SPATIAL PRICE EQUILIBRIUM AND LINEAR PROGRAMMING

By PAUL A. SAMUELSON*

I.—*Introduction*

Increasingly, modern economic theorists are going beyond the formulation of equilibrium in terms of such marginal equalities as marginal revenue equal to marginal costs or wage rate equal to marginal value product. Instead they are reverting to an earlier and more fundamental aspect of a maximum position: namely, that from the top of a hill, whether or not it is locally flat, all movements are downward. Therefore, the real import of marginalism is embodied in the following type of statement: for any produced units of output, extra revenues exceed extra costs; but for any further producible units, extra revenue would fall short of extra costs. These marginal inequalities—which need not apply to small local movements alone—do, in well-behaved cases with smooth slopes, imply the usual marginal equalities. But they are more general, in that from them we can derive most of what is potentially useful in marginal analysis, a point which has been missed by both the defenders and attackers of "marginalism." And more than that, the marginal inequalities can apply to cases (like simple comparative advantage) where the marginal equalities fail.

In recent years economists have begun to hear about a new type of theory called *linear programming*. Developed by such mathematicians as G. B. Dantzig, J. v. Neumann, A. W. Tucker, and G. W. Brown, and by such economists as R. Dorfman, T. C. Koopmans, W. Leontief, and others, this field admirably illustrates the failure of marginal equalization as a rule for defining equilibrium. A number of books and articles on this subject are beginning to appear. It is the modest purpose of the following discussion to present a classical economics problem which illustrates many of the characteristics of linear programming. However, the problem is of economic interest for its own sake and because of its ancient heritage.

The first explicit statement that competitive market price is determined by the intersection of supply and demand functions seems to have been given by A. A. Cournot in 1838 in connection, curiously enough, with the more complicated problem of price relations between

* The author is professor of economics at the Massachusetts Institute of Technology and consultant to the Rand Corporation, whose help in this research is acknowledged.

two spatially separate markets—such as Liverpool and New York.[1] The latter problem, that of "communication of markets," has itself a long history, involving many of the great names of theoretical economics.[2] Dr. Stephen Enke in a recent interesting paper generalized the problem of interspatial markets and gave it an elegant solution.[3]

Proceeding from the Enke formulation, I propose in this paper (1) to show how this purely descriptive problem in non-normative economics can be cast mathematically into a *maximum* problem; and (2) to relate the Enke problem to a standard problem in linear programming, the so-called Koopmans-Hitchcock minimum-transport-cost problem.[4]

Spatial problems have been so neglected in economic theory that the field is of interest for its own sake. In addition, this provides a reasonably easy-to-understand example in the new field of linear programming. But, most important, insight into the fundamental nature of economic pricing is provided by the present discussion.

II.—*Formulation of the Two Problems*

In the Cournot-Enke problem, we are given at each of two or more localities a domestic demand and supply curve for a given product (*e.g.*, wheat) in terms of its market price at that locality. We are also given constant transport costs (shipping, insurance, duties, etc.) for carrying one unit of the product between any two of the specified localities. What then will be the final competitive equilibrium of prices in all the markets, of amounts supplied and demanded at each place, and of exports and imports?

From the description of the above problem, an economist would be tempted to guess that it includes inside it the following Koopmans problem, which I slightly reword to bring out the similarity: A specified total number of (empty or ballast) ships is to be sent out from each of a number of ports. They are to be allocated among a number of other receiving ports, with the total sent in to each such port being specified. If we are given the unit costs of shipment between every two ports, how can we minimize the total costs of the program?

[1] A. A. Cournot, *Mathematical Principles of the Theory of Wealth* (1838), Chap. X.

[2] J. Viner, *Studies in International Trade*, New York, 1937, pp. 589–91 gives references to Cunyngham (1904), Barone (1908), Pigou (1904), and H. Schultz (1935); for a non-graphic literary exposition, see F. W. Taussig, *Some Aspects of the Tariff Question*, Cambridge, Mass. (1915 and 1931), Chap. I.

[3] S. Enke, "Equilibrium Among Spatially Separated Markets: Solution by Electric Analogue," *Econometrica*, Vol. 19 (Jan., 1951), pp. 40–47.

[4] See T. Koopmans, *Activity Analysis of Production and Allocation* (Monograph 13 of the Cowles Commission, published by John Wiley, 1951) for references. The special transport problem itself is dealt with in Chapters XIV and XXIII and independently deals with a problem considered in 1941 by F. L. Hitchcock and in 1942 by a Russian mathematician, L. Kantorovitch. For a readable account, see T. C. Koopmans, "Optimum Utilization of the Transportation System," *Econometrica*, Vol. 17, Suppl. (July, 1949), pp. 136–46.

Note that total shipments in or out of any one place are an unknown in the first problem, whereas they are given in the second. In this sense the first problem is the more general one and includes the second inside itself. Note, too, that, as it stands, the first problem is one of decentralized price-mechanics: innumerable atomistic competitors operate in the background, pursuing their private interests and taking no account of any centralized magnitude that is to be maximized. Yet, even without Adam Smith's "as-if" principle of the Invisible Hand, our teleological faith in the pricing mechanism is such that we should be surprised if the resulting allocations resulted in costly cross-haulages: we instinctively feel that arbitragers could make money getting rid of any such inefficiencies.

A final hint suggests that the first problem, which is definitely not a maximum problem to begin with, might be convertible into a maximum problem. Enke provides a simple ingenious electric circuit for its solution in the case of linear market functions. At least since the work of Clerk Maxwell and Kirchhoff a century ago it has been realized that the equilibrium of simple passive electric networks can be described in terms of an extremum principle—the minimization of "total power-loss."[5]

It is not surprising, therefore, that the Enke problem can be artificially converted into a maximum problem, from which we may hope for the following specific advantages: (1) This viewpoint might aid in the choice of convergent numerical iterations to a solution. (2) From the extensive theory of maxima, it enables us immediately to evaluate the sign of various comparative-statics changes. (*E.g.*, an increase in net supply at any point can never in a stable system *decrease* the region's exports.) (3) By establishing an equivalence between the Enke problem and a maximum problem, we may be able to use the known electric devices for solving the former to solve still other maximum problems, and perhaps some of the linear programming type. (4) The maximum problem under consideration is of interest because of its unusual type: it involves in an essential way such non-analytic functions as absolute value of X, which has a discontinuous derivative and a *corner;* this makes it different from the conventionally studied types and somewhat similar to the inequality problems met with in linear programming. (5) Finally, there is general methodological and mathematical interest in

[5] In its simplest form, such a minimum problem is of conventional interior differentiable ("Weierstrassian") type, and it does not involve the quasi-linear boundaries, inequalities, and vertexes encountered in linear programming. Nonetheless, A. W. Tucker in an unpublished Office of Naval Research memorandum "Analogues of Kirchhoff's Laws" (Stanford, 1950) noted the similarity of the Kirchhoff-Maxwell problem to the linear programming problem of the Koopmans type. Moreover, I gathered from personal conversation with Professor Tjalling Koopmans that when he first solved the transportation problem years ago, before linear programming had been explicitly formulated, the analogy with the network problem readily occurred to him and helped guide his explorations toward a solution. See *Activity Analysis, op. cit.* pp. 258–59.

the question of the conditions under which a given equilibrium problem can be significantly related to a maximum or minimum principle.

