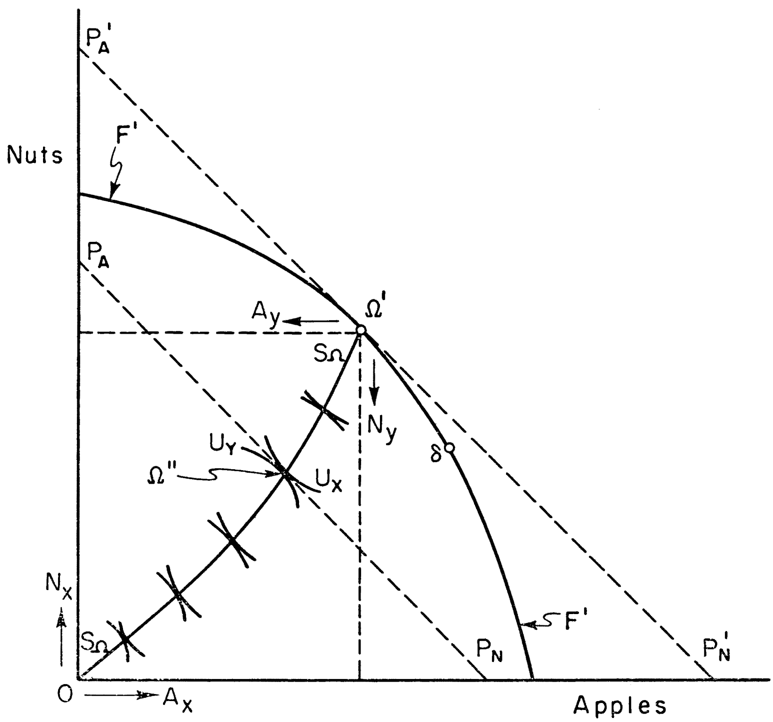

FIGURE 5

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 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, FF of Figure 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.

II. *Prices, Wages and Rents*

The above is antiseptically independent of institutional context, notably of competitive market institutions. It could constitute an intellectual exercise for 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, is 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.[13]

Can these constants—prices, wages, rents—be identified in our diagrammatic representations?[14] 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.[15] 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), and the rental rate of land (r).[16]

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 must be produced at a minimum total cost.[17] The ele-

[13] 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.

[14] 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.

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

[16] Since we are still assuming that all the functions have neoclassical 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 nonconvexities (*e.g.*, increasing returns to scale), yet maintain for the constants some essentially price-like qualities. More on this in Part V.

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

mentary 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 r/w.

Reversing the train of thought, the input blend denoted by the point Ω''' in Figure 1 implies a shadow r/w ratio that just equals the MRS of labor for land in the production of both apples and nuts at that point Ω'''. $\text{MRS}_{\Omega'''}$ is given by the (equalized) slopes of the isoquants at Ω'''. The implicit r/w, therefore, must equal the slope of the line RW that is tangent to (both) the isoquants at Ω'''.[18]

The slope of RW 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 Figure 5, as denoting the p_A/p_N ratio implied by the solution. X, as also 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 p_A/p_N. He, and Y, will pick Ω'' only if p_A/p_N is equal to the absolute slope of the tangent (P_AP_N) at Ω''. This slope, therefore, fixes the Ω-value of p_A/p_N.[19]

Note that this makes p_A/p_N equal to the slope also of the production-possibility curve $F'F'$ at Ω'.[20] This is as it should be. If $p_A/p_N=10$, *i.e.*, if one apple is "worth" ten nuts on the market, it would be odd in-

[18] Again, absolute values of these slopes are implied throughout the argument. Recall from footnote 3 that the labor-for-land MRS, the absolute slope of the isoquants at Ω''' as given bv RO_A/WO_A, is equal to the

$$\left[\frac{\text{Marginal Physical Product of Land}}{\text{Marginal Physical Product of Labor}}\right] \text{ratio.}$$

Our shadow r/w, then, turns out to be just equal to that ratio.

[19] The price-ratio relates reciprocally to the axes: $p_A/p_N=P_AO/P_NO$ in Figure 5. Along, *e.g.*, X's indifference curve (U_X at Ω'') a rise in p_A/p_N, *i.e.*, a steepening of P_AP_N, results in a substitution by X of nuts for apples; ditto for Y.

[20] 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 (see p. 27 above).

deed, in our frictionlessly efficient world of perfect knowledge, if 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 p_A/p_N differed from MRT at Ω'.

We have identified the r/w and 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, *i.e.*, the price, just equals the marginal cost of producing that output.[21] 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 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.

Alternatively, given our assumption that production functions show constant returns to scale, we can make use of Euler's "product exhaustion" theorem. Its economic content is that if constant returns to scale prevails, the total as-if-market-imputed income of the factors of production just "exhausts" the total value of the product. This means, simply, that $wL+rD=p_AA+p_NN$, and it provides a third relationship between w, r, p_A and p_N for the Ω-values of L, D, A and N.[22]

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*

[21] 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.

