STABILITY IN COMPETITION [1]

After the work of the late Professor F. Y. Edgeworth one may doubt that anything further can be said on the theory of competition among a small number of entrepreneurs. However, one important feature of actual business seems until recently to have escaped scrutiny. This is the fact that of all the purchasers of a commodity, some buy from one seller, some from another, in spite of moderate differences of price. If the purveyor of an article gradually increases his price while his rivals keep theirs fixed, the diminution in volume of his sales will in general take place continuously rather than in the abrupt way which has tacitly been assumed.

A profound difference in the nature of the stability of a competitive situation results from this fact. We shall examine it with the help of some simple mathematics. The form of the solution will serve also to bring out a number of aspects of a competitive situation whose importance warrants more attention than they have received. Among these features, all illustrated by the same simple case, we find (1) the existence of incomes not properly belonging to any of the categories usually discussed, but resulting from the discontinuity in the increase in the number of sellers with the demand; (2) a socially uneconomical system of prices, leading to needless shipment of goods and kindred deviations from optimum activities; (3) an undue tendency for competitors to imitate each other in quality of goods, in location, and in other essential ways.

Piero Sraffa has discussed [2] the neglected fact that a market is commonly subdivided into regions within each of which one seller is in a quasi-monopolistic position. The consequences of this phenomenon are here considered further. In passing we remark that the asymmetry between supply and demand, between buyer and seller, which Professor Sraffa emphasises is due to the condition that the seller sets the price and the buyers the quanti-

[1] Presented before the American Mathematical Society at New York, April 6, 1928, and subsequently revised.

[2] " The Laws of Returns Under Competitive Conditions," Economic Journal, Vol. XXXVI. pp. 535–550, especially pp. 544 ff. (December 1926).

ties they will buy. This condition in turn results from the large number of the buyers of a particular commodity as compared with the sellers. Where, as in new oil-fields and in agricultural villages, a few buyers set prices at which they will take all that is offered and exert themselves to induce producers to sell, the situation is reversed. If in the following pages the words " buy " and " sell " be everywhere interchanged, the argument remains equally valid, though applicable to a different class of businesses.

Extensive and difficult applications of the Calculus of Variations in economics have recently been made, sometimes to problems of competition among a small number of entrepreneurs.[1] For this and other reasons a re-examination of stability and related questions, using only elementary mathematics, seems timely.

Duopoly, the condition in which there are two competing merchants, was treated by A. Cournot in 1838.[2] His book went apparently without comment or review for forty-five years until Walras produced his *Théorie Mathématique de la Richesse Sociale*, and Bertrand published a caustic review of both works.[3] Bertrand's criticisms were modified and extended by Edgeworth in his treatment of duopoly in the *Giornale degli Economisti* for 1897,[4] in his criticism of Amoroso,[5] and elsewhere. Indeed all writers since Cournot, except Sraffa and Amoroso,[6] seem to hold that even apart from the likelihood of combination there is an essential instability in duopoly. Now it is true that such competition lacks complete stability; but we shall see that in a very general class of cases the independent actions of two competitors not in collusion lead to a type of equilibrium much less fragile than in the examples of Cournot, Edgeworth and Amoroso. The solution which we shall obtain can break down only in case of an express or tacit understanding which converts the supposed

[1] For references to the work of C. F. Roos and G. C. Evans on this subject see the paper by Dr. Roos, " A Dynamical Theory of Economics," in the *Journal of Political Economy*, Vol. XXXV. (1927), or that in the *Transactions of the American Mathematical Society*, Vol. XXX. (1928), p. 360. There is also an application of the Calculus of Variations to depreciation by Dr. Roos in the *Bulletin of the American Mathematical Society*, Vol. XXXIV. (1928), p. 218.

[2] *Recherches sur les Principes Mathématiques de la Théorie des Richesses*. Paris (Hachette). Chapter VII. English translation by N. T. Bacon, with introduction and bibliography by Irving Fisher (New York, Macmillan, 1897 and 1927).

[3] *Journal des Savants* (1883), pp. 499–508.

[4] Republished in English in Edgeworth's *Papers Relating to Political Economy* (London, Macmillan, 1925), Vol. I. pp. 116–26.

[5] ECONOMIC JOURNAL, Vol. XXXII. (1922), pp. 400–7.