III.—*The Two-Locality Case Graphically Treated*

The two-variable case provides a convenient introduction to the principles involved. The general n variable case then follows without much difficulty. Figure 1 shows the usual textbook back-to-back diagram determining the equilibrium flow of exports from market 1 to 2.

EQUILIBRIUM OF EXPORTS AND IMPORTS

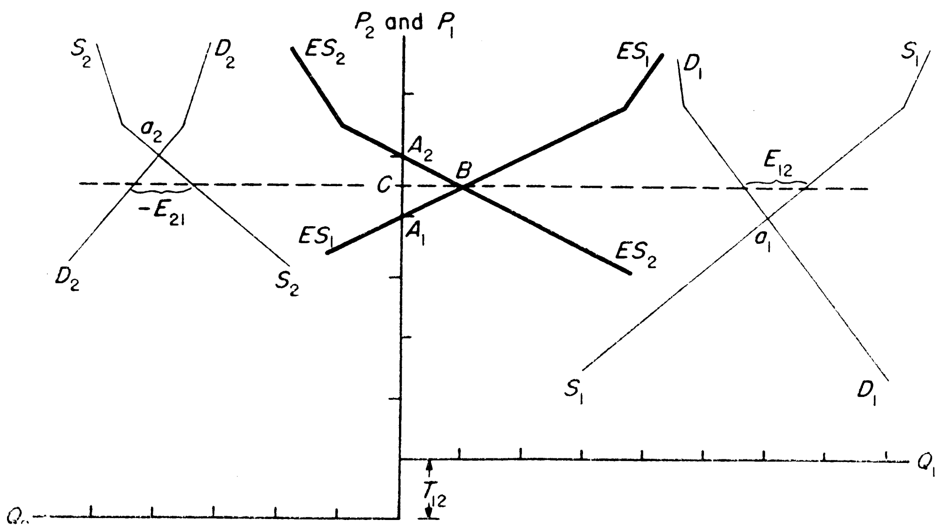

FIGURE 1. Equilibrium is at B where exports of 1 match imports of 2 at the differential between P_2 and P_1 equal to transport costs, T_{12}. Note shift in lower axis of 2.

Before trade, equilibrium would be at $P_1 = A_1$ where supply and demand in the first market just meet; or what is the same thing, where the excess-supply function ES_1 to ES_1, which is equal to the demand curve subtracted laterally at every price from the supply curve, is at its zero point. Likewise, $P_2 = A_2$ if no trade is possible.

But now suppose that goods can move from 1 to 2 for T_{12} dollars per unit, and from 2 to 1 for T_{21} dollars per unit. Since the pretrade price is lower in 1 than 2, trade will obviously never flow from 2 to 1 and so only the T_{12} figure is relevant. Because the initial differential in prices exceeds the transport costs, there will be a positive flow of exports from 1 to 2, and P_2 will come to exceed P_1 by exactly T_{12}. For this reason the axes of market 1 have been displaced relative to those of market 2 by the distance T_{12}.

The new equilibrium is shown at B, where the excess supply or exports of market 1 exactly equal the algebraically negative excess supply or imports of market 2. The bracketed distances E_{12}, and $-E_{21}$, and CB are all exactly equivalent depictions of these flows.

Of course, if A_1 and A_2 had been closer together than T_{12} (or T_{21}), then the markets would have split-up and the separate equilibria would be at (A_1, A_2). Had A_2 been less than A_1 by more than T_{21}, then the flow of exports would have automatically reversed directions so that E_{21} would be positive and E_{12} negative. What makes the problem interesting is its complicated non-linear equilibrium conditions:

(1)a $$\text{If} \quad P_2 = P_1 + T_{12},$$

any non-negative E_{12} may flow; $E_{12}>0$ implies $P_2=P_1+T_{12}$.

(1)b $$\text{If} \quad P_2 < P_1 + T_{12} \quad \text{and if} \quad P_1 < P_2 + T_{21},$$

then $E_{12}=0$ and $E_{21}=0$.

(1)c $$\text{If} \quad P_1 = P_2 + T_{21},$$

then $E_{21} \geqq 0$, depending upon total world supply and demand, etc.; $E_{21}>0$ implies $P_1=P_2+T_{21}$.

Figure 2 provides a new graphical restatement of what is shown in Figure 1. The same excess-supply curves are shown but this time the prices in the two countries are measured *from the same level* rather than with one axis shifted by the amount of transportation cost. However, the transport costs do enter into the problem. Look first at the two upper curves of the figure only. Now the final equilibrium is determined at JK where the two excess-supply curves differ vertically by T_{12}, as shown by the bracket.

This same equilibrium determination of exports and imports (E_{12} and $-E_{21}$) is shown in the lower part of the figure by the heavy NN curve which represents the *vertical* difference between the two excess supply curves. The final equilibrium is at F where the net excess-supply curve hits the curve of discontinuous transport costs $WXYZ$.

Figure 2 has the merit of suggesting how a shift in one or both of the excess-supply curves could cause the equilibrium intersection to be shifted over to the WX interval, with 2 exporting to 1 as in equation (1)c. It also suggests how the intersection *might* be on the XY interval, with exports and imports zero, and with prices related as in (1)b.

IV.—*Defining Social Pay-off*

Figure 2 paves the way toward a maximum or minimum formulation of the problem. An economist looking at these figures would naturally think of some kind of consumers surplus concept. The area A_1JKA_2, and $OMFG$, its equivalent, cry out to be compared with the area under the transport curve, $OMFY$. However, the name consumers surplus has all kinds of strange connotations in economics. To avoid these and to underline the completely artificial nature of my procedure, I shall simply define a "net social pay-off" function, with three components:[5a] NSP = Social Pay-off in 1 + Social Pay-off in 2 − Transport Cost.