[22] 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 a relationship between input price and output price involves a measure of sleight of hand. It is only as a consequence of the relationships between price and marginal productivity (*cf.* the preceding paragraph) that the theorem assures equality of income with value of product.

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 50 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 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.[23]

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 neoclassical model, of a set of constants embedded in the "technocratic" maximum-of-welfare problem, that can be viewed as competitive market prices.[24] 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. *Factor Ownership and Income Distribution*

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 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'', *i.e.*, $p_AA_X+p_NN_X$; 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

[23] For the possibility of inessential indeterminacies, however, see Part IV-2.

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

frontier (*BB* in Figure 4). But only by the sheerest accident will that point on *BB* be better in terms of my *W*-function, or Thomas Jefferson's, or that of a "political consensus," than a multidimensional infinity of other points *on or off BB*. As emphasized above, only one point on *BB* can have ultimate 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.[25]

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 nonhuman means of production are vested in the community, hence where the proceeds of nonwage income are distributed independently of marginal productivity, marginal-rate-of-substitution considerations. 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.[26]

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 per cent of total L; each man working $\frac{1}{2}L$ hours may well not do.[27]

But all this is diversion. For our noninstitutional purposes it is sufficient to determine the particular L_X, D_X, L_Y and D_Y that are consistent

[25] It is of course possible to break the link between factor ownership and "final" income dis tribution by means of interpersonal transfers. Moreover, if such transfers are effected by means of costless lump-sum devices—never mind how feasible—then it is possible, in concept, to attain the Ω-implied distribution irrespective of market-imputations. But no decentralized price-market-type "game" can reveal the pattern of taxes and transfers that would maximize a particular *W*-function. "Central" calculation—implicit or explicit—is unavoidable.

[26] 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 blur. Except in an "Austrian" world, however, it would not vanish. As long as men retain a measure of control over the quality and time-shape of their own services, there will always remain an incentive problem.

[27] All this is based on the "Austrian" assumption that labor is supplied inelastically; further, that such inelasticity is due not to external compulsion, but rather to sharp "corners" in the preference-fields of X and Y in relation to work-leisure choices. More than this, the *W*-function must not be sensitive to variations in the L_XL_Y mix except as these influence income distribution.

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:

$$wL_X + rD_X = p_A A_X + p_N N_X,$$

and

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

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 Euler's theorem taken jointly with the marginal productivity conditions that give the solution for the eight variables, A_X, N_X, A_Y, . . . which are here taken as known. Hence, we have only three independent equations. This is as it should be. It means only that with our 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 nonlabor 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 Figure 1, hence the position of the production-possibility curve of Figures 2 and 5, 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, or many axes for many differently irksome kinds of labor), and the W-function. Three (or more) dimensional diagrams would be needed for a geometrical solution.

IV. *Some Extensions*

We have demonstrated the solution of the maximum problem of modern welfare economics in context of the simplest statical and stationary neoclassical model. Many generalizations and elaborations suggest themselves, even if one remains strictly neoclassical 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.[28]

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, FF, in the production box of Figure 1; similarly, that we would get such an "internal" SS in the trading boxes of Figures 2 and 5. Nothing in our

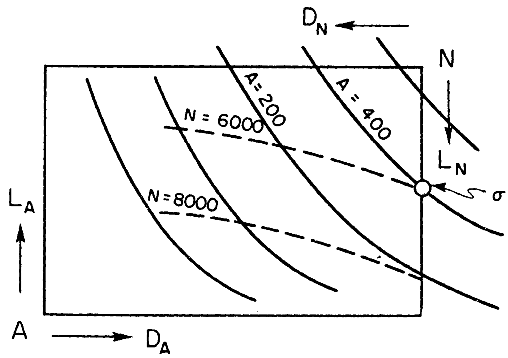

FIGURE 6

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-tangencies along the edges of the box? Figure 6 illustrates this possibility. The maximum feasible output of A, for $N=6000$, occurs at σ, where $A=400$; but at σ 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 Figure 6, and with technology as denoted by the isoquants, it is not possible to reallocate inputs until the MRS of labor for

[28] Rigorous general treatment of the $m \times n \times q$ 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 nonnormative point of view, see P. A. Samuelson, "Prices of Factors and Goods in General Equilibrium," *Rev. Econ. Stud.*, 1953–1954, XXI (1), No. 54, 1–20.

