Eaton, B.Curtis; Lipsey, Richard G.

Working Paper

The Principle of Minimum Differentiation Reconsidered: Some New Developments in the Theory of Spatial Competition

Queen's Economics Department Working Paper, No. 87

Provided in Cooperation with:
Queen’s University, Department of Economics (QED)

Suggested Citation: Eaton, B.Curtis; Lipsey, Richard G. (1972) : The Principle of Minimum Differentiation Reconsidered: Some New Developments in the Theory of Spatial Competition, Queen's Economics Department Working Paper, No. 87, Queen's University, Department of Economics, Kingston (Ontario)

This Version is available at:
https://hdl.handle.net/10419/189072

Queen's Economics Department Working Paper No. 87

THE PRINCIPLE OF MINIMUM DIFFERENTIATION RECONSIDERED: SOME NEW DEVELOPMENTS IN THE THEORY OF SPATIAL COMPETITION

B.Curtis Eaton
University of B.C.

Richard G. Lipsey
Queen's University

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

4-1972

THE PRINCIPLE OF MINIMUM DIFFERENTIATION RECONSIDERED: SOME NEW DEVELOPMENTS IN THE THEORY OF SPATIAL COMPETITION

Discussion Paper No. 87

B. Curtis Eaton - University of B.C.
and
Richard G. Lipsey - Queen's University

NOTE TO THE READER: We have endeavoured to tie our results into the existing literature wherever possible. References to such literature and discussion of points of comparison and contrast are given in the footnotes (which are at the end of the paper). Note also that is some cases our results are asserted in the text and proofs relegated to the footnotes.

ACKNOWLEDGMENT: We are indebted to many colleagues for comments and suggestions. We are also indebted to the Queen's Institute for Economic Research for generous support over a two year period.

I. INTRODUCTION

In his famous 1929 paper, "Stability in Competition" [12], Hotelling presents a model of two firms competing to sell a homogeneous product to customers spread evenly along a linear market. In equilibrium the two duopolists locate very close to each other at the centre of the market rather than being in the locations that would minimize transport costs. Hotelling originally suggested that his model explained a wide variety of social phenomena.

> 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.

Boulding [5], who appears to have been the originator of the term principle of minimum differentiation (called MD hereafter) to describe Hotelling's result, is even more extravagant in his suggestions as to the range of phenomena that are explained by the simple theoretical model.(1)

> This is a principle of the utmost generality. It explains why all the dime stores are usually clustered together, often next door to each other; why certain towns attract large numbers of firms of one kind; why an industry, such as the garment industry, will concentrate in one quarter of a city. It is a principle which can be carried over into other "differences" than spatial differences. The general rule for any new manufacturer coming into an industry is "make your product as like the existing products as you can without destroying the differences." It explains why

> all automobiles are so much alike and why no manufacturer dares make a car in which a tall hat can be worn comfortably. It even explains why Methodists, Baptists, and even Quakers are so much alike, and tend to get even more alike.

Professor Hotelling's model has been criticized and extended in the 40 odd years since its publication. It has also been applied to a number of specific cases. Professor Steiner [23], for example, successfully uses it to explain the similarity between the T.V. programs produced in Britain under conditions of duopoly (with each duopolist controlling one channel) and the variety of radio programs produced under monopoly (with the monopolist controlling three stations).

In our own research we have set ourselves the task of examining in a more systematic fashion than has been done to date the cases in which the principle of minimum differentiation does or does not apply and of discoving other principles applicable to small-group competition where neither MD nor socially-optimal differentiation seems to occur. This paper reports on our theoretical work and, although basic, we view it as preliminary to the systematic empirical applications that are necessary to remove the principle from the status of an interesting theoretical curiosity, with facts used as illustrations rather than as controls for the formulation of a satisfactory theory.

As a first step we consider how robust is the tendency toward MD in the face of changes in the specification of the model that seem empirically relevant, or otherwise interesting. Five assumptions in the Hotelling model seem critical:

(a) the nature of the consumers' demand (either one unit per period at a parametric price or completely inelastic);

(b) each firm adopts zero conjectural variation with respect to the behavior of the other firm;

(c) the number of firms is restricted to two;

(d) the firms compete in a one-dimensional space that has boundaries in each direction; and

(e) the customers are evenly spread throughout the market.

In this paper we study variations in assumptions (b), (c), (d), and (e). Since the effects of altering assumption (a) have already been studied extensively, only cross-references plus very brief discussions are provided in this paper.

We develop both one and two-dimensional models. Within each we distinguish (a) bounded, (b) unbounded but finite, and (c) unbounded, infinite spaces. Among other things, we show: in one-dimension the nature of the space is not, as many investigators have thought, critical; in two-dimensions, however, the very existence of equilibrium is seen to depend upon the nature of the space; the commonly-used rectangular customer density function yields results that do not generalize to any other density function; the existence of multiple equilibria in both one and two-dimensions is a pervasive phenomenon in any of the spaces studied and MD occurs only when the number of firms is restricted to two.

Although the analysis and discussion are in terms of location theory and are concerned with the relationship between equilibrium configuration of firms and the transport-cost minimizing configuration, many of the results generalize to other forms of differentiation. The conditions under which the results generalize are considered in the concluding section of the paper.

The paper is divided into parts: one-dimensional markets and two-dimensional markets. It is necessary to consider one-dimensional markets at

some length; first, because the analysis of such markets, although extensive in the existing literature, is both incomplete and contains some serious errors; and, second, because the one-dimensional market provides a relatively simple benchmark for comparison and contrast with the effects of relaxing certain key assumptions in the context of two-dimensional markets. Nevertheless we have organized the paper so that the analysis of the two-dimensional market is self-contained (other than for comparisons with the one-dimensional case).

We first lay out the assumptions of the basic model. All of these assumptions are maintained throughout the paper except where brief consideration is given to the effects of relaxing assumptions (ii) and (vi).

(i) Customers are distributed throughout the market according to a customer density function $c(X)$, in one dimension, and $c(X,Y)$ in two dimensions (where X and Y are distances measured from an aribitrary origin). The function is assumed to be integrable and once differentiable. Customers do not move.(2)

(ii) Each customer purchases one unit of the homogeneous product per unit of time.(3)

(iii) The customers pay transport costs which are the same increasing function of the distance from each firm to each customer.

(iv) Customers always buy from the firm that quotes the lowest delivered price (mill price plus transport cost) no matter how small is the difference between the delivered prices of different firms.(4)

(v) All firms charge the same parametric mill price.(5)

(vi) Production is at constant marginal cost which is less than the mill price. Thus, the profit maximizing firm seeks to maximize the number of customers that it serves.(6)

(vii) There are no costs of entry and the number of firms in the market is arbitrarily restricted.(7)

(viii) There are no costs of relocation. Presented with the chance of changing its location to change its market from M_1 to M_2, the firm will move if M_2 is preferred to M_1; it will remain where it is if M_1 is the same as or preferred to M_2.

(ix) No more than one firm can occupy a given location.

(x) In choosing its location, the i-th firm conjectures either (a) that all other firms will leave their own location unaltered; or (b) that some other firm will change its location in a way that causes the maximum possible reduction in the i-th firm's market. In case (b) the i-th firm adopts a minimax strategy (MM) seeking to retain the largest possible market after the other firm has made its conjectured move.

Assumption (x) is crucial. Many writers are not explicit on this important point, particularly in "free entry" models.(8) The omission of an assumption about conjectural variations is serious since equilibrium must be undefined in the absence of such an assumption. How could one establish an equilibrium without knowing the assumptions on which firms base their behaviour? In most free entry models it is merely asserted, without an explicit conjectural variation assumption, that entry will proceed until profits disappear. The correct procedure is to make some assumption about the firm's behaviour and to deduce the equilibrium level of profits that results from unrestricted entry. In fact, if firms have zero conjectural variations (ZCV) in either one or two-dimensional space, free entry does not drive profits to zero.(9)

Where conjectural variations have been used, the two assumptions of ZCV and MM are frequently employed.(10) Although not applicable

in all situations, ZCV is a reasonable assumption, either where the equilibrium is approached very rapidly so that firms do not have time to learn their opponents' reactions, or where relocation occurs with a long time lag (because, e.g., it is very costly) as with many locational problems.

Before commencing our analysis a few terms that are used throughout the paper need to be defined.

The market: The space over which potential customers are located.

The market boundaries: Limits beyond which the market does not extend.

A firm's market: The portion of the market within which the firm sells. In our model this is the set of points closer to the firm than to any other firm.

The i-th firm's market boundary: (a) an interior boundary is the locus of points that are equidistant from the i-th firm and one other firm, and not closer either to any other firm or to some portion of the market boundary; (b) an exterior boundary is that portion of the market boundary that is closer to the i-th firm than to any other firm.

An interior firm: A firm whose entire market boundary is an interior boundary.

A peripheral firm: A firm whose market boundary is an exterior boundary over some of its range.

A firm's neighbours: All those firms with whom the firm shares a common boundary.

Paired firms: Two firms are said to be paired when the distance between them is as small as is permitted. The minimum permitted distance, δ, is

arbitrary and its size is unimportant as long as it is "small" in relation to the overall size of the market. What is important is that no third firm can locate between two paired firms. Firms that are paired are also said to be located back-to-back.

Minimum Differentiation: This is said to occur when all firms in the market are separated from their neighbours by the distance δ.

Equilibrium: The i-th firm is in equilibrium when there is no location that is preferred to its present location. The whole market is in equilibrium whenever all n firms are individually in equilibrium.

II. ONE-DIMENSIONAL MARKETS

In the first part of our paper we analyze one-dimensional markets. Some further definitions are required.

The i-th firm's market segment: If the i-th firm is an interior firm, its market extends half the distance to its two neighbours. The length of an interior firm's market is thus half the length of the interval between its two neighbours wherever the firm locates within that interval. If it is a peripheral firm, its market extends all the way to the market boundary in one direction and half way to its one neighbour in the other direction.

The sides of the i-th firm's market segment: The location of the i-th firm divides its market segment into two sides. Where the two sides are unequal they are referred to as the long and short sides of the firm's market. Each side is also referred to as a half-market (whether or not they are of equal length).

The market areas of paired firms: When two firms are paired, the short side of each of their markets if $\delta/2$. It is assumed for ease of analysis that the short side of the market is zero for paired firms (actually it goes to zero as $\delta \to 0$).

MODEL 1: The assumptions that distinguish this model are ZCV and a rectangular customer density function (i.e., the customers are evenly spread along the line).

Definition: A firm is in equilibrium under ZCV when there is no move that will increase the number of customers that it serves.

We first apply model 1 to a line of finite length which gives us Hotelling's model. The length of the market is taken as unity, and its boundaries are at 0 and 1. All distances are measured from the origin. We refer to this type of market as bounded one-dimensional (B, 1-D), since if we proceed far enough in either direction we eventually encounter a boundary beyond which the market does not extend.

Figure 1 illustrates our definitions in this market. The firms numbered 1 and 2 are paired; 1 is a peripheral firm and 2 is an interior firm. The boundary between firms 1 and 2 is located at a point Y distance from the market boundary. The long side of 1's market is Y (equals 1's whole market). Firm 3 is located at 3Y. The boundary between 2 and 3 is at 2Y. Firm 2's long side is thus Y (equals its whole market) while the left hand side of 3's market is also Y (its right hand side is not determined until firm 4 is located).

The necessary and sufficient conditons for equilibrium are:

(1-i) no firm's whole market is smaller than any other firm's half market;

(1-ii) the two peripheral firms are paired.[11]

The necessity of the two conditions is established by showing that if they do not hold some firm will wish to move. Any firm can, by a suitable move, capture a market equal in length to either half market of any other firm, (e.g., if firm i pairs with firm j, firm j's relevant half market becomes i's whole market.) Thus firm i must want to move if its whole market is less than any other firm's half market (condition 1-i). Since an unpaired peripheral firm can always increase its market by moving toward its neighbour, it cannot be in equilibrium unless condition (1-ii) holds.