The social pay-off of any region is defined as the algebraic area under its excess-*demand* curve. This is equal in magnitude to the area under

EQUIVALENT DEPICTION OF SPATIAL EQUILIBRIUM

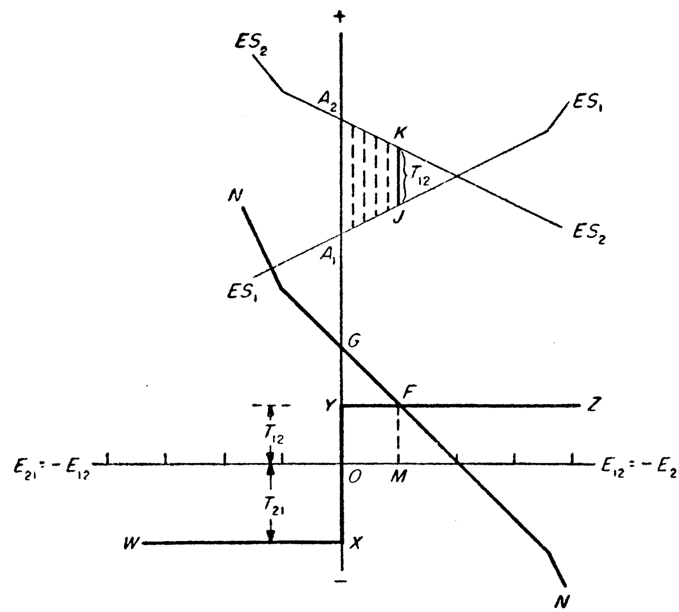

FIGURE 2. The same excess-supply curves appear as in figure 1, but now lower axes are evenly aligned. Transport costs enter through the discontinuous curve $WXYZ$. Equilibrium is where the *net* excess-supply curve for the two markets—NN = vertical subtraction of ES_1 from ES_2—intersects $WXYZ$ at F. (Alternatively, equilibrium is where JK vertical distance equals T_{12}.)

its excess-supply curve but opposite in algebraic sign. However, since the second market has been put back-to-back to the first, the area under the second market's excess-supply curve in Figure 2 does measure the second market's pay-off; and from it we must *subtract* the area under the first country's excess-supply curve. *Hence, the area under the net curve NN in Figure 2 does perfectly measure the combined social pay-off of both markets.*

[5a] This magnitude is artificial in the sense that no competitor in the market will be aware of or concerned with it. It is artificial in the sense that after an Invisible Hand has led us to its maximization, we need not necessarily attach any social welfare significance to the result.

In Figure 3 the curve NON indicates how the combined payoff of the two markets varies with algebraic exports from 1 to 2. From this we subtract the curve of total transport cost UOU. Total transport cost has a corner at the origin because of the discontinuity between T_{12} and T_{21} as shown in $WXYZ$ of Figure 2; the algebraic integral of this discontinuous function leads to the V-shaped total function UOU.[6] We

MAXIMIZING NET SOCIAL PAY-OFF

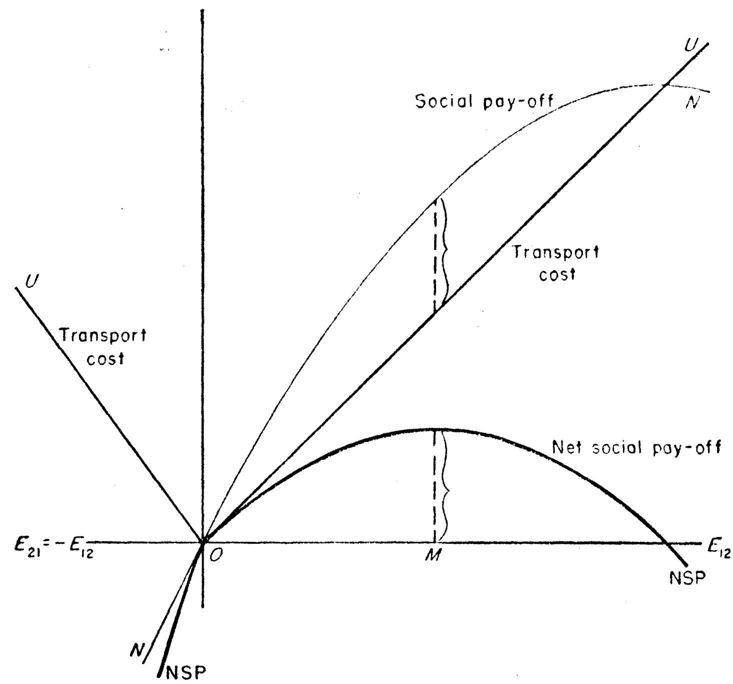

FIGURE 3. Same equilibrium as in previous figures is shown as the maximum of net social pay-off or maximum vertical difference between upper two curves.

find our equilibrium where the vertical distance between the two upper curves is at a maximum. This same optimal level of exports (E_{12}) or imports ($-E_{21}$) is also shown at the maximum point of the net social pay-off curve, the third curve which measures the vertical distance between the upper two.

This completes the two-variable case. Figure 4 illustrates that three possible cases could have emerged, corresponding to equations (1)a,

[6] The mathematical symbolism for this function can be written in many equivalent ways: *e.g.*

$$t_{12}(E_{12}) = T_{12}E_{12} \text{ for } E_{12} \geqq 0$$
$$= T_{21}(-E_{12}) \text{ for } E_{12} \leqq 0$$

and still other equivalent symbolisms (involving absolute values of E_{12}, etc.) can be found.

(1)b, and (1)c. In Figure 4a, region 1 exports to 2 so that the maximum point is smooth; the corner in the curve, due to the discontinuity in the rate of transport cost, is on the vertical axis and does not affect the maximum. Similarly Figure 4c shows a normally smooth maximum

TYPES OF MAXIMA FOR NET SOCIAL PAY-OFF

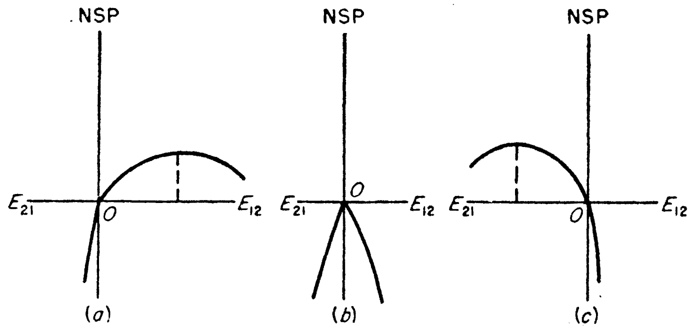

FIGURE 4. Because of transport costs, each curve of net social pay-off has a corner at the origin. In (b) price differentials are too little to surmount transport costs, so trade is zero.

without corners; 2 is then exporting to 1. Figure 4b shows the intermediate case where the maximum point is a cusp with a corner: the transport cost discontinuity is obviously to blame.

* * *

The final point to emphasize is this: Once the separate pay-off functions are set up as areas or integrals of the excess-supply curves and once transport-cost functions are known, a clerk could be given the task of experimentally varying exports so as to achieve a net maximum. He could proceed by trial and error, always moving in a direction that increased the net pay-off, and he would ultimately arrive at the correct equilibrium. The existence of pathological corners would not impair convergence; rather it might accelerate convergence.[7]

[7] Mathematically, calling the excess-supply functions $s_i(E_i)$,

$$NSP = -\int_0^{E_{12}} s_1(x)dx - \int_0^{-E_{12}} s_2(x)dX - t_{12}(E_{12}).$$

By setting $dN/dE_{12}=0$, we arrive at conditions equivalent to equations (1):

$$(1) \qquad -T_{21} \leqq s_2(E_{21}) - s_1(E_{12}) \leqq T_{12}, \quad E_{12}+E_{21}=0$$

and with $E_{12}\neq 0$ implying that one of the equality signs holds. A gradient method of making dE_{12}/dt proportional to dN/dE_{12} would always converge for positive sloping s_i functions.