[6] *Lezioni di Economia Mathematica* (Bologna, Zanichelli, 1921).

competitors into something like a monopoly, or in case of a price war aimed at eliminating one of them altogether.

Cournot's example was of two proprietors of mineral springs equally available to the market and producing, without cost, mineral water of identical quality. The demand is elastic, and the price is determined by the total amount put on the market. If the respective quantities produced are q_1 and q_2 the price p will be given by a function

$$p = f(q_1 + q_2).$$

The profits of the proprietors are respectively

$$\pi_1 = q_1 f(q_1 + q_2)$$

and

$$\pi_2 = q_2 f(q_1 + q_2).$$

The first proprietor adjusts q_1 so that, when q_2 has its current value, his own profit will be as great as possible. This value of q_1 may be obtained by differentiating π_1, putting

$$f(q_1 + q_2) + q_1 f'(q_1 + q_2) = 0.$$

In like manner the second proprietor adjusts q_2 so that

$$f(q_1 + q_2) + q_2 f'(q_1 + q_2) = 0.$$

There can be no equilibrium unless these equations are satisfied simultaneously. Together they determine a definite (and equal) pair of values of q_1 and q_2. Cournot showed graphically how, if a different pair of q's should obtain, each competitor in turn would readjust his production so as to approach as a limit the value given by the solution of the simultaneous equations. He concluded that the actual state of affairs will be given by the common solution, and proceeded to generalise to the case of n competitors.

Against this conclusion Bertrand brought an "objection péremptoire." The solution does not represent equilibrium, for either proprietor can by a slight reduction in price take away all his opponent's business and nearly double his own profits. The other will respond with a still lower price. Only by the use of the quantities as independent variables instead of the prices is the fallacy concealed.

Bertrand's objection was amplified by Edgeworth, who maintained that in the more general case of two monopolists controlling commodities having correlated demand, even though not identical, there is no determinate solution. Edgeworth gave a variety of examples, but nowhere took account of the stabilising effect of masses of consumers placed so as to have a natural preference for one seller or the other. In all his illustrations of competition one merchant can take away his rival's entire business by undercutting his price ever so slightly. Thus discontinuities appear, though a discontinuity, like a vacuum, is abhorred by nature. More typical of real situations is the case in which the quantity sold by each merchant is a continuous function of two variables, his own price and his competitor's. Quite commonly a tiny increase in price by one seller will send only a few customers to the other.

I

The feature of actual business to which, like Professor Sraffa, we draw attention, and which does not seem to have been generally taken account of in economic theory, is the existence with reference to each seller of groups of buyers who will deal with him instead of with his competitors in spite of a difference in price. If a seller increases his price too far he will gradually lose business to his rivals, but he does not lose all his trade instantly when he raises his price only a trifle. Many customers will still prefer to trade with him because they live nearer to his store than to the others, or because they have less freight to pay from his warehouse to their own, or because his mode of doing business is more to their liking, or because he sells other articles which they desire, or because he is a relative or a fellow Elk or Baptist, or on account of some difference in service or quality, or for a combination of reasons. Such circles of customers may be said to make every entrepreneur a monopolist within a limited class and region—and there is no monopoly which is not confined to a limited class and region. The difference between the Standard Oil Company in its prime and the little corner grocery is quantitative rather than qualitative. Between the perfect competition and monopoly of theory lie the actual cases.

It is the gradualness in the shifting of customers from one merchant to another as their prices vary independently which is ignored in the examples worked out by Cournot, Amoroso and Edgeworth. The assumption, implicit in their work, that all buyers deal with the cheapest seller leads to a type of instability which disappears when the quantity sold by each is considered as a continuous function of the differences in price. The use of such a continuous function does, to be sure, seem to violate the doctrine that in one market there can at one time be only one price. But this doctrine is only valid when the commodity in question is absolutely standardised in all respects and when the

" market " is a point, without length, breadth or thickness. It is, in fact, analogous to the physical principle that at one point in a body there can at one time be only one temperature. This principle does not prevent different temperatures from existing in different parts of a body at the same time. If it were supposed that any temperature difference, however slight, necessitates a sudden transfer of all the heat in the warmer portion of the body to the colder portion—a transfer which by the same principle would immediately be reversed—then we should have a thermal instability somewhat resembling the instability of the cases of duopoly which have been discussed. To take another physical analogy, the earth is often in astronomical calculations considered as a point, and with substantially accurate results. But the precession of the equinoxes becomes explicable only when account is taken of the ellipsoidal bulge of the earth. So in the theory of value a market is usually considered as a point in which only one price can obtain; but for some purposes it is better to consider a market as an extended region.