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

$$\left[\frac{\text{marginal productivity of land}}{\text{marginal productivity of labor}}\right] \text{ratio}$$

in apple production exceeds that in nut production even when, as at σ, *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.*, $L_A + L_N \leqq L$), such phenomena of bumping into the axes can be handled; further, that only inessential indeterminacies occur in the implied shadow-price configuration.[29]

[29] All this can perhaps be made clearer by two examples. The essential requirement for A_σ to be at a maximum for $N=6000$ is that the intersection at the boundary be as in Figure 6 rather than as in Figure 7. In the latter, σ' gives a minimum of A for $N=6000$; the true maximum is at σ''. The distinction between σ in 6 and σ' in 7 is between the relative rates of change of the two MRS's. The price indeterminacy implied by the maximum, *i.e.*, the fact that σ is consistent with an 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 $MC=p$ 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 . . . " *Am. Econ. Rev.*, Sept. 1948, XXXVIII, 607–17) this is incorrect: What is important is that at smaller than equilibrium output MC be less than price and at higher outputs MC exceed price. It is true, but quite harmless to the theory, that such a situation does leave a range of indeterminacy in the price

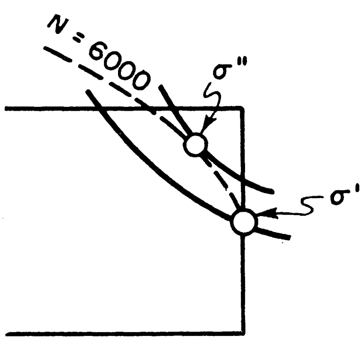

FIGURE 7

3. We stressed, above, the nonexistence of *community* indifference contours such as would provide a unique ranking, for the community as a whole, of various output combinations.[30] Individual marginal rates of substitution between, *e.g.*, apples and silk shirts, equalized along a trading-box contract curve to give a "community" MRS, are likely to be sensitive to the distribution of income[31] between gourmets and dandies; accordingly, community MRS at a given point in commodity space, *i.e.*, the slope of a curve of community indifference, will vary with movements along the associated utility-possibility curve. However, once the most desirable $U_X U_Y$ combination for a given package of A and N is fixed, MRS at that AN-point becomes determinate. It follows, as recently pointed out and proved by Samuelson,[32] that if the observed community continuously redistributes "incomes" in utopian lump-sum fashion so as to maximize, in utility space, over the W-function implied by a political consensus, then there does exist, in output space, a determinate *social* indifference function which provides a ranking for the community as a whole of all conceivable output combinations. This function, which yields conventionally convex social indifference contours, can be treated as though a single mind were engaged in maximizing it. Moreover, in concept and if granted the premise of continuous redistribution, its contours are subject to empirical inference from observed price-market data.

This existence theorem justifies the use of *social* indifference maps—maps "corrected" for distribution—in handling problems of production efficiency, international trade, etc.—a substantial analytical convenience.[33] More important, it provides a conceptual foundation, however abstract, for prescription based not on just any arbitrary ethic, but rather on the particular ethic revealed by a society as reflecting its own political consensus.[34]

4. It is useful, and in a mathematical treatment not difficult, to drop the "Austrian" assumption of inelastically supplied inputs, and intro-

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 simple substitution. On all this, see the literature of linear programming (*e.g.*, citations [10] and [13]).

[30] See fn. 5.

[31] In terms of abstract purchasing power.

[32] See citation [11].

[33] Note, however, that none of this eliminates the need for a W-function: social indifference contours are a convex function of individual taste patterns of the usual ordinal variety taken jointly with an implicit or explicit W-function of "regular" content and curvature. Further, no ultimate superiority attaches to the W-function implied by a particular political consensus. One may disapprove of the power relationships on which such consensus rests, etc.

[34] Needless to say, feasibility is not here at issue. Even on this level of abstraction, however, matters become much more difficult once account is taken of the fact that the world is not stationary.

duce leisure-work choices.[35] 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 nonimposed choice. A dictator, too, has to take account of such choices, if only because of feasibility limitations on coercion.

5. We assumed away joint-product 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.

6. 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" nonproduced 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 is violated by such generalization.[36]

7. What if instead of assuming that production functions show constant returns to scale, we permit diminishing returns to proportional expansion of inputs? This could be due either to inherent nonlinearities in the physics and topography of the universe, or to the existence of some unaccounted-for but significant input in limited, finite-elastic supply.[37]

[35] If we assume only one commodity, say apples, and replace the second good by leisure (or by negative labor input); and if we let the second-good production function be a simple linear relation, our previous geometry will portray the simplest goods-leisure situation.

[36] Analytically, this is done by designating all produced goods as X_1, X_2, X_3 . . . . 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 . . . 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.

[37] 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

Diminishing returns to scale, as distinct from increasing returns, does 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. It does 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.[38]

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- or nut-producing firms could be assumed indeterminate. Every firm could be assumed able to produce any output up to A_Ω (or N_Ω) at constant 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.

It is, however, natural, at least in a context of market institutions, to think of decreasing returns to scale, 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 intramarginal 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.[39]

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.

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

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

At any rate, no serious damage is done to the statical system by decreasing returns to scale. When it is a matter of actually computing a maximum problem the loss of linearity is painful, but the trouble is in the mathematics.[40]

8. 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 (nonpecuniary) external economies or diseconomies of production and consumption. The assumption is reflected in four 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 a 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 beekeeper whose honey output will increase, other things equal, if the neighboring apple producer expands *his* output (hence his apple blossom "supply").[41] 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, 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 nonreversible 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.[42] But on our

[40] 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 setup.

[41] The other type of externality treated in the neoclassical literature, the type Jacob 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.