V.—*The General Case of Any Number of Regions*

Instead of two regions suppose we have $i=1, 2, \cdots, n$ regions. The algebraic amount of exports from i to j we can write as E_{ij} and it will of course be the same thing as the algebraic imports of j from i, $-E_{ji}$. Table I shows the two-way table relating these interregion exports to the total algebraic exports of any region, E_i. (Note that about half the numbers in the table will be negative.) Suppose too that we are given T_{ij}, the transport cost per unit of product moved from region i to region j. These transport costs could also have been arranged in a two-way table. Finally suppose that for each region we have an excess supply function $s_i(E_i)=P_i$, which is calculated by taking the lateral difference between local supply and demand functions.

TABLE I

Region	1	2	$\cdots$	j	$\cdots$	n	Total exports (algebraic)
1	$\cdot$	$E_{12}=-E_{21}$		E_{1j}	$\cdots$	$E_{1n}=-E_{n1}$	$E_1=\sum_j E_{1j}$
2	E_{21}	$\cdot\cdot$		$\cdot$		$\cdot$	$E_2=\sum_j E_{2j}$
$\cdot$							$\cdot$
$\cdot$							$\cdot$
$\cdot$							$\cdot$
i	E_{i1}	$\cdot$		$E_{ij}=-E_{ji}$		E_{in}	$E_i=\sum_j E_{ij}$
$\cdot$							$\cdot$
$\cdot$							$\cdot$
$\cdot$							$\cdot$
n	E_{n1}	$\cdot$		E_{nj}		$\cdot$	$E_n=\sum_j E_{nj}$
Total imports (algebraic)	$-E_1=\sum_i E_{i1}$	$-E_2=\sum_i E_{i2}$	$\cdots$	$-E_i=\sum_i E_{ij}$	$\cdots$	$-E_n=\sum_i E_{in}$	$0=\sum_i\sum_j E_{ij}$ = total net exports

As in the two-variable case, we can define a social pay-off for every region in terms of the area under the excess-demand or excess-supply function. This will depend only upon the total exports of the region and can be written

$$S_i(E_i) = \text{area under the excess-demand curve} = -\int_0^{E_i} s_i(x)\,dx. \tag{2}$$

The transport costs can be written as a function of exports between any two regions; or as $t_{ij}(E_{ij})$ = a V-shaped curve like that shown in Figure 3.[8]

Now we can form a final net social pay-off for all the regions as the sum of the n separate pay-offs minus the total transport costs of all the shipments:

[8] Mathematically, this function has a corner at the zero export point, and will equal $T_{ij}E_{ij}$ if i exports to j and $T_{ji}E_{ji}$ if j exports to i. Note T_{ij} need not equal T_{ji}, but it can be shown that by definition $t_{ij}(E_{ij})=t_{ji}(E_{ji})$.

$$(3) \qquad NSP = \sum_{1}^{n} S_i(E_i) - \sum_{i<j}\sum t_{ij}(E_{ij}).$$

Because $E_i = E_{i1} + \cdots + E_{in}$, this is a function of all the E_{ij}'s and when we have found its maximum we have arrived at the final unique equilibrium trade pattern.[9]

Providing that all domestic supply curves cut demand curves from below (as price rises), which is the so-called case of normal or stable intersection, the excess-supply curves will never be falling curves; and it will necessarily follow that the maximum position will exist and be unique. At the maximum point, we will find

$$(4) \qquad -T_{ij} \leqq s_i(E_i) - s_j(E_j) \leqq T_{ji} \qquad \text{(for all } i, j = 1, \cdots, n)$$

with both inequalities holding only if $E_{ij} = 0 = E_{ji}$; if $E_{ij} > 0$, then the right-hand equality must hold; and if $E_{ji} < 0$, the left-hand equality must hold. Recalling that the s_i's and P's are the same thing, we obviously end up with the proper n-region generalization of equations (1) that were derived for the two-region case.

Our task has now been successfully completed. The problem in descriptive price behavior has been artificially converted into a maximum problem. This maximum problem can be solved by trial and error or by a systematic procedure of varying shipments in the direction of increasing social pay-off.[10]

* * *

Once the exports are determined between any two places, it is obvious that the total exports of any and every place are also determined. Some of the n-regions will end up as positive net exporters; some will end up as net importers (negative net exporters); some may even end up in perfect balance, with zero imports and exports. Reflection will show that we are free to omit all such balanced regions from our further dis-

[9] Since we know that the imports of one region are the exports of another, we do not have to specify all the $n^2 E_{ij}$'s in Table I. Instead we can work with all those that are above the hollow diagonals of the Table, inferring those below by using the identity $E_{ij} = -E_{ji}$. Thus, we may adopt the convention of having $i<j$, and may work with $n(n-1)/2$ unknown E_{ij}'s. Incidentally, most of the E_{ij}'s will turn out to be zero: *i.e.*, a typical export region will export to only a few other regions and a typical importer will import from only one or a few regions. More exactly, it is a theorem of linear programming that the number of positive exports need not exceed $n-1$. It will also be true that if T_{ij} and E_i are all integers, the E_{ij}'s will all be integers.

[10] Even if this were not a market problem, we could set up pretended competitive markets whose supply and demand relations might be used to help compute the correct mathematical solutions. Computing clerks could be instructed to act like brokers and arbitragers, etc. Or we could dispense with all markets, and instead watch how NSP is changing as we change each E_{ij}, continuing to move always in the direction of increasing NSP. Doing this long enough will carry us to the top of the hill.

For the case where the excess-supply curves are straight lines Enke has given a simple electric circuit, consisting of resistances, rectifiers, and batteries, which will give the final solution as a measurement of currents and voltages. See Enke, *op. cit.*, Figure 1, p. 45 and the next section below.

cussion, since so long as they remain in balance they need not export or import from any locality.[11]

It follows that we can divide our n-regions up into $i=1, \cdots, m$ export regions and $i=m+1, \cdots, n$ import regions. Reflection will show that a net import region will never export to any region. (Why send exports out if you have to expensively ship in imports to replace them? Instead ship directly.) Reflection also shows that a net export region will never import from any region. Thus, the only non-zero E_{ij}'s are from an export region i to an import region j.