Consider the following illustration. The buyers of a commodity will be supposed uniformly distributed along a line of

FIG. 1.

Market of length $l = 35$. In this example $a = 4, b = 1, x = 14, y = 16$.

length l, which may be Main Street in a town or a transcontinental railroad. At distances a and b respectively from the two ends of this line are the places of business of A and B (Fig. 1). Each buyer transports his purchases home at a cost c per unit distance. Without effect upon the generality of our conclusions we shall suppose that the cost of production to A and B is zero, and that unit quantity of the commodity is consumed in each unit of time in each unit of length of line. The demand is thus at the extreme of inelasticity. No customer has any preference for either seller except on the ground of price plus transportation cost. In general there will be many causes leading particular classes of buyers to prefer one seller to another, but the ensemble of such consideration is here symbolised by transportation cost. Denote A's price by p_1, B's by p_2, and let q_1 and q_2 be the respective quantities sold.

Now B's price may be higher than A's, but if B is to sell anything at all he must not let his price exceed A's by more than the cost of transportation from A's place of business to his own. In fact he will keep his price p_2 somewhat below the figure p_1 —

$c(l - a - b)$ at which A's goods can be brought to him. Thus he will obtain all the business in the segment of length b at the right of Fig. 1, and in addition will sell to all the customers in a segment of length y depending on the difference of prices and lying between himself and A. Likewise A will, if he sells anything, sell to all the buyers in the strips of length a at the left and of length x to the right of A, where x diminishes as $p_1 - p_2$ increases.

The point of division between the regions served by the two entrepreneurs is determined by the condition that at this place it is a matter of indifference whether one buys from A or from B. Equating the delivered prices we have

$$p_1 + cx = p_2 + cy.$$

Another equation between x and y is

$$a + x + y + b = l.$$

Solving we find

$$x = \tfrac{1}{2}\left(l - a - b + \frac{p_2 - p_1}{c}\right),$$

$$y = \tfrac{1}{2}\left(l - a - b + \frac{p_1 - p_2}{c}\right),$$

so that the profits are

$$\pi_1 = p_1 q_1 = p_1(a + x) = \tfrac{1}{2}(l + a - b)p_1 - \frac{p_1^2}{2c} + \frac{p_1 p_2}{2c},$$

and $$\pi_2 = p_2 q_2 = p_2(b + y) = \tfrac{1}{2}(l - a + b)p_2 - \frac{p_2^2}{2c} + \frac{p_1 p_2}{2c}.$$

If p_1 and p_2 be taken as rectangular co-ordinates, each of the last equations represents a family of hyperbolas having identical asymptotes, one hyperbola for each value of π_1 or π_2. Some of these curves are shown in Fig. 2, where (as also in Fig. 1) we have taken $l = 35$, $a = 4$, $b = 1$, $c = 1$.

Each competitor adjusts his price so that, with the existing value of the other price, his own profit will be a maximum. This gives the equations

$$\frac{\partial \pi_1}{\partial p_1} = \tfrac{1}{2}(l + a - b) - \frac{p_1}{c} + \frac{p_2}{2c} = 0,$$

$$\frac{\partial \pi_2}{\partial p_2} = \tfrac{1}{2}(l - a + b) + \frac{p_1}{2c} - \frac{p_2}{c} = 0,$$

from which we obtain

$$p_1 = c\left(l + \frac{a - b}{3}\right),$$

$$p_2 = c\left(l - \frac{a - b}{3}\right);$$

and $$q_1 = a + x = \tfrac{1}{2}\left(l + \frac{a-b}{3}\right),$$
$$q_2 = b + y = \tfrac{1}{2}\left(l - \frac{a-b}{3}\right).$$

The conditions $\partial^2\pi_1/\partial p_1{}^2 < 0$ and $\partial^2\pi_2/\partial p_2{}^2 < 0$, sufficient for a maximum of each of the functions π_1 and π_2, are obviously satisfied.