Sufficiency of the two conditions is established by showing that if they do hold no firm will wish to move. First consider a move within the interval defined by the firm's present neighbours. The definition of the i-th firm's market area shows that interior firms never gain by such a move, and that peripheral firms always gain by a movement away from the market boundary and towards their one neighbour. There is, however, no scope for such movements by peripheral firms once condition (1-ii) is established. Second, consider a movement to an interior interval between new neighbours j and k. Firm i would obtain half of that interval as its market which is the same as the existing (equal) half markets of j and k within that interval. But by (1-i) firm i's present whole market cannot be less than either j or k's relevant half markets. Thus i cannot gain by a move to any other interior interval if (1-i) holds. Finally, consider a movement into either of the peripheral market segments. The best firm i could do would be to pair with the existing peripheral firm gaining its relevant half market which by (1-i) cannot be larger than i's present whole market.

The application of these equilibrium conditions to various situations distinguished by the number of firms in the market is tricky. It is necessary to consider some of the cases individually.

One Firm: The location of one firm is indeterminant. It captures the whole market wherever it goes, and there is nothing in the model to make it prefer one location to another.

Two Firms: Both firms are peripheral firms and therefore, by condition (1-ii) they must be paired. Condition (1-i) dictates that they be paired at the market's centre. This is Hotelling's MD case.

Three Firms: It is impossible to satisfy the equilibrium conditions when there are three firms in the market. The only way to satisfy conditon (1-ii) is for both peripheral firms to be paired with the interior firm. But this leaves the interior firm with a market area of virtually zero - a violation of condition (1-i).[12]

Four Firms: Condition (1-ii) requires that the peripheral firms be paired and condition (1-i) is satisfied only if the pairs are located at the first and third quartiles. In this configuration, which is illustrated in Figure 2(a), all four firms have equal market areas wholly concentrated in their long sides.

Five Firms: The only possible equilibrium pattern for five firms is obtained by making firms 3 and n-2 in Figure 1 coincident - i.e., they are the same firm. The resulting configuration is shown in Figure 2(b). The peripheral pairs are located at 1/6 and 5/6 and one firm is in the centre of the market. Firm 3 has a market of 1/3 divided into two half markets of 1/6. Each of the other four firms has a market of 1/6 concentrated wholly on its long side. This configuration gives no incentive for any firm to move.

Six Firms: With six or more firms the equilibrium configuration ceases to be unique. Two limiting cases are shown in Figure 2 (c) and (d). Both of these exhibit the necessary symmetry shown in Figure 1. The first case, however, minimizes the distance between the third and fourth firms and thereby maximizes the market lengths of the four firms in the peripheral pairs. The second case maximizes the distance between 3 and 4 and thereby minimizes the (equal) market lengths for peripheral pairs. In the first case, all six firms have equal markets and in the second case the four firms in the peripheral pairs have markets of 1/8 while the inner two firms have markets of 1/4, and since they are located at the middle of these, their half markets of 1/8 just give no incentive for one of the outer firms to relocate to capture one of them.

In equilibrium, firms 3 and 4 can be separated by any distance between the extremes of δ and 1/4.[13] Firms 3 and 4 must both be separated from their neighbouring peripheral pair by the distance 2Y, but they can be separated from each other by any distance up to 2Y. (If they are separated by more than 2Y this gives them a long side of more than Y and creates an incentive for the other firms to locate in the interval between 3 and 4.) This means that firms 3 and 4 can have any market[14] between 1/6 and 1/4.

The general case for six or more firms: In general there is an infinite number of equilibria for each $n > 5$. The two extreme cases (and an intermediate case) are illustrated for seven firms in Figure 2 (e), (f) and (g). In the intermediate case the middle firm, firm 4, can be located anywhere between 4/9 and 5/9 without violating the equilibrium conditions.

To complete the analysis of six or more firms, three questions need to be answered. First, what is the range of possible market lengths for the

various firms that is compatible with equilibrium? Second, what are the equilibrium configurations that minimize, and that maximize transport costs. Third, how do total transport costs compare with the transport costs in the socially-optimal configuration (the configuration that minimizes total transport costs).

Two propositions follow immediately from the equilibrium conditions. (1) No firm can have a market more than twice as large as any other firm's market. (2) No firm can have a market smaller than Y - the market length of the firms in the peripheral pairs.

The minimum and the maximum possible sizes of the i-th firm's market depend upon the number of firms in the market, and upon whether or not the i-th firm is a member of a peripheral pair. The bounds are[15]

$$\frac{1}{2n-4} \leq L_p \leq \frac{1}{n}$$

$$\frac{1}{2n-6} \leq L_i \leq \frac{2}{n+1}$$

where L_p is the length of the market of each of the firms in the peripheral pairs and L_i is the length of the market of any other single firm.[16]

We now answer the second and third questions. The configuration that minimizes transport costs[17] has all firms spread out along the line serving equal markets of length $1/n$ divided into equal half markets of $2/n$. Because the peripheral firms must be paired, this socially-optimal configuration is not an equilibrium one. The equilibrium configuration with the lowest transport costs has two firms paired at each end of the

market and all other firms spread evenly throughout the market (for example, (d) and (f) in Figure 2). The transport costs are at the minimum that would be attainable if there were only n-2 firms in the market. What the configuration does is to 'waste' the transport-cost-reducing potential of one of the two firms in each of the peripheral pairs.

The configuration that maximizes transport costs has all firms paired (or all but one if n is odd). The pairs of firms are located at the socially-optimal location for n/2 firms. This configuration thus wastes the cost-reducing potential of every other firm and gives transport costs that are exactly (n even) double those resulting from the socially-optimal configuration.

We now test the basic conjecture that the behaviour of this model depends critically on the nature of the space by transferring model 1 to a circle whose circumference is unity. This is a one-dimensional space, but if we continue to move along it in one direction or the other we do not encounter a boundary but instead return to our starting place. We refer to this market as being unbounded, finite, one-dimensional (U, F, 1-D). Because there are no boundaries, there are only interior firms; there are neither peripheral firms nor special locational requirements near a boundary. Condition (1-i) is now the necessary and sufficient condition for equilibrium.

One Firm: As on the line, the location of one firm is indeterminant, whatever the firm's location it gets the whole market.

Two Firms: No matter where the second firm locates it gets half the circle as its market. Thus in contrast to (B, 1-D) space, any configuration is an equilibrium one.

Three Firms: Unlike (B, 1-D) space, equilibrium configurations are possible

for $n = 3$. The arc in which the third firm enters depends on the location of the first two firms. Place the first two firms arbitrarily at A and B on the circle in Figure 3. C will now wish to enter on the longer of the two arcs between A and B and wherever it locates it gains half of the arc as its own market. Some locations on this arc produce equilibrium configurations; others do not. To identify these two sets of locations for C in Figure 3, draw a diameter through A and a diameter through B intersecting the circle at C' and C''. Any location for C in the arc $C'\,C''$ produces an equilibrium configuration. Any location for C in the arc $B\,C'$ puts B into disequilibrium, and any location in the arc $A\,C''$ puts A into disequilibrium.(18)

The multiplicity of equilibria persists as n is increased. As with the line for $n > 5$, all we can do is to place some limits on the size of the firm's market. These bounds are

$$\frac{1}{2(n+1)} \leq L \leq \frac{2}{n+1}$$

where L is the length of the market arc for a single firm.(19)

Finally, we determine the minimum and maximum possible displacements from the cost-minimizing location that are compatible with equilibrium on the circle. Unlike the line, the socially-optimal configuration is compatible with equilibrium: all firms are equally spaced around the circle at a distance $1/n$ apart and no firm has an incentive to relocate. The cost-maximizing location is the same as on the line: all firms are paired (n even). The cost-reducing potential of half the firms is wasted so that transport costs are twice what they need to be.

Hotelling's model has been important in the historical development of the subject, and it is an interesting special case that has been applied successfully to certain situations. For these reasons it may be worth summarizing our conclusions.

(1) The boundedness of the space, does exert some influence on the behaviour of the model because it imposes the symmetrical configuration at the market boundaries illustrated in Figure 1. All of the special results on the line for $n \leq 5$ follow from the influence of the boundaries. (a) On the circle, one or two firms are in equilibrium whatever their locations and, for $n > 2$, there is an infinite number of configurations that fulfill the one equilibrium condition (and an infinite number that do not). On the line one firm is in equilibrium whatever its location; when $n = 2$, 4 or 5 there is an unique equilibrium; when $n = 3$ there is no equilibrium. For $n > 5$ there are more half markets than are required to produce the symmetry in Figure 1 and accordingly there is an infinite number of equilibria.[20] (b) On the line no firm can have a market smaller than the market of the four firms in the peripheral pairs. There is no analogue to this result on the circle. (c) On the circle, the socially-optimal configuration can be achieved for any n. On the bounded line the pairing of the peripheral firms makes this impossible (for any $n > 1$). All firms other than those in two peripheral pairs can, however, be spread out so as to be in the middle of equal markets. Thus as n increases, the ratio of <u>minimum transport cost consistent with equilibrium / minimum transport costs attainable by imposing the socially-optimal configuration</u> diminishes steadily as n increases, and goes to unity as n goes to infinity. (d) The nature of the space makes no difference, however, to the equilibrium configuration that maximizes transport costs. All

possible firms are paired and transport costs are twice what they could be.

(2) Some results apply equally to the circle and to the line. (a) The markets of individual firms can be different from one another provided that no firm's whole market is more than twice another firm's whole market, and that no firm's half market is more than another firm's whole market. (b) The principle of minimum differentiation is not compatible with equilibrium for any $n > 2$. The least differentiation that is possible has all firms paired.(21)

Our conclusions suggest rejection of two fundamental conjectures. First, the nature of the space is not critical to the behaviour of the model:(22) as n increases beyond 5 the behaviour on the line becomes increasingly similar to that on the circle. Second, MD is not a characeristic configuration of the linear model for $n > 2$.

MODEL 2: This is the same as model 1 - a bounded linear market with a rectangular customer density function - except that the firms adopt the minimax (MM) strategy. The market remaining to any firm, i, after one other firm, j, has made the move that is maximally damaging to i is, of course, the smaller of i's two half markets.

Definition: A firm is in equilibrium under MM when it has no move available to it that will increase the smaller of its two half markets.

Under ZCV an interior firm is indifferent between any of the locations within the interval between its two neighbours, while peripheral firms wish to be paired with their neighbours. Under MM in model 2, a firm's location is uniquely determined within any interval in which it locates. The firm maximizes its short side by locating in the middle of its own market and this

implies that interior firms locate at the mid point between their two neighbours and that peripheral firms locate one third of the distance from the market boundary to their one neighbour.

This single proposition determines the unique equilibrium configuration for any n. The firms will be spaced along the line so as to have equal market areas of 1/n and equal half markets of 1/2n. If this were not true, at least one firm would not be in the middle of its market area and hence the configuration could not be an equilibrium one.[23] Thus (with the one exception described in the footnote) a minimax strategy leads the firms to locate in the socially-optimal configuration; this configuration occurs whether the firms are guarding against the damage that could be done by a new entrant or by a move from some of the existing firms.[24]

Two firms, j and k, can always pair on either side of a third firm, i, thus reducing i's market to virtually zero, no matter where i is located. For this reason any MM model becomes completely indeterminant in a 1-D market if firm i looks ahead to the maximally damaging moves to be taken by two other firms.