[42] 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

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 setup.[43]

c. X and Y were assumed insensitive, also, to the input-output configuration of producers, except as this 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.

d. 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 system, 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 nonexistence 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

---

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 lifetime services of the trainee; the private benefit to the producer accrues until the man quits to go to work for a competitor.

[43] For an important exception, however, see fn. 44 below.

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 either to blur or vanish with "direct interaction." Most varieties of such interaction destroy the "duality" of the system: the constants embedded in the maximum problem, if any, lose significance as prices, wages, rents. They will not correctly account for all the "costs" and "benefits" to which the welfare function in hand is sensitive.[44]

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 setup 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. *Relaxing the Curvature Assumptions: Kinks and Nonconvexities*

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

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

[44] 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 analogous to those of Part I—conditions that properly account for full social costs and benefits. So-called "public goods," *e.g.*, national defense, which give rise to direct interaction since by definition their consumption is joint—more for X means not less but more for Y—are an important example. Maximizing yields MRS conditions that bear intriguing correspondence to those which characterize ordinary private-good situations. But these very MRS conditions serve to reveal the failure of duality. (Samuelson's is again the original and definitive treatment. See citation [12].)

body of theory.[45] All the efficiency conditions have their counterparts in such a system, and the existence of implicit "prices" embedded in the maximum problem is, if anything, even more striking.[46]

2. Easing of the neoclassical 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 nonincreasing) 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 mathematical economics, is rather simple. Take a single isoquant such as MM in Figure 8a. 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 MM 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 MM 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 Figure 8a. In Figure 8b, however, where the isoquant is 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 Figure 8c, is "convex"—this is why the generalization of (1) above was painless.[47]

What kind of trouble does nonconvexity create? In the case of concave-to-the-origin isoquants, *i.e.*, nonconvex isoquants, the difficulty is

[45] Isoquants in such a setup 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 Figure 8c.

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

[47] 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.

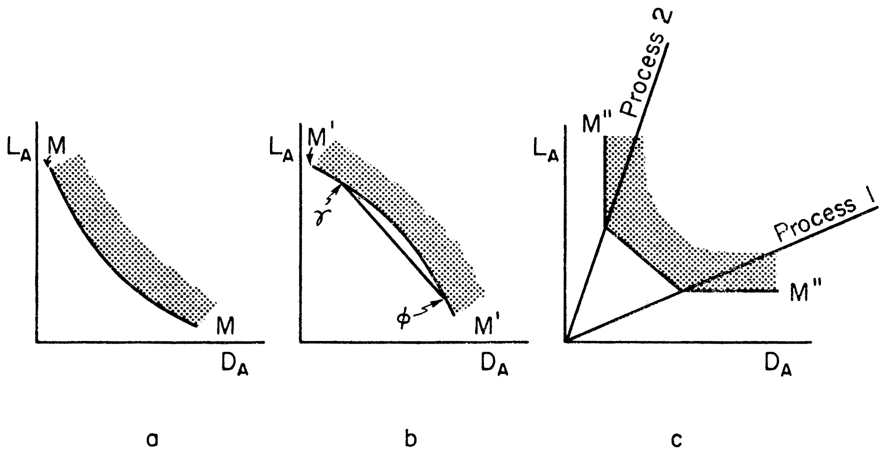

FIGURE 8

easy to see. Look back at Figure 1 and imagine that the old nut-isoquants are really those of apple producers, hence oriented to the southwest, and vice versa for nuts. Examination of the diagram will show that the locus of tangencies, FF, 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.[48]

3. This is not the occasion for extensive analysis of convexity problems. It might be useful, however, to examine one very important variety of nonconvexity: increasing returns to scale in production. Geometrically, increasing returns to scale is denoted by isoquants that are closer and closer together for outward movement along any ray from

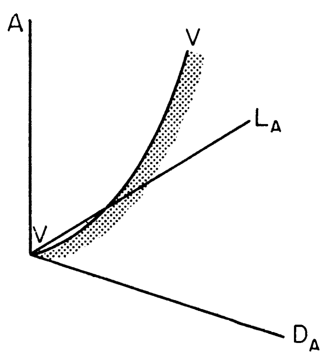

FIGURE 9

[48] 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.

the origin: to double output, you less than double the inputs. Note that the isoquants still bound convex sets in the LD plane (they are still as in Figure 8a). But in the third or output dimension of a two-input, one-output production surface, slices by vertical planes through the origin perpendicular to LD will cut the production surface in such a way as to give a boundary such as VV in Figure 9. It is evident that VV bounds a nonconvex set of feasible points, so the full three-dimensional set of feasible input-output points is not convex.