If the two prices are originally the co-ordinates of the point Q in Fig. 2, and if A is the more alert business man of the two, he

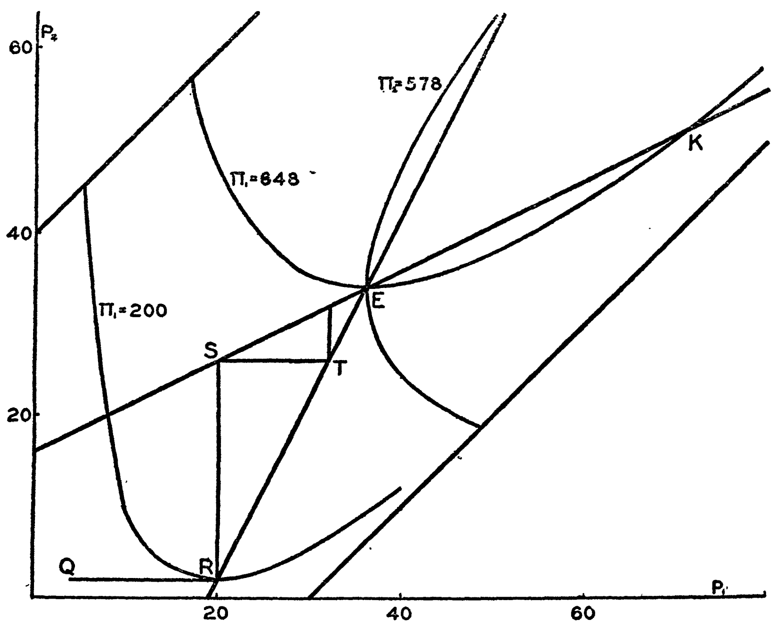

FIG. 2.

Conditions of competition for the market of Fig. 1. The co-ordinates represent the prices at A's and B's shops for the same article. The straight lines through E are the two lines of maximum profit. On one of the curves through E, A's profit is everywhere 648; on the other, B's is 578. The lower curve is the locus on which A's profit is 200.

will change his price so as to make his profit a maximum. This is represented graphically by a horizontal motion to the point R on the line $\partial\pi_1/\partial p_1 = 0$. This line has the property that every point on it represents a greater profit for A than any other point having the same ordinate. But presently B discovers that his profits can be increased by a vertical motion to the point S on his own line of maximum profit. A now moves horizontally to T. Thus there is a gradual approach to the point E at the intersection of the two lines; its co-ordinates are given by the values of p_1 and

p_2 found above. At E there is equilibrium, since neither merchant can now increase his profit by changing his price. The same result is reached if instead of Q the starting point is any on the figure.[1]

Now it is true that prices other than the co-ordinates of the equilibrium point may obtain for a considerable time. Even at this point one merchant may sacrifice his immediate income to raise his price, driving away customers, in the hope that his rival will do likewise and thus increase both profits. Indeed if A moves to the right from E in Fig. 2 he may reasonably expect that B will go up to his line of maximum profit. This will make A's profit larger than at E, provided the representing point has not gone so far to the right as K. Without this proviso, A's position will be improved (and so will B's as compared with E) if only B will sufficiently increase p_2. In fact, since the demand is inelastic, we may imagine the two alleged competitors to be amicably exploiting the consumers without limit by raising their prices. The increases need not be agreed upon in advance but may proceed by alternate steps, each seller in turn making his price higher than the other's, but not high enough to drive away all business. Thus without a formal agreement the rivals may succeed in making themselves virtually a monopoly. Something of a tacit understanding will exist that prices are to be maintained above the level immediately profitable in order to keep profits high in the long run.