What does this mean for Table I? It means that we can label our regions so as to give the first m numbers to our export regions and the

TABLE II

Regions	$1 \cdots m$	$m+1,$	$m+2$	$\cdots$ n	Total Exports
exporters					
1		$E_{1,m+1}$	$E_{1,m+2}$	$\cdots$ $E_{1,n}$	(E_1)
2	zeros	$E_{2,m+1}$	$E_{2,m+2}$	$E_{2,n}$	(E_2)
.		.			
.		.			
.		.			
m		$E_{m,m+1}$	$E_{m,m+2}$	E_{mn}	(E_m)
importers					
$m+1$					
.	redundant				
.	negative		zeros		
.	numbers				
n					
Total Imports		$(-E_{m+1})$	$(-E_{m+2})$	$\cdots(-E_n)$	

last $n-m$ numbers to the import regions. Table I will then divide itself up into 4 major blocks: positive numbers will then be in only the upper right-hand block relating the exporting countries to the importing countries; the two blocks relating exporters to exporters and importers to importers will necessarily be full of zeros, and can be neglected. The block relating importers to exporters will consist of negative numbers only and will simply duplicate our positive numbers.

[11] That is $E_i=0$ implies $E_{ij}=0=E_{ji}$ for all j. Note that I am adopting the convention of not treating cargo shipped *through* a port as at all part of its exports or imports. Thus, cargo going from London to San Francisco is not to be treated as both an import and export of Panama. A similar philosophy tells me that T_{ij} data have already been adjusted so that it is no longer cheaper to send cargo from i to j via a third port k: such an "indirect" route would already have been defined to be the cheapest direct route and by our convention the port k would not be explicitly involved. Such preliminary adjustments of the definition of T_{ij} have made it satisfy the "Pythagorean" relations $T_{ij} \leqq T_{ik}+T_{kj}$, etc.

Table II shows the relevant configuration. Note that all the numbers shown by symbol are positive: Thus $(-E_n)$ represents positive imports because algebraic exports E_n are negative for an importing country. Note that a table of T_{ij}'s is definable for *all* countries and all blocks; but with the given positive exports and imports shown in the margins of the table within parentheses, we would be interested only in the T_{ij}'s in the upper right-hand corner.

VI.—*Relation to the Koopmans Linear Programming Problem*

In this and the next section, I shall try to relate the results of our international trade problem to the newly developed theory of linear programming (defined austerely by the mathematician as "the maximization of a linear expression subject to linear inequalities"). For the theorist these are important sections, and to an economist who has heard about this new field and would like to get an idea of what it is all about, these will serve to indicate its general flavor. Though I have tried to use only the most elementary tools, these two sections are not summer-hammock reading; and for this reason, the remaining sections, from VIII on, have been arranged so that a reader can go directly on to them, skipping the more technical material.

Imagine now that the positive totals in parentheses $(E_1, \cdots, E_m)$ and $(-E_{m+1}, \cdots, -E_n)$ were given to us by Enke while at the same time he concealed from us the entries $E_{i,m+j}$ giving the detailed breakdowns. Then for us to find the missing numbers so as to minimize total transport cost would be precisely the Koopmans-Hitchcock problem in linear programming.[12]

How do we know that Enke's solution for E_{ij}'s does truly minimize transport cost? Since I have shown that Enke does maximize net social pay-off, and since the expressions $\sum_1^n S_i(E_i)$ in NSP of equation (3) depend only on the regional totals of exports $(E_1 \cdots, E_n)$, it follows that maximizing NSP would be impossible unless the E_{ij}'s were optimal for minimizing transport cost. Thus the Cournot-Enke problem does have inside it the Koopmans problem.[13]

* * *

[12] We minimize

$$\sum_{i=1}^{m} \sum_{j=m+1}^{n} T_{ij}E_{ij}$$

subject to

$$\sum_{j=m+1}^{n} E_{ij} = E_i \qquad (i = 1, \cdots, m)$$

and

$$\sum_{i=1}^{m} E_{ij} = (-E_j) \qquad (j = m+1, \cdots, n) \qquad E_{i,m+j} \geqq 0.$$

[13] A close analogy is the Yntema-Robinson problem of determining best outputs for a monopolist discriminating among independent markets. This includes inside it the problem of best allocating a given total output among markets.

Despite its likeness to the Kirchhoff-Maxwell quadratic maximum property of electric networks, the Koopmans problem cannot be solved by simple Kirchhoff networks unless entirely new laws of resistance can be inserted into such a network.[14] However, we can utilize the Enke network, which uses standard resistances and rectifiers, to solve any Koopmans problem. But we must work backwards: we must experimentally vary Enke's A_i's, which he puts into his network as prescribed voltages, until we achieve the requisite n totals $(E_1, \cdots, E_m, -E_{m+1}, \cdots, -E_n)$.[15] We can read off the required interregional exports as electric currents from the resulting Enke analogue network.

The above method is of interest because it suggests that any linear programming problem may be solvable as an unconstrained extremum problem provided we can imbed it in a suitably generalized problem.[16] Since the Air Forces and Bureau of Standards have set up their electronic calculators so as to solve transportation problems in relatively short time, there is no need to pursue the computational advantage of analogue networks any farther here.

VII.—*Equilibrium Prices and the Dual Problem*

The equilibrium prices as given by equations (4) arise naturally in the Enke problem but are completely absent from the initial formulation of the transport problem. However, as an economist, Koopmans sought to introduce some kind of price mechanism into the calculation

[14] See *Activity Analysis*, p. 259.

[15] This can always be done. The theory of the "dual problem" in linear programming assures us that there exists a set of parallel shifts of the excess-supply curves that will bring about any prescribed E_i configuration. The slopes of the excess-supply curves, Enke's b's can be arbitrary positive numbers. The procedure here described has the drawback of involving in effect a need to solve a set of simultaneous equations between the A's and E's. This might be mechanized by servomechanisms. (Query: does the "dual network" to Enke's solve the "dual problem" of linear programming?)

[16] Thus, if we use ordinary matrix notation and are given m arbitrary c's and $m(\geqq n)$ arbitrary x's, suppose we seek to maximize $Z=b'x$ subject to $Ax=c$ and $x\geqq 0$. Then for any nxn positive definite matrix B, there exists an n-vector v such that

$$Z^* = c'Bc - v'c + b'x = x'A'BAx - v'Ax + b'x$$

will be at a maximum for $x\geqq 0$ only if $Ax=c$. Having somehow found such unknown v's, we can solve Z by solving Z^*, an unconstrained extremum problem. For special A's and B's this can be converted into a simple analogue network problem. In every case we can solve the dual problem by taking proper combinations of the optimal derivatives of Z^*. Actually Z^* need not be quadratic but can be more general. A natural generalization of the Koopmans problem would be to have the flow of shipments vary directly with the marginal costs of transport $T_{ij}+P_i-P_j$. This would differ from the Enke problem in that E_i would then be a function of all P's and not of P_i alone. Dr. Martin Beckmann has written a number of interesting Cowles Commission Memoranda dealing with the case of continuous markets located everywhere in the plane. The partial differential equations of equilibrium correspond very closely to those of the Koopmans problem and a potential function plays a similar rôle. For the continuous case too, we can imbed the problem into a more general situation in which there is an excess-supply function at every point.

and he did succeed in defining a set of prices or potentials that correspond to our s_i's and P_i's. In doing this he was guided by economic and electric analogies and he was able to anticipate the mathematicians' theories of what is called the "dual problem" in linear programming. Mathematical theory assures us that every minimum problem in linear programming can be, so to speak, turned on its side and can be converted into a related maximum problem. The answer to this maximum problem also gives the correct answer for the quantity that was to be minimized.