We now transfer model 2 to the (U, F, 1-D) market of the circumference of a circle. Since peripheral firms are not paired on the bounded line, the removal of the boundaries has little effect on the behaviour of the model. There are a few differences, however, when n = 1 or 2. Unlike the line, when n = 1 the location of the firm is not determined on the circle. Since there are no market boundaries, one firm loses half its market to a new entrant wherever either of the firms locate. As with the line, the location of the two firms is determined, and is socially optimal, when they are guarding against entry by a third firm.[25]

In summary, MM (with firms looking only one move ahead) produces the socially-optimal configuration.[26] There is no absence of equilibrium for n = 3, nor any special cases for n > 2. Also the conjecture that the nature of the space critically affects the behaviour of a linear model with a MM strategy must be rejected.[27]

MODEL 3: This is model 1 - ZCV in a (B, 1-D) market - but with customer density functions that are not rectangular. We originally conjectured that the assumption of a rectangular density function was not critical on the arguments that local clusters could always be created by making the density function multi-modal and that, while a uni-modal distribution might pull the firms in towards the centre, it would not seriously upset any configuration established for a rectangular function. (The general acceptance of some such conjecture seems necessary to explain the considerable attention that continues to be paid to rectangular customer density functions.) This conjecture, however, is mistaken. In fact with a uni-modal density function there can be no equilibrium for more than two firms!

We first state the necessary and sufficient conditions for equilibrium in Model 3. Since, however, we only need the necessity of these conditions for our subsequent analysis, we omit the proof of their sufficiency. The conditions are:

(3-i) no firm's whole market is less than another's half market;

(3-ii) peripheral firms are paired;

(3-iii) if i is an unpaired interior firm $c(B_L) = c(B_R)$;

(3-iv) if i is a paired firm $c(B_{SS}) \geq c(B_{LS})$; where B_L and B_R denote the lefthand and righthand boundaries, and B_{SS} and B_{LS} denote

the shortside and longside boundaries.

PROOF: The arguments for the necessity of (3-i) and (3-ii) are identical to those given for model 1. To establish the necessity of (3-iii), form the expression for the i-th firm's market area (MA_i): when the i-th firm is located at X_i

$$MA_i = \int_{B_L}^{X_i} c(X)dX + \int_{X_i}^{B_R} c(X)dX \qquad \text{(Eqn. 1)}$$

$$= C(B_R) - C(B_L)$$

where $C(X) = \int c(X)dX$. We know that $B_L = (X_{i-1} + X_i)/2$ and $B_R = (X_{i+1} + X_i)/2$. Differentiating MA_i with respect to X_i:

$$\frac{\partial MA_i}{\partial X_i} = \frac{1}{2}c(B_L) - \frac{1}{2}c(B_R) \ .$$

The first order condition for a maximum is then $c(B_L) = c(B_R)$ which establishes the necessity of (3-iii). To establish the necessity of (3-iv) let a paired firm consider moving within the interval between its two neighbours. By assumption it cannot move closer to the firm with which it is paired. Let it consider moving away from that firm. By the argument immediately above the rate change of its market area is

$$\frac{\partial MA_i}{\partial X_i} = \frac{1}{2}c(B_{LS}) - \frac{1}{2}c(B_{SS}) \ .$$

If the inequality in (3-iv) holds then $\frac{\partial MA_i}{\partial X_i} < 0$ for movements away from the firm with which i is paired, and if the inequality does not hold $\frac{\partial MA_i}{\partial X_i} > 0$

for such movements. This establishes the necessity of (3-iv).

Now consider possible equilibrium configurations. Whatever the shape of the customer density function, there is always an equilibrium for n = 2; the two firms are paired at the median of the density function. Also, there is never an equilibrium for n = 3 since the pairing of both peripheral firms violates condition (3-i) for the interior firm.

There are no other possible equilibria on a customer density function that is strictly monotonic increasing from each market boundary to a single mode. It may be helpful to consider the example of four firms locating on a symmetric uni-modal customer density function. (There is a unique equilibrium with n = 4 and a rectangular density function.) The only configuration that satisfies (3-i) and (3-ii) has the firms paired at the quartiles of the density function. This is illustrated in Figure 4 where the areas R, S, T and U are equal. But this configuration leaves condition (3-iv) unsatisfied for the two interior firms. By relocating at the mode either firm can satisfy (3-iii). Assume that C makes this move before D, it will now pay D to relocate outside B; but it also pays A to come back and pair with C, after which it pays B to cross over to pair with A, and D will then come in to pair with C. This creates a grouping of the four firms at the middle of the market and an outward leapfrogging of pairs of firms towards the quartiles begins. The outward movement continues until one of the interior firms can gain more by moving back to the mode rather than relocating on the periphery of the market. The collapse to the centre of the market described above then reoccurs (although it happens before the firms have reached the quartiles).

We now consider variable density functions more generally and begin by establishing the following theorem.

Theorem: With a variable customer density function that is not rectangular over any finite range of X, a necessary condition for equilibrium is that the number of firms does not exceed the number of modes.[28]

Proof: Equilibrium condition (3-iii) implies that the market interval of any unpaired firm must contain at least one turning point as an interior point in the interval. Furthermore, the turning point must be a maximum since if it were a minimum

$$\frac{\partial^2 MA_i}{\partial X_i^2} = \frac{1}{2}\frac{\partial c(B_R)}{\partial X_i} - \frac{1}{2}\frac{\partial c(B_L)}{\partial X_i} < 0$$

which means that the firm's market area is a minimum. Equilibrium condition (3-iv) implies that any paired firm whose customer density is increasing away from the firm in the direction of the firm's long side market boundary must have a maximum point in the customer density function as an interior point in that firm's market. Since it must always be true for one of any pair of firms, not located at a mode, that $c(X)$ is increasing for a movement away from the firm's location towards its long side boundary, the long side of the market of one of the firms in each pair must include a mode. Since every unpaired firm and one member of every paired firm must have a market that includes the mode as an interior point (or as a short side boundary if the firms are paired at the mode) it is impossible to fullfill this necessary condition with $n > 2M$.

$n \leq 2M$ is only a necessary condition. For density functions with $M > 1$ there may or may not be stable equilibrium configurations even if $4 \leq n \leq 2M$![29] Everything depends on the precise shape of the function. A complete taxonomy would be tedious but two examples of the absence of equilibrium in bimodal distributions with $n = 4$ are illustrated in Figure 5. In each case, the firms are shown paired at the quartile of the density function (the areas

R, S, T, and U are equal). In Figure 5(a) conditions (3-i) and (3-iv) are not satisfied for firm C, while in Figure 5(b) conditions (3-i) and (3-ii) are satisfied for all firms but condition (3-iv) is not satisfied for either firms B or C.

A further point relating to the dynamics of a disequilibrium system such as the one considered above for the uni-modal density function is worth noting. If a firm enters the market, or considers moving, and its market is not to include the mode as an interior point it will always wish to pair with another firm. When the system is necessarily in disequilibrium ($n > 2M$) some firms must always be unable to include the mode in their markets and so will always seek to pair with another firm (who will then wish to move away from that firm). This phenomenon of pairing in disequilibrium situations is so pervasive, especially in 2-D markets, that it seems reasonable to refer to a <u>principle of pairing</u> as a basic characteristic of disequilibrium models.

When the number of firms is 2M, conditions (3-iii) and (3-iv) require that the firms all be paired. Thus transport costs are <u>necessarily</u> twice their socially-optimal level. When $n < 2M$, there is some indeterminancy in the location of some of the firms (at least for some density functions) and a full taxonomy of locations and transport costs does not seem worthwhile.

We now transfer model 3 from the bounded line to the circumference of the circle. Conditions (3-i), (3-iii) and (3-iv) — but not (3-ii) — are the necessary and sufficient conditions for equilibrium on the circle. Since the proof that $n \leq 2M$ is a necessary condition for equilibrium does not employ condition (3-ii) this result generalizes immediately to the circle.

Consider for example equilibrium with a customer density function that is strictly monotonic-increasing in both directions around the circle from a single minimum to a single maximum. The maximum number of firms compatible with the existence of equilibrium is again two. The two firms must be paired, and their location is uniquely determined by the condition that there should be equal areas under the density function over the two semi-circles defined by drawing a diameter from their point of location. If the density function is symmetrical about its maximum (and thus also its minimum) the two firms locate at the mode. If not, the firm whose customer density is increasing away from the firm towards its long side boundary must have the mode as an interior point in its market area.

Thus, the change from bounded to unbounded finite space in no way affects the behaviour of the model with ZCV and a variable customer density function. The variable customer density function does, however, produce results that differ significantly from those obtained for the rectangular density function (which can now be regarded as a special case of the variable density function in which the number of modes is infinite).

MODEL 4: Model 4 combines the MM strategy with a variable customer density function and is applied first to the bounded linear market. With a variable density function it is necessary to distinguish between the length of a firm's market segment and the number of customers in that segment, just as the firm's location divides its market segment into a long and a short side, its location divides the distribution of its customers into a large and a small market side.

In order to establish the conditions for equilibrium with a variable density function we first distinguish between a local equilibrium (no move in the neighbourhood of the firm's present location will increase its short-side market) and a global equilibrium (no move will increase the firm's own short-

side market.) We turn first to conditions for a local equilibrium. There are two distinct possibilities for local equilibrium.

Type I local equilibrium conditions: The i-th firm is located at X_i such that

$$\int_{B_L}^{X_i} c(X)dX = \int_{X_i}^{B_R} c(X)dX \quad . \qquad (4\text{-i})$$

For interior firms

$$2\, c(X_i) \leq c(B_L),\ c(B_R) \quad . \qquad (4\text{-ii-a})$$

For a left-hand peripheral firm

$$2\, c(X_i) \leq c(B_R) \quad . \qquad (4\text{-ii-b})$$

For a right-hand peripheral firm

$$2\, c(X_i) \leq c(B_L) \quad . \qquad (4\text{-ii-c})$$

If these conditions hold, a small move in either direction will decrease the firm's short side. If condition (4-i) does not hold, condition (4-ii) does, the firm cannot be in equilibrium since it can always increase either one of its market sides (and hence its small side if there is one) by moving towards the boundary of the other of its market sides. If condition (4-i) holds but (4-ii) does not hold for one boundary, the firm can increase both sides of its market by moving towards the boundary for which $c(B) > 2\, c(X_i)$. Let this boundary be B_R so that $c(B_R) > 2\, c(X_i)$ and $c(B_L) < 2\, c(X_i)$. It follows from equation (1) (and the fact that $B_R = (X_{i+1} + X_i)/2$) that

$$\frac{\partial MA_i^R}{\partial X_i} = \frac{1}{2} c(B_R) - c(X_i) > 0$$

and $\frac{\partial MA_i^L}{\partial X_i} = c(X_i) - \frac{1}{2}c(B_L) > 0$ (where superscripts on MA identify left and right hand half-markets). Thus X_i cannot represent an equilibrium location for the firm.

The possibility of condition (4-ii) being unsatisfied when (4-i) is satisfied (because the density function is "two steeply sloped") gives rise to the second type of MM equilibrium.

Type II local equilibrium conditions

An interior firm, i, must satisfy either

$$c(X_i) = \frac{1}{2}c(B_R) \text{ and } c(X_i) \leq \frac{1}{2}c(B_L) \qquad \text{(4-iii-a)}$$

or

$$c(X_i) = \frac{1}{2}c(B_L) \text{ and } c(X_i) \leq \frac{1}{2}c(B_R) . \qquad \text{(4-iii-b)}$$

A right-hand peripheral firm must satisfy

$$c(X_i) = \frac{1}{2}c(B_R) . \qquad \text{(4-iv-a)}$$

A left-hand peripheral firm must satisfy

$$c(X_i) = \frac{1}{2}c(B_L) . \qquad \text{(4-iv-b)}$$

As firm i approaches its right hand neighbour $B_R - X_i$ approaches $\delta/2$ (where δ is the arbitrary distance separating paired firms). Thus providing that $c(X_{i+1} - \delta) < 2\,c\left(\frac{X_{i+1} - \delta}{2}\right)$ equilibrium type II is established before the i-th firm becomes paired with its neighbours. The possibility of pairing is ignored in what follows.