But understandings between competitors are notoriously fragile. Let one of these business men, say B, find himself suddenly in need of cash. Immediately at hand he will have a resource: Let him lower his price a little, increasing his sales. His profits will be larger until A decides to stop sacrificing business

[1] The solution given above is subject to the limitation that the difference between the prices must not exceed the cost of transportation from A to B. This means that E must lie between the lines $p_1 - p_2 = \pm c(l - a - b)$ on which the hyperbolic arcs shown in Fig. 2 terminate. It is easy to find values of the constants for which this condition is not satisfied (for example, $l = 20$, $a = 11$, $b = 8$, $c = 1$). In such a case the equilibrium point will not be E and the expressions for the p's, q's and π's will be different; but there is no essential difference either in the stability of the system or in the essential validity of the subsequent remarks. A's locus of maximum profit no longer coincides with the line $\partial\pi_1/\partial p_1 = 0$, but consists of the portion of this line above its intersection with $p_1 - p_2 = c(l - a - b)$, and of the latter line below this point. Likewise B's locus of maximum profit consists of the part of the line $\partial\pi_2/\partial p_2 = 0$ to the right of its intersection with $p_2 - p_1 = c(l - a - b)$, together with the part of the last line to the left of this point. These two loci intersect at the point whose co-ordinates are, for $a > b$,

$$p_1 = c(3l - 3a - b), \quad p_2 = 2c(l - a),$$

and the type of stability is the same as before.

and lowers his price to the point of maximum profit. B will now be likely to go further in an attempt to recoup, and so the system will descend to the equilibrium position E. Here neither competitor will have any incentive to lower his price further, since the increased business obtainable would fail to compensate him.

Indeed the difficulties of maintaining a price-fixing agreement have often been remarked. Not only may the short-sighted cupidity of one party send the whole system crashing through price-cutting; the very fear of a price cut will bring on a cut. Moreover, a price agreement cannot be made once for all; where conditions of cost or of demand are changing the price needs constant revision. The result is a constant jarring, an always obvious conflict of interests. As a child's pile of blocks falls to its equilibrium position when the table on which it stands is moved, so a movement of economic conditions tends to upset quasi-monopolistic schemes for staying above the point E. For two independent merchants to come to an agreement of any sort is notoriously difficult, but when the agreement must be made all over again at frequent intervals, when each has an incentive for breaking it, and when it is frowned upon by public opinion and must be secret and perhaps illegal, then the pact is not likely to be very durable. The difficulties are, of course, more marked if the competitors are more numerous, but they decidedly are present when there are only two.

The details of the interaction of the prices and sales will, of course, vary widely in different cases. Much will depend upon such market conditions as the degree of secrecy which can be maintained, the degree of possible discrimination among customers, the force of habit and character as affecting the reliance which each competitor feels he can put in the promises of the other, the frequency with which it is feasible to change a price or a rate of production, the relative value to the entrepreneur of immediate and remote profits, and so on. But always there is an insecurity at any point other than the point E which represents equilibrium. Without some agreement, express or tacit, the value of p_1 will be less than or equal to the abscissa of K in Fig. 2; and in the absence of a willingness on the part of one of the competitors to forgo immediate profits in order to maintain prices, the prices will become the co-ordinates of E.

One important item should be noticed. The prices may be maintained in a somewhat insecure way *above* their equilibrium values but will never remain *below* them. For if either A or B has a price which is less than that satisfying the simultaneous equations it will pay him *at once* to raise it. This is evident from the figure. Strikingly in contrast with the situation pictured by Bertrand, where prices were for ever being cut below their calculated values, the stabilising effect of the intermediate customers who shift their purchases gradually with changing prices makes itself felt in the existence of a pair of minimum prices. For a prudent investor the difference is all-important.

It is, of course, possible that A, feeling stronger than his opponent and desiring to get rid of him once for all, may reduce his price so far that B will give up the struggle and retire from the business. But during the continuance of this sort of price war A's income will be curtailed more than B's. In any case its possibility does not affect the argument that there is stability, since stability is by definition merely the tendency to return after *small* displacements. A box standing on end is in stable equilibrium, even though it can be tipped over.

II

Having found a solution and acquired some confidence in it, we push the analysis further and draw a number of inferences regarding a competitive situation.

When the values of the p's and q's obtained on p. 46 are substituted in the previously found expressions for the profits we have

$$\pi_1 = \frac{c}{2}\left(l + \frac{a-b}{3}\right)^2, \quad \pi_2 = \frac{c}{2}\left(l - \frac{a-b}{3}\right)^2.$$

The profits as well as the prices depend directly upon c, the unit cost of transportation. These particular merchants would do well, instead of organising improvement clubs and booster associations to better the roads, to make transportation as difficult as possible. Still better would be their situation if they could obtain a protective tariff to hinder the transportation of their commodity between them. Of course they will not want to impede the transportation of the supplies which come to them; the object of each is merely to attain something approaching a monopoly.