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

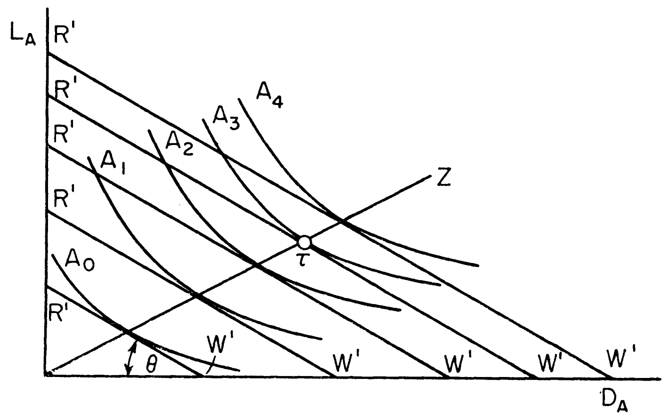

FIGURE 10

a. *Increasing returns to scale and* AC *curves.* It is a necessary consequence of increasing returns to scale that at the maximal configuration of inputs, outputs and input prices, producers' AC curves decline with increasing output. By the definition of increasing returns to scale at a given point τ of a production function, successive isoquants in the neighborhood of τ lie closer and closer together for movement "north-east" along the ray from the origin through τ (Z in Figure 10). As Figure 10 is drawn, the ray Z happens also to correspond to an expansion path for the particular r/w ratio denoted by the family of isocost lines $R'W'$: each $R'W'$ is tangent to an isoquant along Z. Given $r/w = |\text{tangent } \theta|$, a profit-maximizing apple producer will calculate his minimum total cost for various levels of output from input-output points along Z. But along Z the equal cost $R'W'$ tangents in the neighborhood of τ 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 AC curve at τ for $r/w = |\text{tangent } \theta|$ must be falling.

Suppose the expansion path for $r/w = |\text{tangent } \theta|$ happened not to correspond to the ray Z, but only to cross it at τ. 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 Figure 10: the negative effect on AC would be reinforced. The point is, simply, that if for movement along a ray from the origin cost per unit of output declines, AC 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 τ? It has just been shown that AC at τ is falling. A falling AC implies a marginal cost curve (MC) that lies *below* the average. But if τ is the Ω'''-point, the shadow-p_A will just equal MC of τ. It follows that the maximum-of-W configuration requires $p_A < AC$, *i.e.*, perpetual losses. Losses, however, are incompatible with real life (perfect) markets; hence where increasing returns to scale prevails 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.[49]

Increasing returns to scale has also a "macro" consequence that is associated with $p < AC$. For constant returns to scale, we cited Euler's 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: $rD + wL > p_A A + p_N N$.[50]

b. *Increasing returns to scale and* MC *curves*. Where nonconvexity of the increasing-returns-to-scale variety results in falling AC 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 AC is to fall, MC must lie below AC, but at the requisite Ω-output, MC's may nevertheless be rising, as for example at ϵ in Figure 11. 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 sys-

[49] 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.

[50] 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.

tem 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 AC's due to increasing returns to scale either one or both industries will show losses, but these are irrelevant to optimal allocation.[51]

What if for a maximum-of-W producers are required to produce at points such as ϵ', where $p=MC$ but MC is declining?[52] The fact that ϵ' shows $AC>MC=p$, hence losses, has been dealt with above. But more is involved. By the assumption of a falling MC-curve, the horizon-

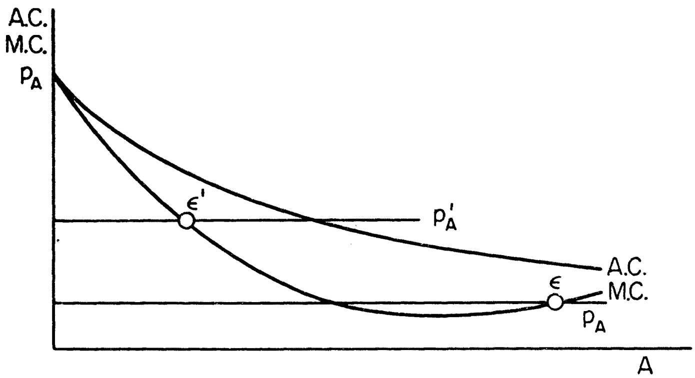

FIGURE 11

tal price line at ϵ' cuts the MC curve from below, hence profit at ϵ' 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 MC 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.[53]

Paradoxically enough, the correct rule for all industries whose MC is

[51] 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 be not at $p=MC$ 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 $MC>p$ everywhere, that distinguishes Lange-Lerner systems from "real-life" perfect markets, both viewed as "analogue computers."

[52] This would necessarily be the case, for instance, with Cobb-Douglas type increasing-returns-to-scale functions. Such functions imply ever-falling MC curves, for whatever r/w ratio.

[53] Note that a falling MC curve is simply a reflection of nonconvexity in the total cost curve.

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 raising (lowering) of prices by the responsible "Ministries" according to whether supply falls short of (exceeds) demand, is all that is needed.[54] 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 MC'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 $AC > MC$ with declining MC, 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 Figure 2). While a proportional expansion of L_A and D_A by a factor of two would more than double apple output, 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 FF in Figure 1 makes this clear for the constant-returns-to-scale case. As we move from any initial point on FF toward more A and less N, the L_A/D_A and L_N/D_N proportions change.[55]

The point is that if, as in Figure 1, land is important relative to labor in producing apples, and vice versa for nuts, expansion of apple production 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

[54] 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?)