Now consider the existence of global equilibria. Such equilibria can

be shown to exist for some density functions. An example can be obtained from Figure 6[30] by cutting the Figure off at X_5 and assuming that 3.66 is the modal point of a symmetrical distribution whose left hand side is shown by the Figure over the range $[0, X_5]$.

Global equilibrium configurations do not always exist for any n on any density function. Figure 6 provides an example. Seven firms are located in the left hand side of a particular symmetrical density function in the unique configuration that satisfies the local equilibrium conditions. (The seventh firm is at the mode and the remaining six firms are not shown.) Although the local equilibrium conditions are everywhere satisfied, firm 1 is not in global equilibrium at X_1 since its two equal half-markets (of .50) are less than the two equal half markets it could obtain by locating between firms 6 and 7.[31]

Finally, we turn to the relationship between the MM equilibria and the transport-cost-minimizing configurations.

Theorem: The MM equilibrium minimizes transport costs if and only if each firm is in type I local equilibrium.[32]

Proof: Total transport costs for n firms are:

$$TC = \sum_{i=1}^{n} \left[\int_{B_L}^{X_i} c(X)\,(X_i - X)\,dX + \int_{X_i}^{B_R} c(X)\,(X - X_i)\,dX \right] \quad \text{(Eqn. 2)}$$

There are n necessary conditions for minimizing transport costs involving n unknowns, X_i, $(i = 1, \ldots, n)$ that can be generated from (2):

$$\frac{\partial TC}{\partial X_i} = 0 , \quad (i = 1, \ldots, n) .$$

If i is an interior firm, the terms involving i in equation (2) are:

$$\int_{X_{i-1}}^{(X_{i-1}+X_i)/2} c(X)\,(X-X_{i-1})\,dX \;+\; \int_{(X_{i-1}+X_i)/2}^{X_i} c(X)\,(X_i-X)\,dX \;+\; \int_{X_i}^{(X_{i+1}+X_i)/2} c(X)\,(X-X_i)\,dX$$

$$+\; \int_{(X_{i+1}+X_i)/2}^{X_{i+1}} c(X)\,(X_{i-1}-X)\,dX \quad ,$$

since the boundary between two firms is at the midpoint of the interval between them. Letting $G(X) = \int Xc(X)dX$ and $C(X) = \int c(X)dX$ the above terms can be written as

$$X_{i-1}C(X_{i-1}) - G(X_{i-1}) - X_{i-1}C\left(\frac{X_i+X_{i-1}}{2}\right) + 2X_iC(X_i) - 2G(X_i)$$

$$- X_iC\left(\frac{X_i+X_{i-1}}{2}\right) - X_iC\left(\frac{X_i+X_{i+1}}{2}\right) + 2G\left(\frac{X_i+X_{i-1}}{2}\right) + 2G\left(\frac{X_i+X_{i+1}}{2}\right)$$

$$- X_{i-1}C\left(\frac{X_i+X_{i+1}}{2}\right) + X_i\,C\left(X_{i+1}\right) - G\,X_{i-1} \quad .$$

Differentiating with respect to X_i and simplifying, we obtain the i-th transport minimizing condition:

$$2C(X_i) - C\left(\frac{X_i+X_{i-1}}{2}\right) - C\left(\frac{X_i+X_{i+1}}{2}\right) = 0 \quad .$$

This can be written

$$C(X_i) - C\left(\frac{X_i + X_{i-1}}{2}\right) = C\left(\frac{X_i + X_{i+1}}{2}\right) - C(X_i) \quad . \quad \text{(Eqn. 3)}$$

(3) is immediately seen to be equivalent to condition (4-i).[33] Thus, if type I equilibrium prevails, it is transport-cost minimizing. Finally, if type II equilibrium prevails for any firm, that firms' half markets are not equal. Hence the configuration is not transport-cost minimizing.

We have thus shown that for a given n and a given density function, the minimax strategy may or may not lead to an equilibrium configuration, and, if it does, the configuration may or may not be the transport-cost minimizing configuration.

Model 4 can be transferred to a circle without changing any of the results that we have reached for the line. (a) Global equilibria are possible for some n on at least some density functions. (Assume for example, that Figure 6 depicts one half of a symmetrical density function on a circle with a minimum at X = 0 and a maximum at X = 3.66 and that there are nine firms in the market. (b) For some n and some density functions global equilibrium is impossible. (Assume, for example, that Figure 6 depicts one half of a circle that goes as far as 4.66 on either side of the minimum at X = 0 before reaching its single maximum point, and that there are 13 firms in the market.) (iii) Equilibria of type I are, while equilibria of type II are not, transport-cost minimizing configurations.

Again, a change in the nature of the space from bounded to finite unbounded does not exert a major influence on the behaviour of the model.

Models 1 - 4 Applied to Unbounded, Infinite, 1-D Space: If we remove the bounds from the line (or what is the same thing, let the circumference of the

circle go to infinity) we obtain an infinitely extensible 1-D space in which travel in a straight line neither encounters a boundary nor returns to the starting place. Infinitely extensible space is historically a very important case in the development of 2-D locational models and we have several new results to add for 2-D space. For completeness, we should extend the four basic 1-D models to unbounded, infinitely extensible, 1-D space.

Fortunately, infinite 1-D space poses no serious problems. With a rectangular customer density function the results for ZCV and MM are the same as on the circle. For ZCV, an infinite number of configurations is possible providing no firm's whole market is smaller than any other firm's half market. For MM, the firms must be in the middle of their individual market segments and this requires that all firms have equal markets.(34)

Variable customer density makes no sense in this space unless there is some function defined over a finite segment of the line that merely repeats itself idefinitely in either direction. The result then follows that with ZCV, the number of firms in any interval of the line, that is consistent with equilibrium cannot exceed twice the number of modes in that interval. For MM, results analogous to those for the line and circle can be established.

CONCLUSIONS FOR ONE-DIMENSIONAL SPACE: This completes our study of one-dimensional markets and some of the most important conclusions are summarized below.

(1) The wide range of generalizations of the Hotelling model suggested by Boulding and others appears suspect. The results are very sensitive to changes in the number of firms, to changes in conjectural variation, and to changes in the distribution of customers throughout the market. Surprisingly,

however, only a few of the results appear sensitive to the existence or non-existence of market boundaries.

(2) Genuine MD appears to be a very special case in the linear model, existing only for n = 2. This suggests that a critical step is to test the conjecture that MD will reassert itself when the market is extended to a two-dimensional space.(35)

(3) With a rectangular density function, ZCV produces multiple equilibria. The equilibrium set includes the socially-optimal configuration on the circle but does not on the line. With a rectangular density function, MM produces a unique equilibrium which is the socially optimal configuration.

(4) The most surprising set of conclusions relates to the effects of abandoning the rectangular customer density function. In the ZCV models equilibrium cannot exist if the number of firms exceeds twice the number of modes in the density function. Under MM equilibrium does not necessarily exist, nor where it exists, is it necessarily socially-optimal.

III. TWO-DIMENSIONAL MARKETS

Having investigated in some detail how our models behave in one-dimensional space it seems critical to see how their behaviour is affected if competition occurs in a space of two dimensions. Most locational problems are two-dimensional, and most problems of product differentiation are at least two-dimensional.

Our objectives are, of necessity, much less ambitious in 2-D space, than in 1-D space both because the literature in 1-D space is more extensive and because the location problem is much simpler in 1-D space. Our two-dimensional work is limited to the effects of transferring Model 1 — ZCV and a constant customer density function [c(X, Y) = K] — to 2-D space. We investigate the questions of existence and uniqueness of equilibrium in (a) unbounded, infinite,

and (b) bounded, finite two-dimentional space.

Some terms require redefinition in 2-D space.

The boundary between any two firms: this is the locus of points that are equi-distant from the two firms, and it is given by the perpendicular bisector of the line joining them.

The i-th firm's market area: this is given by the convex set of points contained within the set of boundaries that can be reached by travelling in a straight line from the i-th firm without crossing a boundary.

MODEL 1 IN UNBOUNDED, INFINITE TWO-DIMENSIONAL (U, I, 2-D) SPACE: This is the space considered by Lösch in his pioneering work [16] and it is the space that has commanded most attention in location theory.(36) Because of its historical importance we use it as our benchmark for comparison with bounded, two-dimensional space. The market area now consists of an infinitely - extensible plane.

Three possible configurations have been considered in the literature: The firms are located so that each firm's market area is (a) an equilateral triangle, (b) a square, and (c) a regular hexagon. We refer to these as triangular, square, and hexagonal configurations. These are the only space-filling, regular polygons in 2-D space. It is presumably because (a) they are space-filling and (b) they leave all firms with identical markets that they have received the almost exclusive attention of location theorists.(37) Among the three the hexagonal configuration minimizes transport costs, and a strong presumption has arisen that it represents the unique equilibrium configuration. This presumption has not been satisfactorily demonstrated.- in fact, it has not even been demonstrated that the hexagonal configuration is an equlibrium configuration. One reason for this failure is that the concept of equilibrium is not defined in many of the existing locational models. Location theorists

have not explicitly introduced a conjectural variation assumption - rather they have imposed some form of "densest packing" and zero profits as conditions of equilibrium. The correct procedure is not to assert equilibrium conditions, but to lay out the behavioural assumptions of the model and establish the equilibrium conditions. A second reason for the failure is that it is extremely difficult to check that any configuration fulfills condition (i) using conventional analytical means. This difficulty can be overcome with numerical simulation techniques. The core of the simulation routine is to discover the size of a firm's market if it locates at any arbitrary point (X_0, Y_0) given that there are n - 1 other firms located at fixed points in the relevant part of the space. The algorithm is described in Appendix.A.

Where there is an arbitrary number of firms, n, per unit of space the equilibrium condition in model 1 is

(II-I): no firm can increase its market area by relocating.

We now consider the existence and uniqueness of equilibria where n is arbitrarily determined. Let n firms be located in what we conjecture to be an equilibrium configuration. Allow one firm to consider a large number of alternative locations throughout the market. Calculate the firm's market area for each of these alternative locations. The configuration is, in fact, an equilibrium one if the firm can find no move that increases its market area.

The firm can move within the area defined by its present neighbours in which case its own movement leaves a gap in the regular lattice of firms. It could move beyond this area in which case it would have to fit into an already complete lattice. Intuition suggests, and calculation confirms, that the firm is always better to stay within the area defined by its present neighbours than to invade any other part of the market. Within this area, calculations were made and detailed maps drawn for the market the firm would obtain in up to 400 alternative locations. The Figures reproduced here show the market area for only a small number of representative locations. We refer to these Figures as market area maps.

Figures 7, 8 and 9 show market area maps when the firms are placed in the triangular, the square and the hexagonal configurations and one firm considers

alternative locations. The firm's neighbours are shown by circled crosses. The unbracketed numbers indicate the market area that would be obtained by the firm in its various alternative locations. (The bracketed numbers give the scales on the X and Y axes.) The broken lines indicate the firm's market boundaries when it locates at the origin, thus completing the regular lattice of firms.

Figure 9 confirms that the Löschian hexagonal configuration is an equilibrium. No firm can increase its market area by changing its location. Figure 8 shows, however, that the same is true for the square configuration. Thus, although the hexagonal configuration is an equilibrium configuration, it is not an unique one. Figure 7 shows that the triangular configuration is not an equilibrium one: if all firms are placed in this configuration, each would wish to change its location.

An even more striking result occurs if we consider configurations that give rise to identical, but non-regular polygons. If we let firms be located so that their market areas are identical rectangles, we discover that such configurations are (within limits) also equilibrium configurations. Figure 10 reproduces one such rectangular equilibrium configuration. (The ratio of the long to short side of the rectangle is 9/5.) This result is extremely surprising in the light of the existing literature; it is not so surprising, however, when viewed against the infinity of equilibria established for Model 1 in 1-D space.