Another observation on the situation is that incomes exist which do not fall strictly within any of the commonly recognised categories. The quantities π_1 and π_2 just determined may be classified as monopoly profits, but only if we are ready to extend the term " monopoly " to include such cases as have been considered, involving the most outright competition for the marginal customer but without discrimination in his favour, and with no sort of open or tacit agreement between the sellers. These profits certainly do not consist of wages, interest or rent, since we have assumed no cost of production. This condition of no cost is not essential to the existence of such profits. If a constant cost of production per unit had been introduced into the calculations above, it would simply have been added to the prices without affecting the profits. Fixed overhead charges are to be subtracted from π_1 and π_2, but may leave a substantial residuum. These gains are not compensation for risk, since they represent a minimum return. They do not belong to the generalised type of "rent," which consists of the advantage of a producer over the marginal producer, since each makes a profit, and since, moreover, we may suppose a and b equal so as to make the situation symmetrical. Indeed π_1 and π_2 represent a special though common sort of profit which results from the fact that the number of sellers is finite. If there are three or more sellers, income of this kind will still exist, but as the number increases it will decline, to be replaced by generalised "rent" for the better-placed producers and poverty for the less fortunate. The number of sellers may be thought of as increasing as a result of a gradual increase in the number of buyers. Profits of the type we have described will exist at all stages of growth excepting those at which a new seller is just entering the field.

As a further problem, suppose that A's location has been fixed but that B is free to choose his place of business. Where will he set up shop? Evidently he will choose b so as to make

$$\pi_2 = \frac{c}{2}\left(l + \frac{b-a}{3}\right)^2$$

as large as possible. This value of b cannot be found by differentiation, as the value thus determined exceeds l and, besides, yields a minimum for π_2 instead of a maximum. But for all smaller values of b, and so for all values of b within the conditions of the problem, π_2 increases with b. Consequently B will seek to make b as large as possible. This means that he will come just as close to A as other conditions permit. Naturally, if A is not exactly in the centre of the line, B will choose the side of A towards the more extensive section of the market, making b greater than a.[1]

[1] The conclusion that B will tend to gravitate *infinitesimally* close to A requires a slight modification in the particular case before us, but not in general. In the footnote on p. 48 it was seen that when A and B are sufficiently close together, the analytic expressions for the prices, and consequently the profits,

This gravitation of B towards A increases B's profit at the expense of A. Indeed, as appears from the expressions on p. 46, if b increases so that B approaches A, both q_2 and p_2 increase while q_1 and p_1 diminish. From B's standpoint the sharper competition with A due to proximity is offset by the greater body of buyers with whom he has an advantage. But the danger that the system will be overturned by the elimination of one competitor is increased. The intermediate segment of the market acts as a cushion as well as a bone of contention; when it disappears we have Cournot's case, and Bertrand's objection applies. Or, returning to the analogy of the box in stable equilibrium though standing on end, the approach of B to A corresponds to a diminution in size of the end of the box.

It has become common for real-estate subdividers in the United States to impose restrictions which tend more or less to fix the character of future businesses in particular locations. Now we find from the calculations above that the total profits of A and B amount to

$$\pi_1 + \pi_2 = c\left[l^2 + \left(\frac{a - b}{3}\right)^2\right].$$

Thus a landlord or realtor who can determine the location of future stores, expecting to absorb their profits in the sales value of the land, has a motive for making the situation as unsymmetrical as possible; for, the more the lack of symmetry, the greater is $(a - b)^2$, which appears in the expression above for $\pi_1 + \pi_2$.