[55] Only if FF 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.

phenomenon explains why with constant returns to scale in both functions the production-possibility curve shows concave-to-the-origin curvature. Only if FF in Figure 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 farther towards the origin. In fact, as long as the FF curve has shape and curvature as in Figure 1, the production-possibility curve, $F'F'$ in Figures 2 and 5, 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. This is true in two senses: (1) They still reflect marginal rates of substitution and transformation. Any package of L, D, A and N worth \$1 will, *at the margin*, be just convertible by production and exchange into any other package worth \$1, no more, no less: a dollar is a dollar is a dollar. . . . [56] (2) The total value of maximum-welfare "national" output: p_AA+p_NN, valued at these shadow-price constants, will itself be at a maximum. A glance at Figure 5 makes this clear: at the price-ratio denoted by the line $P'_AP'_N$, 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 Figure 12.[57] This exhibits nonconvexity in output space. 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 Figure 12 confirms our geometric intuition that in situations of nonconvex 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_AA+p_NN.[58]

[56] For the infinitesimal movements of the calculus.

[57] Try two functions which are not too dissimilar in "factor intensity."

[58] For $p_A/p_N=|\text{tangent } \psi|$, (p_AA+p_NN) is at its maximum at the intersection of $F''F''$ with the A-axis. Recall, incidentally, that in situations of falling MC producers were required to ***minimize*** profits.

But with nonconvexity 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 MC's. Take a one-person situation such as in Figure 13. 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

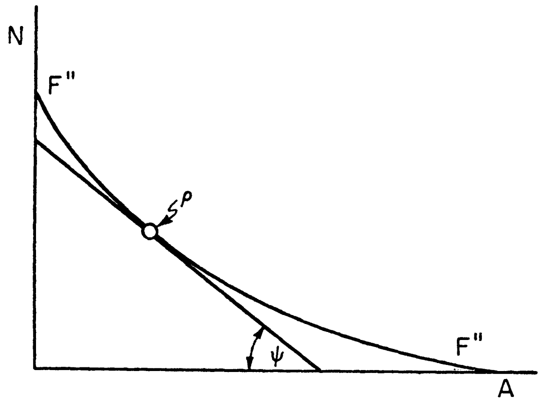

FIGURE 12

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 nonconvex situations *relative* curvatures are crucial: tangency points may as well be minima as maxima.[59]

[59] 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.

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 MC curves. It is true that the $[MC_A/MC_N]$ ratio associated with a point such as Ω' in Figure 5 must be smaller than $[MC_A/MC_N]$ 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 $pA/p_N = [MC_A/MC_N]$, and at Ω' the slope is less steep than at δ. It is also true that along a nonconvex production-possibility curve, such as that of Figure 12, an increase in A and a decrease in N are associated with a *decline* in $[MC_A/MC_N]$. But it does not follow, *e.g.*, in the first case of Figure 5, that at Ω' MC_A must be rising for an increase in A sufficiently to offset a possibly falling MC_N. (Remember, in moving from Ω' to δ we move to the right along the A-axis but to the left along the N-axis.) For any departure from Ω' will, in general, involve a change in input shadow-prices, hence *shifts* in the MC 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 Figure 5, both MC_N and MC_A are falling, though $F'F'$ is convex.

So much for nonconvexity. In its mildest form, if isoquants and indifference curves retain their normal curvature and only returns to scale "increase," nonconvexity 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 nonmarginal *total* conditions for effective real-life market functioning, *e.g.*, that all producers have at least to break

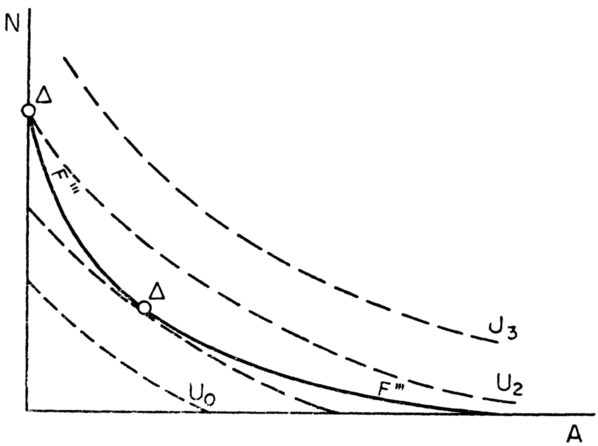

FIGURE 13

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 obtains, 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.[60]

VI. *Dynamics*

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.[61] We

[60] 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 *AC* 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.

[61] See fn. 11.

have 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 possible here, but some indication of the directions which such extension might take is perhaps warranted:

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 days 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 cannot be exhausted instantaneously but only over time. In a descriptive sense, a 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.*, steamrollers, 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.[62] 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

[62] 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 a given kind 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.

takes time, hence seed cannot 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 of any instant are not manna from heaven. Their supply depends on past output decisions. Next year's production possibilities will depend, in part, on the supply of machine tools; this, in turn, partly depends on the resources devoted this year to the construction of new machine tools. This is the problem of investment. From today's point of view 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 determine 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.[63]

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,[64] consistent with

[63] 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?