For sake of brevity I shall simply refer the reader to the theory of linear programming for interpretations and proofs of the following remarks: (1) The P's or s's of equation (4) are the dual variables to the E_{ij}'s of the transport problem;[17] (2) $P_i - P_j$ can be interpreted in the Koopmans problem as an element of indirect marginal cost to be added to direct marginal cost T_{ij}—this term takes into account the money advantages of having an empty ship at j rather than at i; (3) If we change some E_i's, then the resulting increase in total transport cost is related to the P's; (4) Pricing in a competitive market could conceivably keep a system in the proper optimal configuration.[17a]

[17] Defining the profitability of any export from i to j as $\pi_{ij} = P_j - P_i - T_{ij}$, then for all $\pi_{ij} \leqq 0 \leqq P_k$, the maximum of $Z = -\sum P_k E_k$ will equal the minimum achievable transport costs. For any $\pi_{ij} < 0$, $E_{ij} = 0$, and for any $E_{ij} < 0$, $\pi_{ij} = 0$. Z can be rewritten $\sum\sum E_{ij}(P_j - P_i)$ so that we can think of the problem as that of finding the price differentials that will maximize the total *gain* in value of amounts shipped, subject to non-positive profitabilities on each shipment. It may be mentioned that only price differences are determinate and that the "dual variables" corresponding to export regions are related to their prices.

[17a] The present techniques, generalized in the manner indicated by cited work of Beckmann, can be used to throw light on many of the problems arising in connection with the basing point controversy. Oligopolists selling a homogeneous product want some pattern of pricing that will lead to a single *unambiguous price* at any geographical point. Usually, too, the pattern must be such as to lead to a not-too-obviously-wasteful flow of transport, and it must provide for a fairly stable and not-too-lopsided sharing of the market by the different producers. Historically, so long as most production is in fact concentrated in one advantageous place, such as Pittsburgh, a single basing pattern meets these criteria and is not too inefficient. Such a pattern of course encourages consumers to move toward the base; but to the extent that customers do not move, it encourages producers to move out toward them so as to receive in the form of higher net prices the transport costs saved or "phantom freight."

In any case, in a dynamic world any one locality will usually lose its dominant advantages in time and many plants will be operating away from the single basing point. Consequently, the system will lead increasingly to an inefficient pattern of transport and will become increasingly vulnerable to public criticism and to competition. A multiple basing point system may then come into effect and this will represent a compromise between FOB pricing and delivered pricing: within each region, the basing point pattern will prevail but at the shiftable boundaries of the regions there may be price competition between regions.

The requirements for the regional pattern of prices and flow of commodities is so strictly determined by perfect-competition assumptions that it is easy to show how basing point patterns deviate from the spirit and the letter of the competitive pattern: thus when producers all over the country respect a single basing point pattern, the contour lines of equal-delivered-price surround the basing point in a manner superficially similar to a perfect-competition pattern; but the flow of transports is *not* perpendicular to the contour lines of price and so we instantly detect the absence of perfect competition. For any given pattern of producer and con-

VIII.—*Illustrative Case Study*

A numerical example may help to clarify the theoretical relations. Unless there are at least four regions, no problem of optimizing the pattern of exports can arise. So consider regions I, II, III, IV characterized by the following transport costs (in $ per unit of shipment).

$$\text{A} \qquad \left[\begin{array}{cc:cc} \cdot & T_{12} & T_{13} & T_{14} \\ T_{21} & \cdot & T_{23} & T_{24} \\ \hdashline T_{31} & T_{32} & \cdot & T_{34} \\ T_{41} & T_{42} & T_{43} & \cdot \end{array}\right] = \left[\begin{array}{cc:cc} \cdot & 9 & 8 & 7 \\ 4 & \cdot & 4 & 1 \\ \hdashline 5 & 1 & \cdot & 5 \\ 6 & 2 & 6 & \cdot \end{array}\right]$$

Let us suppose that initial local conditions of demand and supply are such as to make Regions I and II exporters and III and IV importers. Suppose further that I initially exports only to III while II exports to both III and IV. In that case the data in the upper right-hand block of the transport cost array are alone relevant. The resulting export shipments are perhaps as shown in the following table:

B

Regions	III	IV	Exports	Local Price
I	38	0	(38)	($10)
II	2	50	(52)	$14
Imports	(40)	(50)	90	
Local Price	$18	$15		

If only the data in parentheses in Table B were given, we could from the cost data in Table A deduce the remaining numbers in B. To do so

sumer location and observed flow of product, linear programming permits us to calculate the optimal minimum cost of transport and to compare it with the actual.

The optimum pattern never permits of overlapping markets of differently located plants; so it is obvious that an omnipotent combine of oligopolists would never use the arbitrary pattern of price discrimination implied by basing points. Nonetheless, there is nothing at all surprising about the use of this pattern; for after all an omnipotent combine of oligopolists would not engage in competitive advertising and many other things which we expect actual imperfect competitors to engage in. Except in times of formal or informal price control, imperfect competitors customarily administer their prices at more than marginal cost; and since perfect non-discrimination is only one out of an infinity of patterns, we should not be at all surprised by the existence of price discrimination. Moreover, contrary to some views in the literature, the pattern of price discrimination implied by adhering to basing point formulas is *not* particularly a strange or arbitrary one: it can be shown to presuppose that the elasticity of demand of customers for the delivered output of the basing point mills increases slightly with distance from the basing point, an assumption that no one would defend as exact but also an assumption that no assistant vice-president would be motivated to denounce strongly. Under oligopoly the most precious of all devices are those that lead to an informal consensus and the basing point system has the great virtue that within a region it reduces the whole pattern of pricing *to a single unknown*, around which sentiments can form and in terms of which mere price-competing is obviated.

we would have to solve the Koopmans problem and its so-called dual problem; however, in solving the more general Enke problem, all P's would already have been determined by supply and demand conditions, as would have all the unknowns of the problem.

To understand the necessary pattern of price, Figure 5a is useful. It represents what the mathematical topologist calls a "tree," and it shows the flow of exports by means of arrows. Note that all four markets are connected and that competition will freeze all P's as soon as any one P in the tree is known.