Furthermore, note that the conditions that all firms should have identical markets, or even equal market areas, is arbitrary. Condition II-i does not appear to be sufficient to establish that either of these restrictions will hold in all equilibrium configurations. (This opens a range of hitherto unsuspected locational equilibria which we have not yet considered.)

Two other questions should be asked. First, will any of these equilibrium configurations be established by a dynamic adjustment process if the firms are initially placed in a disequilibrium pattern? This is a very difficult question in (I, U, 2-D) space since simulation of a dynamic process with an infinite number of firms is inconceivable. Second, we could ask if

some model in which firms enter until further entry is unprofitable would produce the hexagonal (or any other) pattern as a unique equilibrium. This seems to have been the belief of many writers. The consideration of free entry requires a second equilibrium condition.

(II-ii): No new entrant anticipates that it will earn positive profits. The assumption in our models of constant marginal costs less than a parametric mill price implies positive profits for any finite number of firms, and we must restrict n arbitrarily. Thus we cannot handle free entry, but we do know that it is possible to impose equilibrium configurations other than the hexagonal one that give no incentive for entry. For example, any configuration which satisfies conditon II-i and gives all firms equal market areas can be forced to satisfy the free entry condition (condition II-ii) by imposing sufficiently large fixed costs of production (or for any given fixed costs packing the space sufficiently densely with firms so that new entrants could not expect to cover their fixed costs).

MODEL 1 IN FINITE BOUNDED TWO-DIMENSIONAL SPACE: Model 1 is now transferred to a space that is the area contained within a unit circle, i.e., the space is a disc. Since, if one proceeds in a straight line in any direction one always reaches a boundary (rather than returning to ones starting point as on the surface of a sphere), we refer to this space as bounded, two-dimensional (B, 2-D).[38]

One Firm: As with the 1-D markets, a single firm captures the whole market wherever it locates in the disc and is thus in equilibrium wherever it locates.

Two Firms: There is a unique equilibrium with two firms. They are paired in the centre of the market. To see this assume that firm A is located anywhere other than at the market centre and draw a diameter through A. If firm B now enters the market and pairs with A locating on the diameter and just

closer to the centre than A, firm B, then captures more than half of the entire market. It now pays A to relocate on the same diameter but just inside B, thus capturing more than half the market. If both firms are free to move they continue to "leapfrog" inwards along the diameter until they are located at the centre. At this point they split the market equally between themselves and no relocation can increase either firm's market area.(39)

Thus two firms in the disc exactly reproduce the Hotelling result: B's entry creates MD even if no relocation is possible, and the equilibrium with relocation produces MD with both firms located at the centre of the market.(40)

Three or more Firms: If a third firm, C, enters, when A and B are in equilibrium at the centre of the market, C will pair with either of the two existing firms. The firm that is paired with both of the other firms now has virtually no market and it will pay it to relocate outside of one of the other two firms. This could produce a leapfrogging outwards along a diameter that is exactly analogous to what happens with three firms in the (B, 1-D) market. In the 2-D market, however, the firms are not constrained to remain on a single diameter. Thus it is not obvious how three or more firms will behave in the disc. Further analysis requires that we use the technique already applied to infinitely extensible 2-D space: we conjecture an equilibrium configuration, determine its exact location and then test the conjecture.

Three configurations seemed worth investigating as candidates for equilibrium. Configuration I: All firms are evenly spaced around a circle whose radius is less than unity. Configuration II: This is the same as Configuration I except that there is an additional firm located at the centre of the disc. Configuration III: Those configurations that give equilibrium in (U, I, 2-D) space.

CONFIGURATION I: We conjecture that the firms will be regularly spaced around a fairly small circle, concentric with the market boundary. The firms thus lie at the tips of a regular, n-sided polygon and their market areas are pieces of pie all meeting at the centre of the market. To set the firms in this configuration and check the conjecture, we need to discover the radius, r, of their circle of location. This is done as follows.

Equilibrium in any configuration requires that if any firm is free to move it will choose not to move. If Configuration I is to be an equilibrium, then if n - 1 of the firms are located at the tips of an n sided regular polygon, defined by the circle of radius r, the nth will choose to locate at the vacant tip of the polygon. r is obtained by relying on this property of equilibrium.

Let n firms be located at the points of a regular, n sided polygon whose centroid is the origin. Rotate the axis so that one firm is located on the Y axis, and let this firm consider alternative locations, but constrain it to locate somewhere on the Y axis. In Figure 11 the firm's two neighbours are located a distance r along rays through the origin which form angles of $2\pi/n$ with the Y axis. Let the nth firm choose any location, P_o, on the Y axis. Its market area, MA, is bounded above by the circle, on the left by the line QR which is the perpendicular bisector of $P_o P_2$, and on the right by RS, the perpendicular bisector of $P_o P_1$. This area is a function of only three variables, r, the radius of location, n, the number of firms (which determines ϕ and θ in Figure 11), and Z, the distance along the Y axis at which the n-th firm locates. Only Z is within the control of the n-th firm and the first order condition for maximizing MA with respect to Z is discovered by setting $\partial MA/\partial Z$ equal to zero. This produces the fairly formidable expression shown as equation (5) in Appendix B.

If our original conjecture is correct, the firm that can choose its location along the Y axis will want to be the same distance, r, away from the origin as the other firms, provided that r is at its equilibrium value. If so we will have r = Z in equilibrium. Substituting this equality into the first order condition given in the appendix produces the expression

$$r = 1/2\sqrt{(1 + \sin \phi)/2}\,. \qquad \text{(Eqn. 4)}$$

If r is set at any value other than that given by the above expression, then $\partial M/\partial Z \neq 0$ evaluated at Z = r, and <u>any</u> firm would wish to move.

To determine if the configuration is an equilibrium one with respect to a small movement of the firm in <u>any</u> direction, is an almost impossible task using analytical methods, and in any case much more is required to establish global equilibrium.[41] We therefore use the numerical methods described in the previous section. We locate n firms on a circle of radius r. We then allow one firm to consider a large number of alternative locations and calculate its market area for each of these.

The market area maps produced by this technique reveal that although Configuration I is a local equilibrium, it is not a global equilibrium configuration. The four firm case is illustrated[42] in Figure 12. Three firms are located on a circle of radius .354. The numbers in the diagram give A's market area for each indicated location. (Since the disc is of unit radius its area is π.) The diagram shows that the point (0, .354) is a weak local maximum but that it is not a global maximum. Global maxima occur at two points very close to the firm's neighbours. From n = 3 to at least 17 the same problem occurs: if the firms are located on a circle of radius r and any one is free to move, it will wish to relocate next to either one of its neighbouring firms.[43]

Thus if we impose the circular configuration, it immediately breaks up. The way in which it breaks up suggests a principle of pairing similar to that found in (1-D) markets. The sub-optimal differentiation is a dis-equilibrium phenomenon since the other paired firm will immediately wish to shift its location. Configuration I is thus rejected as a possible equilibrium configuration.

CONFIGURATION II: One firm is located in the centre of the circle, the remaining firms are regularly spaced out around a circle of radius r'. A procedure analogous to that outlined in Appendix B was used to determine r'. We checked this configuration for up to 17 firms and the results are as follows.

(1) The firm in the central location is not even at a local maximum for $n = 3$. For $3 < n < 9$ the central firm is at a local maximum but not at a global maximum: small movements lower its market area but its market is maximized by moving just outside of one of the firms on the circle. For $n > 8$ the central firm is in a global maximum: a movement to any other position reduces its market area.

(2) The $n - 1$ firms located symmetrically on the circle are always at a local maximum but never at a global maximum for any n up to 17. (It did not seem worthwhile checking for larger value of n.) Any of the $n - 1$ firms loses by a small movement in the neighbourhood of its present location but gains by relocating very close to either of its neighbours.

We thus reject Configuration II as a possible equilibrium configuration. (It is worth noting, however, that Configuration II provides a stronger local maximum than does Configuration I in the two senses that in Configuration II the firms on the circle lose more for small movements away from their symmetrical location, and gain less by moving to the global maximum than they do in Configuration I.)

CONFIGURATION III: The hexagonal configuration (which provides the most familiar equilibrium configuration in (U, I, 2-D space) is adapted to the disc in the following way. Populate an infinitely extensible 2-D space with firms spaced out in the Löschian manner. Drop a circle centred on one firm. The firms left outside of this circle cease to exist.

The first four configurations that are obtained in this manner are of 1, 7, 11 and 15 firms. One firm is in equilibrium anywhere in the disc but none of the configurations for $n > 1$ are equilibrium configurations. For $n = 7$ Configurations II and III are identical. Figure 13 illustrates the absence of equilibrium for $n = 15$. It shows the initial pattern and the configuration after one round of relocations. Clearly, the Löschian pattern has broken up completely: there are five closely grouped pairs of firms and one group of three; only two firms are without a very near neighbour.(44)

Thus for n up to 15 we have discovered that the Löschian pattern is not an equilibrium configuration in bounded (B, 2-D) space. We have also transferred other patterns, such as squares and rectangles, that give equilibrium configurations in (U, I, 2-D) space onto the disc using the techniques described in the text. The patterns always break up and the reason is always the same: firms on the periphery will prefer to pair with a neighbour rather than stay where they are. The number of real cases for which the infinitely extensible plane is the correct analogue must be rather small. The great interest in the hexagonal configuration can only be explained by the assumption, sometimes made explicitly but more often implicitly, that the results obtained from (U, I, 2-D) space transfer to (B, 2-D) space. This assumption is mistaken. The existence of boundaries to the market is critical to the behaviour of the model in 2-D space (although not in 1-D space).(45)

All three conjectured equilibrium configurations have been rejected and we advance the hypothesis that there is no equilibrium configuration for Model 1 on the disc for $n > 2$. It is now necessary to study the dynamic behaviour of the model. We do this for two reasons: (1) There may be equilibrium configurations, the nature of which we have not guessed, but to which our dynamic model might quickly converge; (2) if we discover a pattern of perpetually recurring oscillations, we will have disproved the existence of an equilibrium configuration that is obtainable independent of initial conditions. Indeed, if our starting configuration is not just chosen haphazardly, but is in some sense a likely configuration, we will have thrown strong doubt on the possibility of ever attaining an equilibrium.

The procedure for studying the dynamic behaviour is as follows. The first firm is placed in the centre of the market and the second firm is allowed to pair with it. Each additional firm is then allowed to enter the market one at a time in its market-maximizing location. Thus, the initial conditions are those that arise if all firms enter the market in their best ZCV locations before any firm is allowed to relocate. After all n firms have entered, the existing firms are allowed to relocate in the sequence in which they entered. The first firm calculates it market areas for a large number of possible locations spaced evenly over the disc and relocates where its market is largest. The second firm then goes through a similar set of calculations and relocates in its market-maximizing location, and so on.

Briefly our results are as follows. For $n = 3$ all three firms begin on a diameter through the origin which we take as the X axis. They then leapfrog outwards exactly as in the 1-D space. The outward movement continues to a point where it finally pays one firm to depart slightly from the X axis.

It then pays the next firm also to depart from the X axis. The third firm then finds it most profitable to return to the centre of the disc. The other two firms immediately follow recreating the MD grouping at the centre. The outward leapfrogging then begins again and the pattern repeats endlessly.

The four firm case is shown in Figure 14. The firms enter so as to create an MD configuration which may be referred to as a "main-street". They then leapfrog out along the diameter, but the pattern soon breaks up into apparent confusion (but the principle of pairing remains clearly observable). Soon, however, it pays someone to move near the centre of the circle and the others immediately follow. They line up on a new "main-street" and the outward leapfrogging begins again. The Figure shows one such sequence. After four such sequences, however, they line up in a "main-street" that exactly reproduces the initial conditions.