Our example has also an application to the question of capitalism *v.* socialism, and contributes an argument to the socialist side. Let us consider the efficiency of our pair of merchants in serving the public by calculating the total of transportation charges paid by consumers. These charges for the strip of length a amount to $c\int_0^a t\,dt$, or $\frac{1}{2}ca^2$. Altogether the sum is

$$\tfrac{1}{2}c(a^2 + b^2 + x^2 + y^2).$$

are different. By a simple algebraic calculation which will not here be reproduced it is found that B's profits π_2 will increase as B moves from the centre towards A, only if the distance between them is more than four-fifths of the distance from A to the centre. If B approaches more closely his profit is given by $\pi_2 = bc(3l - a - 3b)$, and diminishes with increasing b. This optimum distance from A is, however, an adventitious feature of our problem resulting from a discontinuity which is necessary for simplicity. In general we should consider q_1 and q_2 as continuous functions of p_1 and p_2, instead of supposing, as here, that as $p_2 - p_1$ falls below a certain limit, a great mass of buyers shift suddenly from B to A.

Now if the places of business are both fixed, the quantities a, b and $x + y$ are all determined. The minimum total cost for transportation will be achieved if, for the given value of $x + y$, the expression $x^2 + y^2$ is a minimum. This will be the case if x and y are equal.

But x and y will not be equal unless the prices p_1 and p_2 are equal, and under competition this is not likely to be the case. If we bar the improbable case of A and B having taken up symmetrical positions on the line, the prices which will result from each seeking his own gain have been seen to be different. If the segment a in which A has a clear advantage is greater than b, then A's price will be greater than B's. Consequently some buyers will ship their purchases from B's store, though they are closer to A's, and socially it would be more economical for them to buy from A. If the stores were conducted for public service rather than for profit their prices would be identical in spite of the asymmetry of demand.

If the stores be thought of as movable, the wastefulness of private profit-seeking management becomes even more striking. There are now four variables, a, b, x and y, instead of two. Their sum is the fixed length l, and to minimise the social cost of transportation found above we must make the sum of their squares as small as possible. As before, the variables must be equal. This requires A and B to occupy symmetrical positions at the quartiles of the market. But instead of doing so they crowd together as closely as possible. Even if A, the first in the field, should settle at one of these points, we have seen that B upon his arrival will not go to the other, but will fix upon a location between A and the centre and as near A as possible.[1] Thus some customers will have to transport their goods a distance of more than $\frac{1}{2}l$, whereas with two stores run in the public interest no shipment should be for a greater distance than $\frac{1}{4}l$.

If a third seller C appears, his desire for as large a market as possible will prompt him likewise to take up a position close to A or B, but not between them. By an argument similar to that just used, it may be shown that regard only for the public interest would require A, B and C each to occupy one of the points at distances one-sixth, one-half and five-sixths of the way from one end of the line to the other. As more and more sellers of the same commodity arise, the tendency is not to become distributed in the socially optimum manner but to cluster unduly.

The importance and variety of such agglomerative tendencies

[1] With the unimportant qualification mentioned in the footnote on p. 48.

become apparent when it is remembered that distance, as we have used it for illustration, is only a figurative term for a great congeries of qualities. Instead of sellers of an identical commodity separated geographically we might have considered two competing cider merchants side by side, one selling a sweeter liquid than the other. If the consumers of cider be thought of as varying by infinitesimal degrees in the sourness they desire, we have much the same situation as before. The measure of sourness now replaces distance, while instead of transportation costs there are the degrees of disutility resulting from a consumer getting cider more or less different from what he wants. The foregoing considerations apply, particularly the conclusion that competing sellers tend to become too much alike.

The mathematical analysis thus leads to an observation of wide generality. Buyers are confronted everywhere with an excessive sameness. When a new merchant or manufacturer sets up shop he must not produce something exactly like what is already on the market or he will risk a price war of the type discussed by Bertrand in connection with Cournot's mineral springs. But there is an incentive to make the new product very much like the old, applying some slight change which will seem an improvement to as many buyers as possible without ever going far in this direction. The tremendous standardisation of our furniture, our houses, our clothing, our automobiles and our education are due in part to the economies of large-scale production, in part to fashion and imitation. But over and above these forces is the effect we have been discussing, the tendency to make only slight deviations in order to have for the new commodity as many buyers of the old as possible, to get, so to speak, *between* one's competitors and a mass of customers.