[64] Capital is characterized not only by the fact of durability, but also by lumpiness or indivisibility "in scale." Such lumpiness results in nonconvexity, hence causes serious analytical troubles.

the convexity assumptions required for solubility and duality. The results, which are the fruit of some very recent and pathbreaking work by R. M. Solow and P. A. Samuelson (soon to be published), define intertemporal 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 intertemporal 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.[65]

Incidentally, the geometric techniques of Part I are fully adequate to the task of handling a Solow-Samuelson dynamical setup for a 2 by 2 by 2 world. Only now the dimensions of the production box and 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.[66] There are, also, important intertemporal 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 nonstationary dynamics: the intrinsic uncertainty that attaches to the notion of future.[67] 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.

A Historical Note on the Literature

Note: For a short but substantive history of the development of thought in this field, the reader is referred to Samuelson's synthesis (nonmathematical), pp. 203–19 of *Foundations* [1].

[65] For possible exception to this, due to sensitivity of the volume of saving, hence of investment, to "as imputed" income distribution, *cf.* my "On Capital Productivity, Input Allocation and Growth," *Quart. Jour. Econ.*, Feb. 1957, LXXI, 86–106.

[66] 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.

[67] 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.

See also Bergson, "Socialist Economics," *Survey of Contemporary Economics*, Vol. I [2] and Boulding, "Welfare Economics," *Survey*, Vol. II [3].

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* [4]. Pigou, the last of the great premoderns was also, as witness the *Economics of Welfare* [5], among the first of the moderns. But he was not the first. Vilfredo Pareto, writing during the first years of the century, has a pre-eminent claim [6]. It is his work, and Enrico Barone's after him [7]—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 [8]. Bergson, in particular, in a single article in 1938, was the first to make us see the structure whole. More recently, Kenneth Arrow has explored the logical underpinnings of the notion of a social welfare function in relation to social choice [9]; T. C. Koopmans, Gerard Debreu and others have tested more complicated systems for duality [10]; Samuelson has developed a meaningful species of social indifference function [11] and derived efficiency conditions for "public goods" [12]; and Robert Solow and Samuelson, in work soon to be published, have provided a dynamical extension [13, 14].

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. L. von Mises [15] and later F. A. Hayek [16] were the principal proponents of unadulterated *laissez faire*, while H. D. Dickinson, Oscar Lange, Lerner and Maurice Dobb stand out on the other side [17]. The decentralized socialist pricing idea, originally suggested by Barone and later by F. M. 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, Nicholas Kaldor [18] and J. R. Hicks [19] took up Lionel Robbins' [20] 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.[68] Tibor Scitovsky pointed out an important asymmetry in the Kaldor-Hicks test [21] and Samuelson in the end demonstrated that a "welfare-function" denoting an ethic was

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

needed after all [22]. I. M. D. Little tried, but I think failed, to shake this conclusion [23].[69] 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 Harold Hotelling, Ragnar Frisch, J. E. Meade, W. A. Lewis, are devoted to the question of optimal pricing, marginal-cost or otherwise, in public utility (M.C.<A.C.) situations [24]. Hotelling, H. P. Wald, M. F. W. Joseph, E. R. Rolph and G. F. Break, Little, and more recently Lionel McKenzie, have, in turn, analyzed alternative fiscal devices for covering public deficits [25]. Last, a number of the above, notably Lerner, Kaldor, Samuelson, Scitovsky, Little, McKenzie and, most exhaustively, Meade, as well as R. F. Kahn, Lloyd Metzler, J. de V. Graaf, H. G. Johnson and others have applied the apparatus to questions of gains from international trade, optimal tariffs, etc. [26].

Bibliography

[1] P. A. Samuelson, *Foundations of Economic Analysis* (Cambridge, 1947).

[2] A. Bergson, "Socialist Economics," in H. S. Ellis, ed., *A Survey of Contemporary Economics*, Vol. I (Philadelphia, 1948).

[3] K. E. Boulding, "Welfare Economics," in B. F. Haley, ed., *A Survey of Contemporary Economics*, Vol. II (Homewood, Ill., 1952).

[4] A. C. Pigou, *Wealth and Welfare* (London, 1912).

[5] ———, *The Economics of Welfare*, 4th ed. (London, 1932).

[6] V. Pareto, *Manuel d'économie politique* (Paris, 1909).

[7] E. Barone, "The Ministry of Production in the Collectivist State," transl. in F. A. Hayek, *Collectivist Economic Planning* (London, 1935).

[8] See A. P. Lerner, *The Economics of Control* (London, 1944); A. Bergson (Burk), "A Reformulation of Certain Aspects of Welfare Economics," *Quart. Jour. Econ.*, Feb. 1938, LII, 310–14, reprinted in R. V. Clemence, ed., *Readings in Economic Analysis*, Vol. I (Cambridge, 1952); P. A. Samuelson, *op. cit.*, Ch. 8.