Five firms is even more complex and we have taken the model through 70 individual moves. The firms leapfrog outwards, break up into apparent confusion, and finally regroup in a main-street near the origin. This sequence continues with each main-street configuration being near the origin but in a slightly different location than the previous one. We have not carried the dynamic model beyond n = 5.(46)

We strongly suspect, but as yet cannot prove, the non-existence of any equilibrium configurations in the disc beyond n = 2. Certainly, for up to n = 17 none of the three configurations that seemed likely to produce equilibrium actually did so. Also up to n = 5 there appear to be regular, cyclic oscillations.

During the whole disequilibrium process the firms tend to be clustered into several unstable groupings and all of the firms are well within the circle the location on which would minimize the costs of transport.(47) This reinforces

the conjecture that the principle of pairing (or possibly a more general principle of "local clustering")[48] should replace the principle of minimum differentiation. It also suggests the further conjectures that the absence of equilibrium may be important in many locational contexts, and that sub-optimal locations may be a persistent result through all of the dynamic fluctuations in locational patterns.

IV. CONCLUSIONS

As the title and analytical structure of our paper indicate, we set out to consider the principle of minimum differentiation when various assumptions in the Hotelling model were relaxed. Our conclusions with respect to the principle of minimum differentiation can be succinctly stated: minimum differentiation is a property of models in which firms pursue a strategy of zero conjectural variation (ZCV) and where the number of firms is restricted to 2. It does not occur in ZCV models when the number of firms exceeds two and never occurs in the other models we have considered.

In the course of our analysis we did, however, discover a more general phenomenon, the principle of pairing (or possibly of local clustering). When a new firm enters a market, or when an existing firm relocates, there is a strong tendency for that firm to locate as close as possible to another firm. This principle is applicable both to equilibrium and to disequilibrium situations. The principle of minimum differentiation is thus a special case of the principle of pairing when the number of firms in the market is restricted to two. The principle of pairing applies in (1) one-dimensional markets with non-rectangular customer density under ZCV, (2) in the bounded, two-dimensional market under ZCV, and (3) we suspect that it also applies in unbounded two-dimensional market under ZCV with a non-rectangular customer density function.

We have several times discovered the existence of multiple equilibria. Under ZCV with a rectangular distribution of customers, an infinite number of equilibria exist in unbounded, one-dimensional space, and in bounded one-dimensional space when $n > 5$. More importantly, multiple equilibria exist in unbounded, infinite, two-dimensional space under ZCV and an even distribution of customers. The hexagonal market packing configuration which has received so much attention in the literature is only one among an infinity of possible equilibrium configurations.[49] The importance of the multiplicity of equilibria is that the socially-optimal configuration, the configuration which minimizes transport costs, will not, in general, prevail.

We also encounter many situations in which no equilibrium exists. For $n > 2$ in the disc under ZCV and an even distribution of customers there do not appear to be any equilibrium configurations. When the customer density function in the one-dimensional models is non-rectangular the possibility of perpetual disequilibrium exists. If firms adopt a ZCV strategy in a one-dimensional market, the maximum number of firms consistent with equilibrium is twice the number of modes in the density function. Even under the minimax strategy equilibrium does not exist for some n and some non-rectangular density functions. We suspect that a non-even distribution of customers may remove the possibility of equilibrium in unbounded, two-dimensional markets.

The conditions under which our results generalize to non-spatial forms of differentiation (product characteristics) are of some interest. It is difficult to imagine unbounded product-characteristic spaces and so we restrict our attention to bounded spaces. Let the bounded line represent a continuum of some non-spatial characteristic, colour for example, and let the Cartesian coordinates of any point in a rectangle represent a combination of two characteristics,

smoothness and alcoholic content of whiskey, for example. Then let the customer density function describe the distribution of customers' most-preferred points through the appropriate space. A firm's location is also described by the characteristic(s) of the product it produces. For our results to generalize we require that consumers buy from the firm which is nearest to their most-preferred point in the characteristic(s) space. In one-dimension this requirement is easily understood, and in two the requirement implies that a monotonic transformation of the scales on either or both of the axes can be found such than an individual's indifference curves are circular around his most preferred point.

Our analysis of two-dimensional markets is obviously incomplete. We should like to apply models 2, 3 and 4 to the disc but this requires a more sophisticated numerical model than the one used in this paper. We should also like to study the behaviour of all four models in an unbounded, finite 2-D space (e.g., the surface of a globe). It would be interesting to discover if model 1 behaved on the globe as it does in the infinite plane (a multiplicity of equilibrium configurations) or as it does in the disc (apparently no equilibrium configurations for $n > 2$) or in some other way (e.g., the socially-optimal configuration is the unique equilibrium configuration).

Furthermore, we have at various points in the paper suggested possible extensions of the analysis. We note here some of the extensions which seem to us to be most important and which would give rise to models beyond the 4 considered in this paper. (1) The assumed inelasticity of demand ought to be relaxed in order to investigate the conjecture that the introduction of a high enough elasticity in the individual demand functions may result in a stable but sub-optimally differentiated configuration in bounded, 2-D space as it does in 1-D space. The requirement that customers always buy from the nearest firm regardless of how small the difference in delivered price, ought also to be

relaxed to test the conjecture that the introduction of a small "zone of indifference" will tend to lessen the tendency toward pairing. (3) Our models have abstracted from price competition by requiring that all firms charge the same mill price. The effects of price competition in these models deserve attention. (4) The models assume that goods are transported in a straight line from firm to customer. For analysis of the location of retailing industries, where customers must travel to the firm on streets laid out in a block system, it seems more appropriate to measure distance from firm to customer as the sides of a right triangle. (5) The assumption that each firm has only a single plant is undesirably restrictive. (6) The nature of the location and product choice decision in higher spaces is of particular interest since, as we have already discovered, the behaviour of a model in 1-D space is not a reliable indication of how it will behave in 2-D space. We are currently developing a more sophisticated simulation model which will facilitate the analysis of most of these problems.

The wide variety of theoretical results suggests that careful and detailed specification of the behaviour of firms, of the nature of the space, and of the distribution of customers is essential. Contrary to many of the conjectures in the literature the results obtained from one model do not easily generalize to other models. The principle of pairing in disequilibrium systems is, however, suggestive of some real world behaviour. The frequency with which the American automobile industry changes styles and the apparent close similarity of certain models between firms at any point in time, together with considerable but slow changes over time, seem to be consistent with the principle of local clustering. The dynamic behaviour of the model in bounded, two-dimensional space is suggestive of the patterns of the growth and decay of retailing centres within urban areas. "Main-street" and a great deal of "local

clustering", are recurring themes in the dynamics of these models. We emphasize that these are only conjectures, and that careful theoretical and empirical work is needed to establish the generality and explanatory power of the principle of pairing and of local clustering.

FIGURE 1

The Location of Firms Near the Market Boundaries.

(Firms are indicated by numbers and the boundaries of the firm's market by broken lines.)

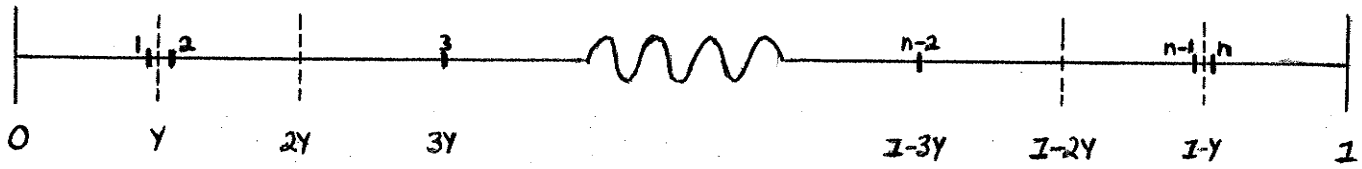


FIGURE 2

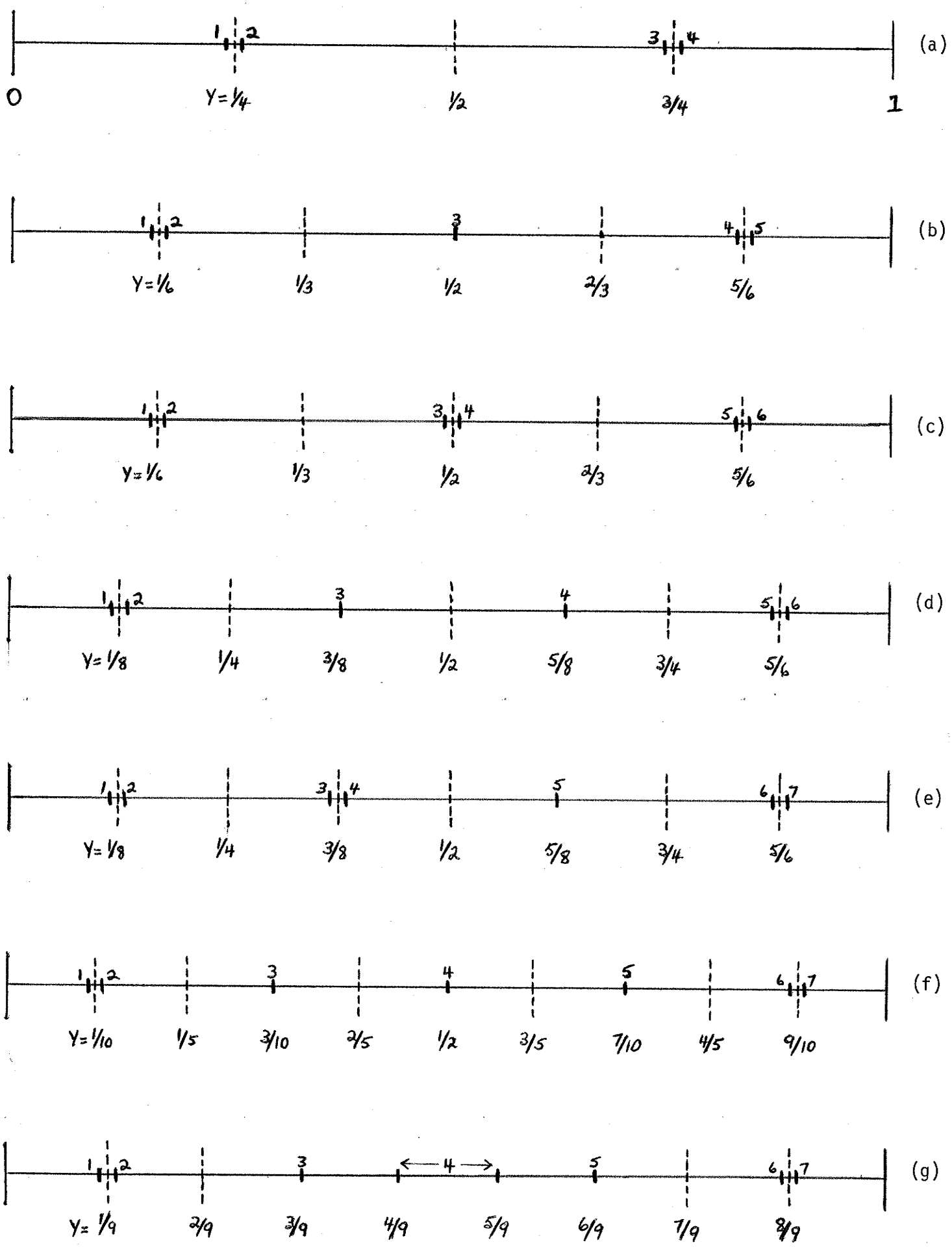

(a)
1
2
3
4
0
Y=1/4
1/2
3/4
1
(b)
1
2
3
4
5
Y=1/6
1/3
1/2
2/3
5/6
(c)
1
2
3
4
5
6
Y=1/6
1/3
1/2
2/3
5/6
(d)
1
2
3
4
5
6
Y=1/8
1/4
3/8
1/2
5/8
3/4
5/6
(e)
1
2
3
4
5
6
7
Y=1/8
1/4
3/8
1/2
5/8
3/4
5/6
(f)
1
2
3
4
5
6
7
Y=1/10
1/5
3/10
2/5
1/2
3/5
7/10
4/5
9/10
(g)
1
2
3
4
5
6
7
Y=1/9
2/9
3/9
4/9
5/9
6/9
7/9
8/9