FOUR-REGION TRADE EQUILIBRIA

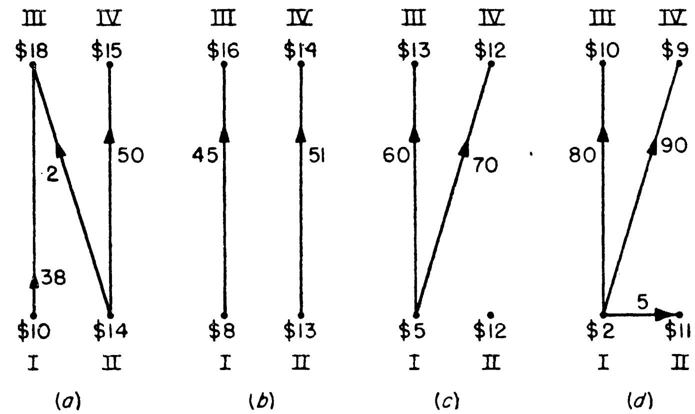

FIGURE 5. The changing network of trade as excess-demand at 1 increases. (Arrows show direction of trade while nearby numbers show its magnitude. Dollar figures refer to regional prices.)

Now let us suppose that there is an increase in domestic supply and a decrease in domestic demand in I, so that I's excess supply function increases. It will be shown in the next section that this must depress P_1. At first, therefore, the *qualitative* pattern of trade will be unchanged; and hence, all P's must decrease by exactly the same amount as P_1 decreases. (Incidentally, algebraic exports of *all* other regions must at first go down if their markets are to be cleared.)

However, if the excess-supply function at I increases enough, region II's export market to III may be completely captured by I. This case is shown in Figure 5b. Our tree has now split off into two trees, or what the mathematician calls a "forest." Now P_2 and P_4 are interconnected but they are independent of P_1 and P_3. This is called the case of "degeneracy" in the theory of linear programming, and its theory is well understood.[18] From this point on further increases in excess-supply at I

[18] This "degeneracy" is not quite the same thing as the following phenomenon: there may for some E's be more than one optimal trade pattern, with choice between the different optima

will depress P_1 and P_3 equally but will leave P_2 and P_4 unchanged, and will not affect exports or consumption in those regions.

But if excess-supply in I increases by still more, it will finally so depress P_1 as to permit I to capture II's remaining export market in IV. II is now isolated as in Figure 4c, which shows a lopsided forest. In this region P_1, P_3, and P_4 all fall equally, and imports of III and IV gradually increase. P_2 and II's consumption remain unchanged, with her export remaining zero.

In the ultimate stage shown by Figure 5d, excess-supply increases so much in I that we again have a tree. I has become an exporter to everybody; all prices are from now on locked together and go down equally.

To clinch his understanding of the principles here shown, the reader should draw the diagrams that would precede Figure 5 as the excess-supply function in I is *reduced*. He can show how all P's rise together until I ceases to export and becomes isolated; how all other P's then remain unchanged, until finally I becomes an importer from II and a tree is formed; still further increases in P_1 will finally cause III and IV to become exporters to I rather than importers, finally putting I in the ultimate importer stage. The reader may also experiment with *ceteris paribus* shifts in excess-supply in some other region. If he rigorously specifies a local excess-supply curve at each point, he can rigorously work out the equilibrium solution at each stage; however, it will be a sufficient test of his understanding if he can correctly infer the qualitative direction in which P's and E's must shift.

IX.—*Comparative Statics*

We can now cash in on our success in converting the spatial equilibrium system into a maximum problem: for such systems it is easy to make rigorous predictions as to the qualitative direction in which the variables of the system will change when some change is made in the data of the problem.

Let me begin with a simple case. Suppose the transport cost rises between i and j. What effect will this alone tend to have on the trade pattern? I think anyone will guess that, however other variables may change, exports from i to j must certainly decrease or at worst remain

being a matter of indifference. Thus, Koopmans, *op. cit.* p. 253, has called attention to the following kind of a situation: shipments from London to San Francisco meet shipments from New York to Honolulu in Panama; provided we are dealing with homogeneous shipments of wheat or empty vessels, it is obviously *indifferent* whether we reroute some of the London-Frisco shipments to Honolulu—provided an equal and opposite rerouting of New York cargo is simultaneously made. The cost data of Table A provide such an example when III is exporting say 15, IV exporting 30, I importing 25 and II importing 20. It then becomes a matter of indifference as to how III's exports are divided between I and II, provided that IV's shipments compensate. Professor Robert Solow points out to me that creating a fictitious set of two intermediate ports—V' which receives 45 and V'' which exports 45—will get rid of the indeterminacy.

the same; such exports can certainly not increase.[19] This happens to be a correct conclusion. But how can we be sure that our intuition is correct in a system that may involve hundreds of unknowns? The analytic theory of maximum systems worked out in my *Foundations of Economic Analysis* provides us with just such assurance.

And as a matter of fact, if we try to discover why our intuition suggested the answer it did, we will discover, I believe, that we have consciously or unconsciously been already identifying the system with a maximum and we have been heuristically inputing teleology and wisdom into the system. Again, this happens to be a rigorously correct procedure once the system has been rigorously identified as a maximum one. Indeed, in the physical sciences, the somewhat mystical principle of Le Chatelier, which says that a system tends to react to a stress so as to minimize and counter its effects, is just such a heuristic teleological principle and derives its validity from the maximum conditions underlying thermodynamic equilibrium.

* * *

A more interesting application of the above kind of analysis is provided by the problem of a shift in the excess-supply function at one place alone, say at region 1. Our intuition tells us that an increase in excess-supply at 1 should cause total exports from that region, E_1, to grow.

There is no loss in generality in assuming that the shift in the excess-supply function at 1 is a parallel vertical shift, so that we simply subtract a constant a_1 from the $s_1(E_1)$ expression in order to give the excess-supply curve the desired rightward and downward shift, with more being supplied at every price. Previously our net social pay-off could be written as a function of E_1 and other variables. This is still true, but now in the expression for NSP there will be an additional term $+a_1E_1$. The interested reader may be referred to Chapter 3 of *Foundations of Economic Analysis* for a full statement of the reasoning. It is enough to state here that the algebraic sign of the change in E_1 with respect to a change in a_1 must then be positive, or at worst zero. This rigorously confirms the conjecture that an increase in excess-supply at any point tends to increase the algebraic exports of a region and decrease its algebraic imports.

(1) How will such an increase in excess-supply affect price at 1? (2) How will it affect prices everywhere else? (3) How will it affect algebraic exports, E_i, everywhere else? Our intuition does not respond so easily to these further questions. But anyone who has worked through

[19] In the Enke formulation of the problem, the total E_i's are not held constant: they are changing as they must to restore the equilibrium. Of course, a similar theorem can be stated in the Koopmans case: an increase in T_{ij}, with all E_i's constant, can never increase E_{ij}.

the numerical example of the previous section will be able to make some fairly shrewd conjectures.