For other works, see references in Samuelson, *op. cit.*, p. 219, and in Bergson's and Boulding's *Survey* articles [2, 3].

[9] See K. J. Arrow, *Social Choice and Individual Values* (New York, 1951).

[10] P. A. Samuelson, *Market Mechanisms and Maximization* (unpublished, RAND Corporation Research Memo., 1949).

T. C. Koopmans, *Activity Analysis of Production and Allocation* (New York, 1951); also R. Dorfman, "Mathematical or 'Linear' Programming," *Am. Econ. Rev.*, Dec. 1953, XLIII, 797–825.

[11] P. A. Samuelson, "Social Indifference Curves;" *Quart. Jour. Econ.*, Feb. 1956, LXX, 1–22.

[12] ———, "The Pure Theory of Public Expenditure," *Rev. Econ. Stat.*, Nov. 1954, XXXVI, 387–89.

———, "Diagrammatic Exposition of a Theory of Public Expenditure," *Rev. Econ. Stat.*, Nov. 1955, XXXVII, 350–56.

[13] R. Dorfman, R. M. Solow and P. A. Samuelson, *Linear Programming and Economic Analysis* (RAND Corporation, forthcoming), esp. Ch. 11, 12. Ch. 14 contains a most elegant exposition by R. M. Solow of modern welfare theory in linear programming terms.

[14] Four other works should be mentioned: M. W. Reder, *Studies in the Theory of Welfare*

[69] While I find Little's alternative to a welfare function ("an economic change is desirable if it does not cause a bad redistribution of income, and if the potential losers could not profitably bribe the potential gainers to oppose it" [p. 105]) no alternative at all, his is a provocative evaluation of modern welfare theory. For an evaluation, in turn, of Little, see K. J. Arrow, "Little's Critique of Welfare Economics," *Am. Econ. Rev.*, Dec. 1951, XLI, 923–34.

Economics (New York, 1947), is a book-length exposition of modern welfare theory; Hla Mynt's *Theories of Welfare Economics* (London, 1948), treats classical and neoclassical writings; W. J. Baumol in *Welfare Economics and the Theory of the State* (London, 1952), attempts an extension to political theory; in a different vein, Gunnar Myrdal's *Political Elements in the Development of Economic Theory*, transl. by Paul Streeten (London, 1953), with Streeten's appendix on modern developments, is a broad-based critique of the premises of welfare economics.

[15] For the translation of the original 1920 article by Mises which triggered the controversy, see F. A. Hayek, ed., *Collectivist Economic Planning* (London, 1935).

[16] See esp. F. A. Hayek, "Socialist Calculation: The Competitive Solution," *Economica*, May 1940, VII, 125–49; for a broad-front attack on deviations from *laissez faire* see Hayek's polemic, *The Road to Serfdom* (Chicago, 1944).

[17] H. D. Dickinson, "Price Formation in a Socialist Economy," *Econ. Jour.*, Dec. 1933, XLIII, 237–50; O. Lange, "On the Economic Theory of Socialism" in Lange and Taylor, *The Economic Theory of Socialism*, B. E. Lippincott, ed. (Minneapolis, 1938); A. P. Lerner, *op. cit.*; M. Dobb, "Economic Theory and the Problem of the Socialist Economy," *Econ. Jour.* Dec. 1933, XLIII, 588–98.

[18] N. Kaldor, "Welfare Propositions in Economics and Interpersonal Comparisons of Utility," *Econ. Jour.*, Sept. 1939, LXIX, 549–52.

[19] J. R. Hicks, "The Foundations of Welfare Economics," *Econ. Jour.*, Dec. 1939, LXIX, 696–712 and "The Valuation of the Social Income," *Economica*, Feb. 1940, VII, 105–23.

[20] L. Robbins, *The Nature and Significance of Economic Science* (London, 1932).

[21] T. Scitovsky, "A Note on Welfare Propositions in Economics," *Rev. Econ. Stud.*, 1941–1942, IX, 77–78.

———, "A Reconsideration of the Theory of Tariffs," *Rev. Econ. Stud.*, 1941–1942, IX, 89–110.

[22] P. A. Samuelson, "Evaluation of Real National Income," *Oxford Econ. Papers*, Jan. 1950, II, 1–29.

[23] I. M. D. Little, *A Critique of Welfare Economics* (Oxford, 1950).

[24] H. Hotelling, "The General Welfare in Relation to Problems of Taxation and of Railway and Utility Rates," *Econometrica*, July 1938, VI, 242–69, is the first modern formulation of the problem that was posed, in 1844, by Jules Dupuit ("On the Measurement of Utility of Public Works," to be found in *International Economic Papers*, No. 2, ed. Alan T. Peacock *et al.*).

[25] See esp. Little, "Direct versus Indirect Taxes," *Econ. Jour.*, Sept. 1951, LXI, 577–84.

[26] For a comprehensive treatment of the issues, as well as for references, see J. E. Meade, *The Theory of International Economic Policy*, Vol. II: *Trade and Welfare* and *Mathematical Supplement* (New York, 1955).