FIGURE 3

The Range of Indeterminacy for Three Firms on a Circle

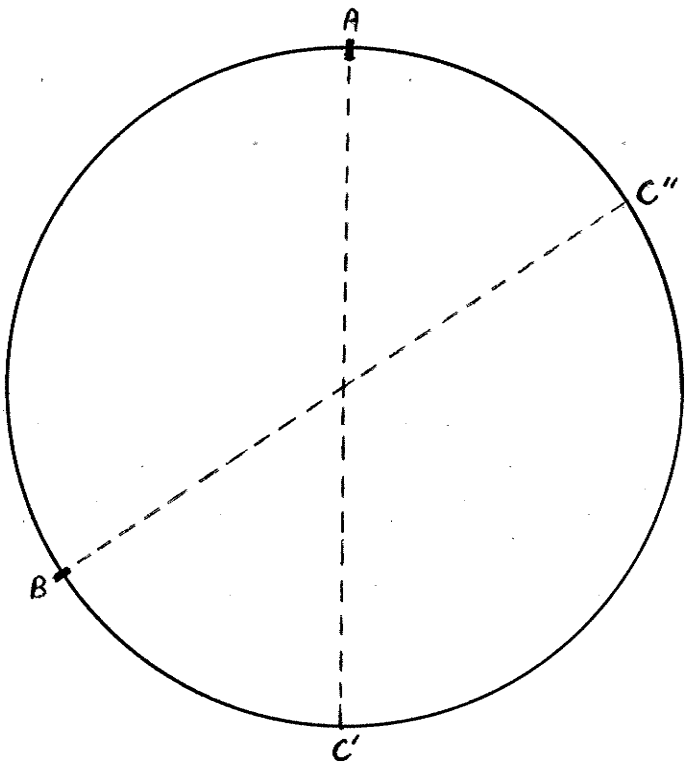


FIGURE 4

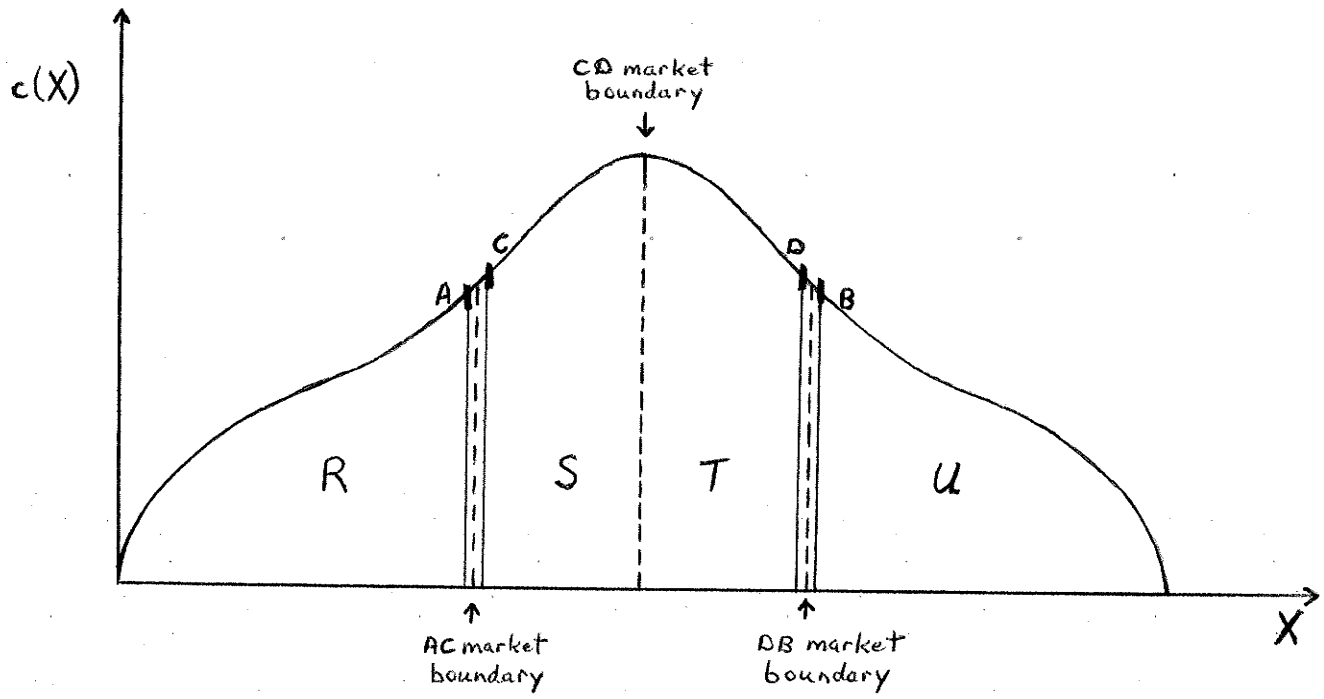
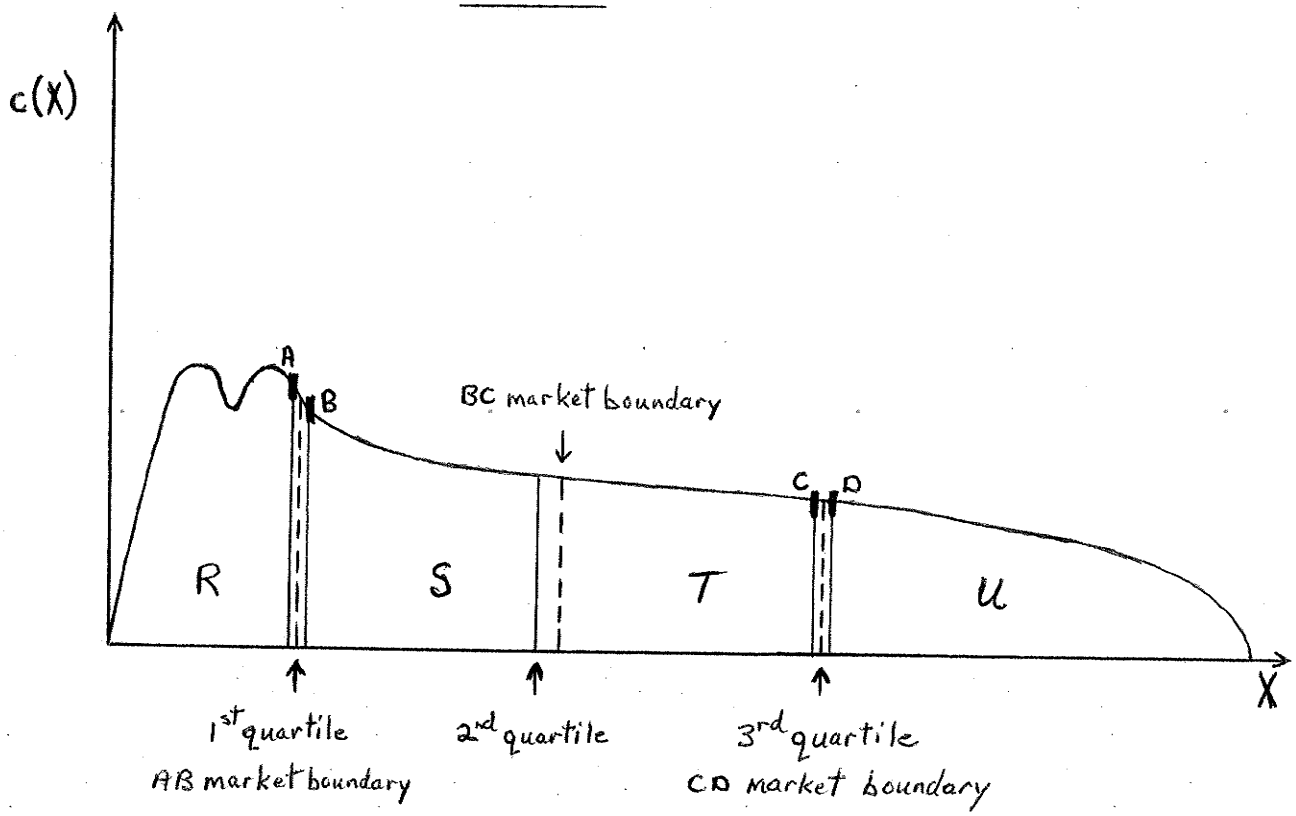

c(X)
CD market boundary
C
D
A
B
R
S
T
U
X
AC market boundary
DB market boundary

FIGURE 5


c(X)
A
B
BC market boundary
C
D
R
S
T
U
X
1st quartile
AB market boundary
2nd quartile
3rd quartile
CD market boundary