First, we would guess that an increase in excess-supply at 1 will depress P_1. Surprisingly, this turns out to be quite difficult to prove. For the result does not depend upon what happens at 1 alone, and no amount of graphical shifting of the curves in the 1 market can be relied upon for valid inferences. If region 1 were self-contained, an increase in its positively sloping excess-supply curve would certainly have to depress its price in order to get the market cleared. But in the previous paragraph, we have already seen that the net exports out of 1 have gone up, which by itself tends to relieve the redundancy of local supply and to increase rather than depress the local price. Which effect will be larger—the direct depressing effect on P_1 or the indirect upward effect on P_1 of the increased exports?[20]

The correct answer tells us that the final effect on P_1 of enhanced excess-supply at 1 *must* be downward. But how do we know that this is the correct answer? Not, I think, by simple maximum reasoning alone. The following considerations are more rewarding: (1) How does region 1 "force" an increase in its exports on the rest of the world if its price does not actually fall? Surely P_1 must decline or we shall have a contradiction to the rigorously proved result that 1's exports do go up. (2) As to the change of all other P's, how does the rest of the world absorb extra algebraic imports unless prices there have "on the whole" fallen?

Actually, our previous numerical example and the theory of the dual problem in linear programming show us that very stringent conditions must be satisfied by the network of prices in spatial equilibrium. Consequently, a much stronger result can be asserted:

> An increase in excess-supply at 1 must have a downward effect on every single price, or at worst leave it unchanged. The downward effect on other prices cannot exceed the downward effect on its own price: for all regions that stay continuously connected by direct or indirect trade with 1, the changes in P_i must exactly equal the drop in P_1; but any regions that at any time remain disconnected from 1 (as in Figure 5b above) the change in P's will be less than the drop in P_1. And so long as we assume "normal" positive sloping excess-supply curves everywhere, we can confidently assert that an increase of excess-supply in region 1 must decrease algebraic exports everywhere else, or at worst leave some of them unchanged.

The proof of these statements can be supplied in a straight-forward fashion by anyone who has mastered the reasoning of the previous

[20] Mathematically, $P_1 = -a_1 + s_1(E_1)$ so that $dP_1/da_1 = -1 + s_1'(dE_1/da_1)$. These last two terms are of opposite sign so the final sign is in doubt.

numerical example relating to spatial price equilibrium. One of the remarkable features of the present model is the fact that economic intuition will lead to correct inferences in a system involving complex interdependence between any number of variables.[21]

X.—*Generalized Reciprocity Relations*

One last relation concerning reciprocity may be mentioned. Consider the effect on E_j of a unit change in the excess-supply curve at i. And let us compare it with the effect on E_i of a unit change in the excess-supply function at j. Qualitatively, these two effects have been shown to be of the same sign: increased excess-supply at any one point tends to decrease algebraic exports at any other point. We may, however, state a much more astonishing truth: a change in i's excess-supply function has *exactly the same quantitative effect* on E_j that a change in j's excess-supply function would have on E_i.

This reciprocity condition follows immediately from the maximum nature of the problem. Similar relations are known to hold in the field of physics due to the work of Maxwell, Rayleigh, and others. In the economic theory of consumer's behavior similar "integrability conditions" play an important rôle, as has been recognized by Slutsky, Hotelling, H. Schultz, Wold, Houthakker, and others.

I do not imagine that many people would be able to have derived such quantitative relations by intuitive reasoning. Nonetheless, these should not be regarded simply as some rather amusing and paradoxical relations turned up by mathematical reasoning. From a deeper methodological viewpoint, the way that we may test whether a given set of observations has arisen from a maximizing or economizing problem is

[21] Perhaps some theoretical economists will feel that the answer to the effect on P's of an increase in 1's excess-supply should have been immediately deducible from the J. R. Hicks stability conditions of *Value and Capital* (Oxford, 1946), Chaps. 5 and 8 and their appendixes. As far as the effect on P_1 itself is concerned, we must beware that we are not simply rewording the problem: "imperfect stability" is a definition and we must answer and not beg the question of whether the concrete specified Enke-Cournot system does or does not enjoy the property of being at least imperfectly stable. The problem is not quite so hopeless as this may sound, in view of the recognition in the 1946 edition of *Value and Capital* of the sufficiency of a maximum position to guarantee perfect and imperfect stability in a wide variety of cases. The present example is *not* directly one such case but it should be possible to make the necessary extensions to the theory so that it would be able to handle cases like the present. Among other things, the present example has the interesting feature of involving functions with corners where no derivatives can be uniquely defined and yet the important logical relations of a maximum position still prevail.

In connection with answering the question of the effect on prices other than P_1, problems of "complementarity" rather than stability are involved. The fact that all P's change in the same direction is related to the Mosak theorem concerning systems in which all goods are substitutes. J. L. Mosak, *General-Equilibrium Theory in International Trade Theory* (Bloomington, 1944) pp. 44, 49–51. Similar matrices have been studied by Leontief, Machlup and Metzler, Frobenius, Minkowski, Hawkins-Simon, Woodbury, Markoff, *et al.*, as discussed in a forthcoming paper by Robert Solow.

in terms of such reciprocity relations.[22] To me, one of the most interesting of the present problems is the fact that the fundamental reciprocity relations implied by a maximum problem turn out to transcend the case where partial derivatives exist.[23]

XI.—*Conclusion*

The problem of interconnected competitive regional markets is one of the rare cases where a reasonably simple and self-contained theoretical treatment is possible. It is a good case to demonstrate the powers of the theory of linear programming since it enables us to give rigorous proofs to plausible conjectures. It so happens that much of the literary discussion of the effects of tariffs and of exchange depreciations is, in the first instance at least, couched in terms of just this model. Thus when a journalist or economist tells you that depreciation of the pound will tend to increase U.K. exports to the United States, to decrease U.K. imports, and to raise pound prices on both categories of goods while depressing dollar prices on them, he is either repeating from memory what someone has worked out from a Cournot-type model or he is himself thinking in terms of such a model.[24] Needless to say, the partial equilibrium assumptions involved in the domestic demand schedules and the neglect of aggregative relations constitute a serious defect of such a model except as a rough first approximation to the answers in especially favorable cases. Good economic theory will recognize these limitations rather than be predisposed to neglect them.

[22] See Henry Schultz, *The Theory and Measurement of Demand*, Chicago, 1938, Chaps. I, XVIII, XVIX.

[23] A more general theory shows that even at corners, where partial derivatives are not uniquely defined, we can extend the definition of a derivative to include all the admissible slopes of "supporting lines or planes" and then generalized reciprocity relations of the following type hold. Plot E_j against a_i and also plot E_i against a_j. The former curve may at some points have a corner so that its generalized partial derivative is anything between, say, $-.33$ and -5.92. It will then turn out that the second curve must also have a corner at the point corresponding to the corner of the first, and its range of indeterminacy of slope must also be between $-.33$ and -5.92. Analogous relations hold in many variables. Thus, within the field of linear programming there exist quite a number of natural generalizations of the relationships that hold for regular differentiable functions.

[24] For elaborations and criticisms see G. Haberler, "The Market for Foreign Exchange and the Stability of the Balance of Payments," *Kyklos*, Vol. III (1949), pp. 193–218; J. Viner, *Studies*, Chicago, 1938, pp. 590–91. Mention should also be made of work by Yntema, Joan Robinson, J. J. Polak, and others.