So general is this tendency that it appears in the most diverse fields of competitive activity, even quite apart from what is called economic life. In politics it is strikingly exemplified. The competition for votes between the Republican and Democratic parties does not lead to a clear drawing of issues, an adoption of two strongly contrasted positions between which the voter may choose. Instead, each party strives to make its platform as much like the other's as possible. Any radical departure would lose many votes, even though it might lead to stronger commendation of the party by some who would vote for it anyhow. Each candidate "pussyfoots," replies ambiguously to questions, refuses to take a definite stand in any controversy for fear of losing votes. Real differences, if they ever exist, fade gradually with time though the issues may be as important as ever. The Democratic party, once opposed to protective tariffs, moves gradually to a position almost, but not quite, identical with that of the Republicans. It need have no fear of fanatical free-traders, since they will still prefer it to the Republican party, and its advocacy of a continued high tariff will bring it the money and votes of some intermediate groups.

The reasoning, of course, requires modification when applied to the varied conditions of actual life. Our example might have been more complicated. Instead of a uniform distribution of customers along a line we might have assumed a varying density, but with no essential change in conclusions. Instead of a linear market we might suppose the buyers spread out on a plane. Then the customers from one region will patronise A, those from another B. The boundary between the two regions is the locus of points for which the difference of transportation costs from the two shops equals the difference of prices, *i.e.* for which the delivered price is the same whether the goods are bought from A or from B. If transportation is in straight lines (perhaps by aeroplane) at a cost proportional to the distance, the boundary will be a hyperbola, since a hyperbola is the locus of points such that the difference of distances from the foci is constant. If there are three or more sellers, their regions will be separated from each other by arcs of hyperbolas. If the transportation is not in straight lines, or if its cost is given by such a complicated function as a railroad freight schedule, the boundaries will be of another kind; but we might generalise the term hyperbola (as is done in the differential geometry of curved surfaces) to include these curves also.

The number of dimensions of our picture is increased to three or more when we represent geometrically such characters as sweetness of cider, and instead of transportation costs consider more generally the decrement of utility resulting from the actual commodity being in a different place and condition than the buyer would prefer. Each homogeneous commodity or service or entrepreneur in a competing system can be thought of as a point serving a region separated from other such regions by portions of generalised hyperboloids. The density of demand in this space is in general not uniform, and is restricted to a finite region. It is not necessary that each point representing a service or commodity shall be under the control of a different entrepreneur from every other. On the other hand, everyone who sells an article in different places or who sells different articles in the same place may be said to control the prices at several points of the symbolic space. The mutual gravitation will now take the form of a tendency of the outermost entrepreneurs to approach the cluster.

Two further modifications are important. One arises when it is possible to discriminate among customers, or to sell goods at a delivered price instead of a fixed price at store or factory plus transportation. In such cases, even without an agreement between sellers, a monopoly profit can be collected from some consumers while fierce competition is favouring others. This seems to have been the condition in the cement industry about which a controversy raged a few years ago, and was certainly involved in the railroad rebate scandals.

The other important modification has to do with the elasticity of demand. The problem of the two merchants on a linear market might be varied by supposing that each consumer buys an amount of the commodity in question which depends on the delivered price. If one tries a particular demand function the mathematical complications will now be considerable, but for the most general problems elasticity must be assumed. The difficulty as to whether prices or quantities should be used as independent variables can now be cleared up. This question has troubled many readers of Cournot. The answer is that either set of variables may be used; that the q's may be expressed in terms of the p's, and the p's in terms of the q's. This was not possible in Cournot's example of duopoly, nor heretofore in ours. The sum of our q's was constrained to have the fixed value l, so that they could not be independent, but when the demand is made elastic the constraint vanishes.

With elastic demand the observations we have made on the solution will still for the most part be qualitatively true; but the tendency for B to establish his business excessively close to A will be less marked. The increment in B's sales to his more remote customers when he moves nearer them may be more than compensation to him for abandoning some of his nearer business to A. In this case B will definitely and apart from extraneous circumstances choose a location at some distance from A. But he will not go as far from A as the public welfare would require. The tempting intermediate market will still have an influence.

In the more general problem in which the commodities purveyed differ in many dimensions the situation is the same. The elasticity of demand of particular groups does mitigate the tendency to excessive similarity of competing commodities, but not enough. It leads some factories to make cheap shoes for the poor and others to make expensive shoes for the rich, but all the shoes are too much alike. Our cities become uneconomically large and the business districts within them are too concentrated. Methodist and Presbyterian churches are too much alike; cider is too homogeneous.

HAROLD HOTELLING

Stanford University,
California.