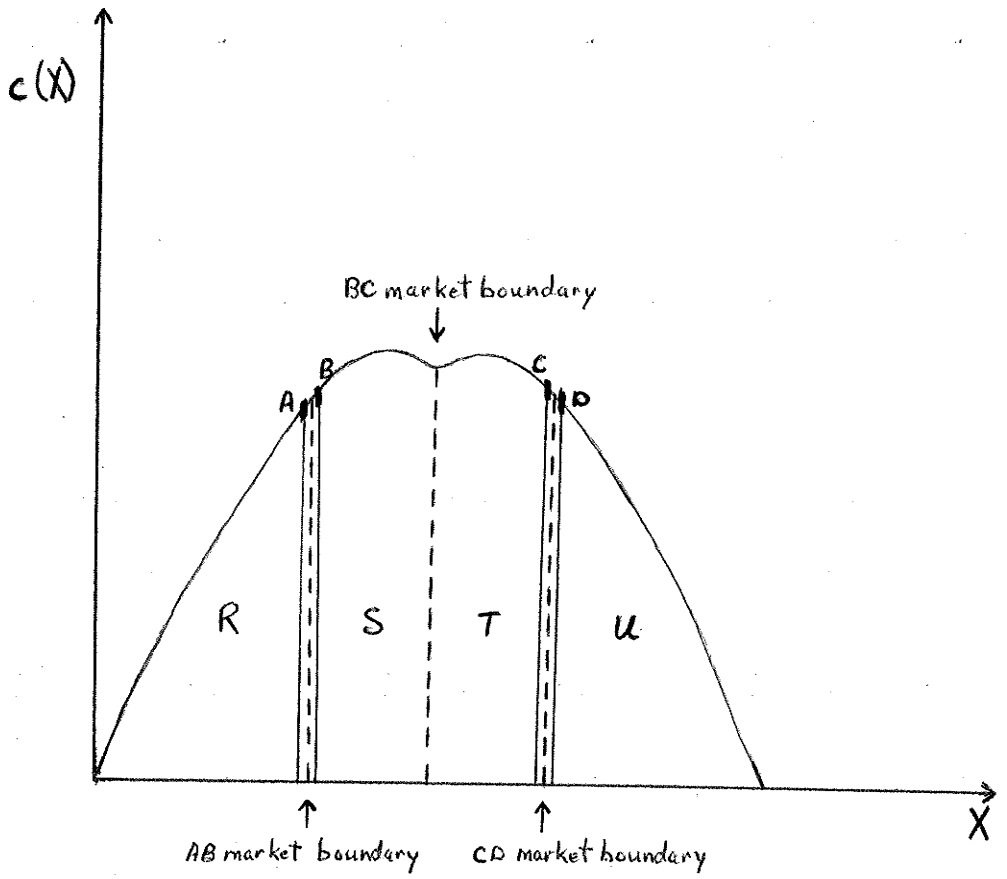

c(X)
BC market boundary
A
B
C
D
R
S
T
U
X
AB market boundary
CD market boundary

FIGURE 6

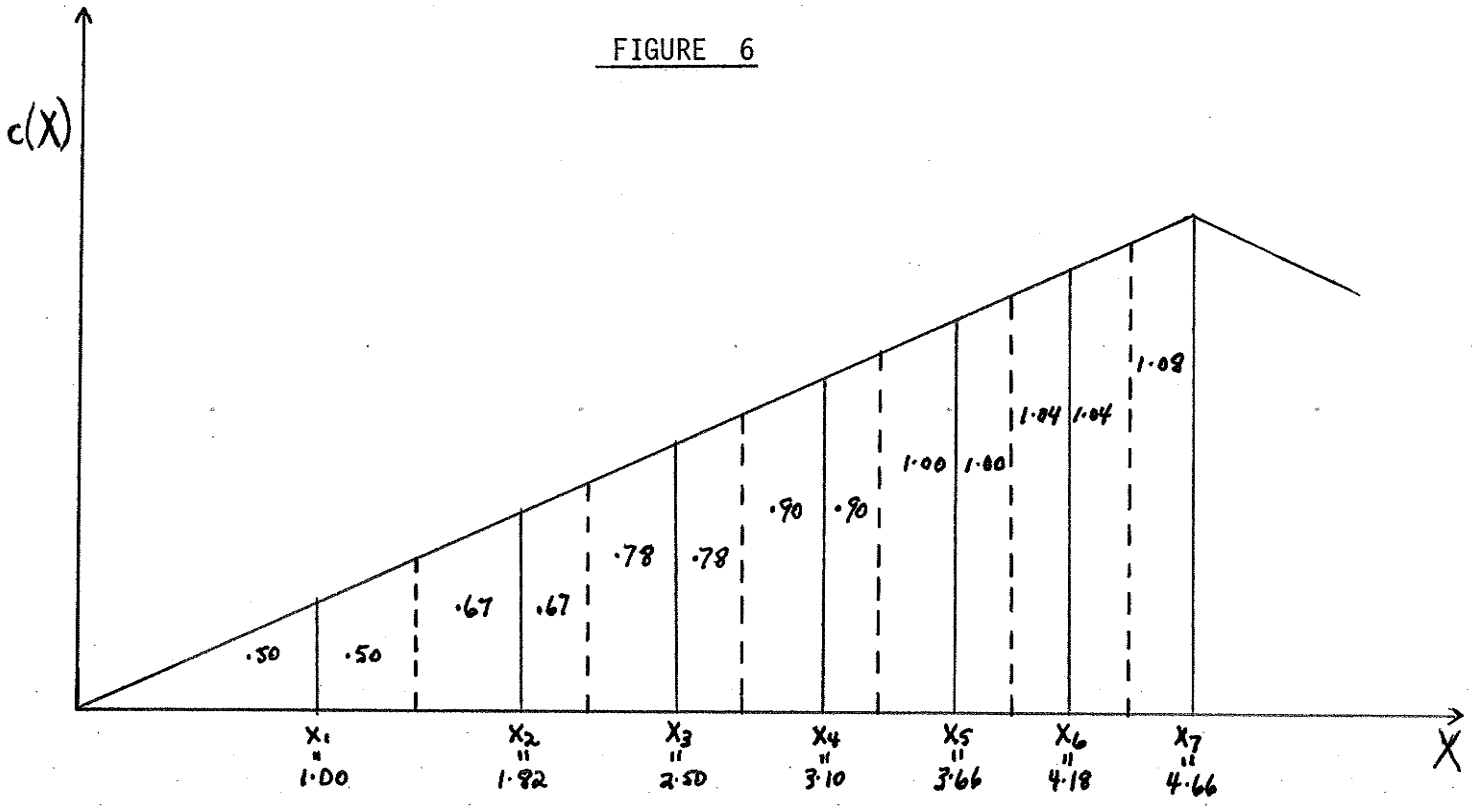

c(X)
.50
.50
.67
.67
.78
.78
.90
.90
1.00
1.00
1.04
1.04
1.08
X1 = 1.00
X2 = 1.92
X3 = 2.50
X4 = 3.10
X5 = 3.66
X6 = 4.18
X7 = 4.66
X

FIGURE 7

Market areas for alternative locations of one firm when all other firms are placed in a regular triangular configuration.

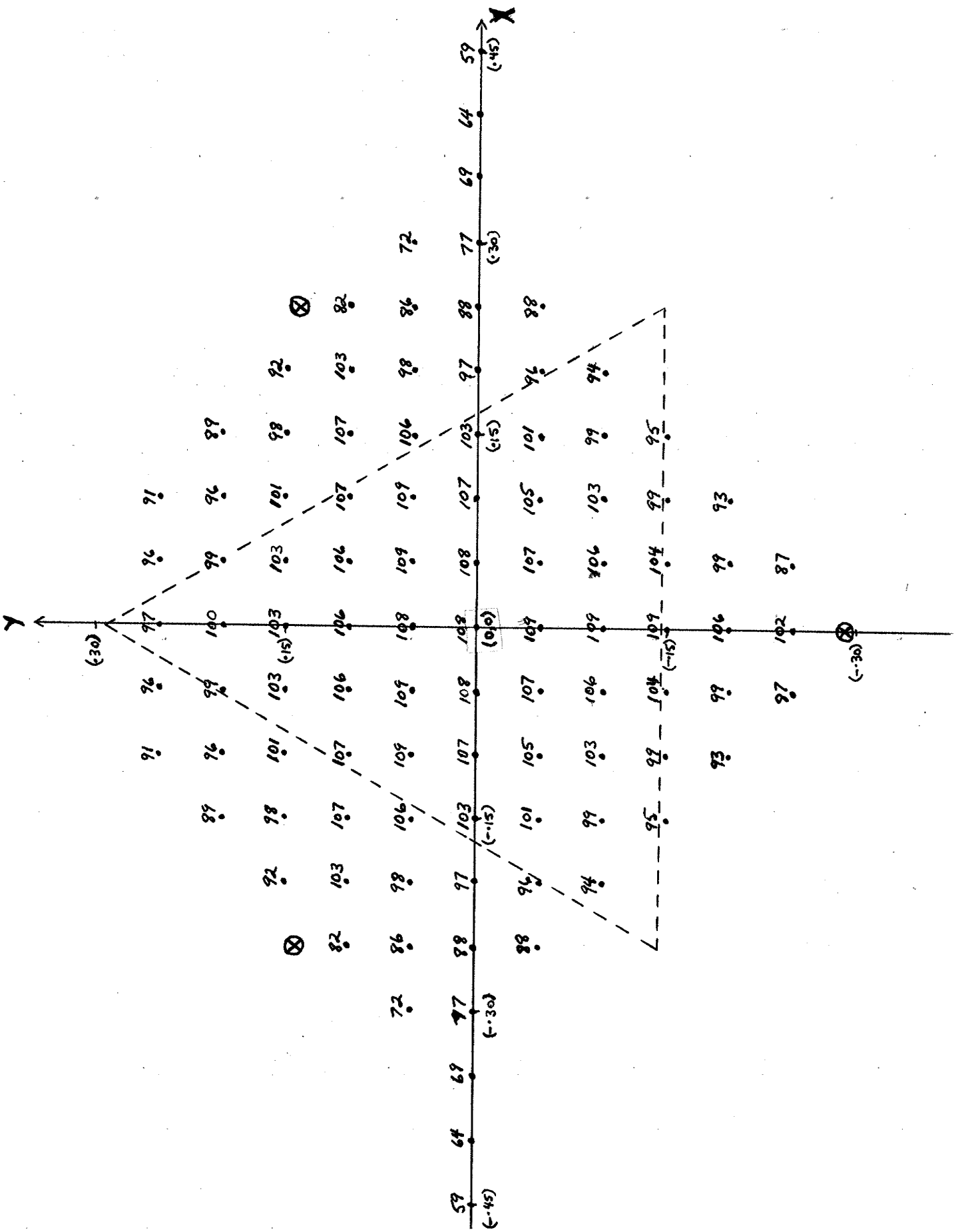


FIGURE 8

Market areas for alternative locations of one firm when all other firms are placed in a square configuration.

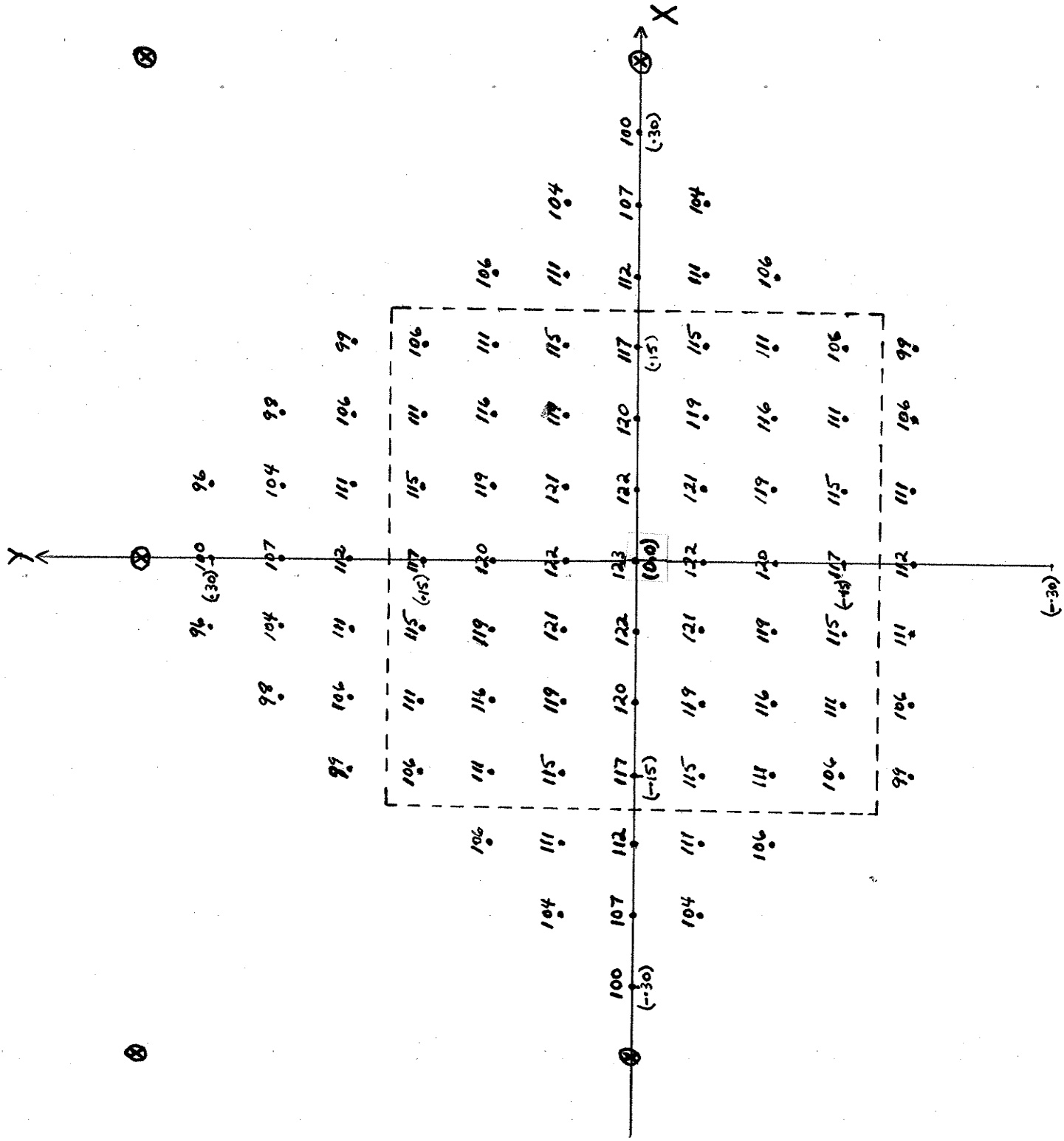


FIGURE 9

Market areas for alternative locations of one firm when all other firms are placed in a regular hexagonal configuration.

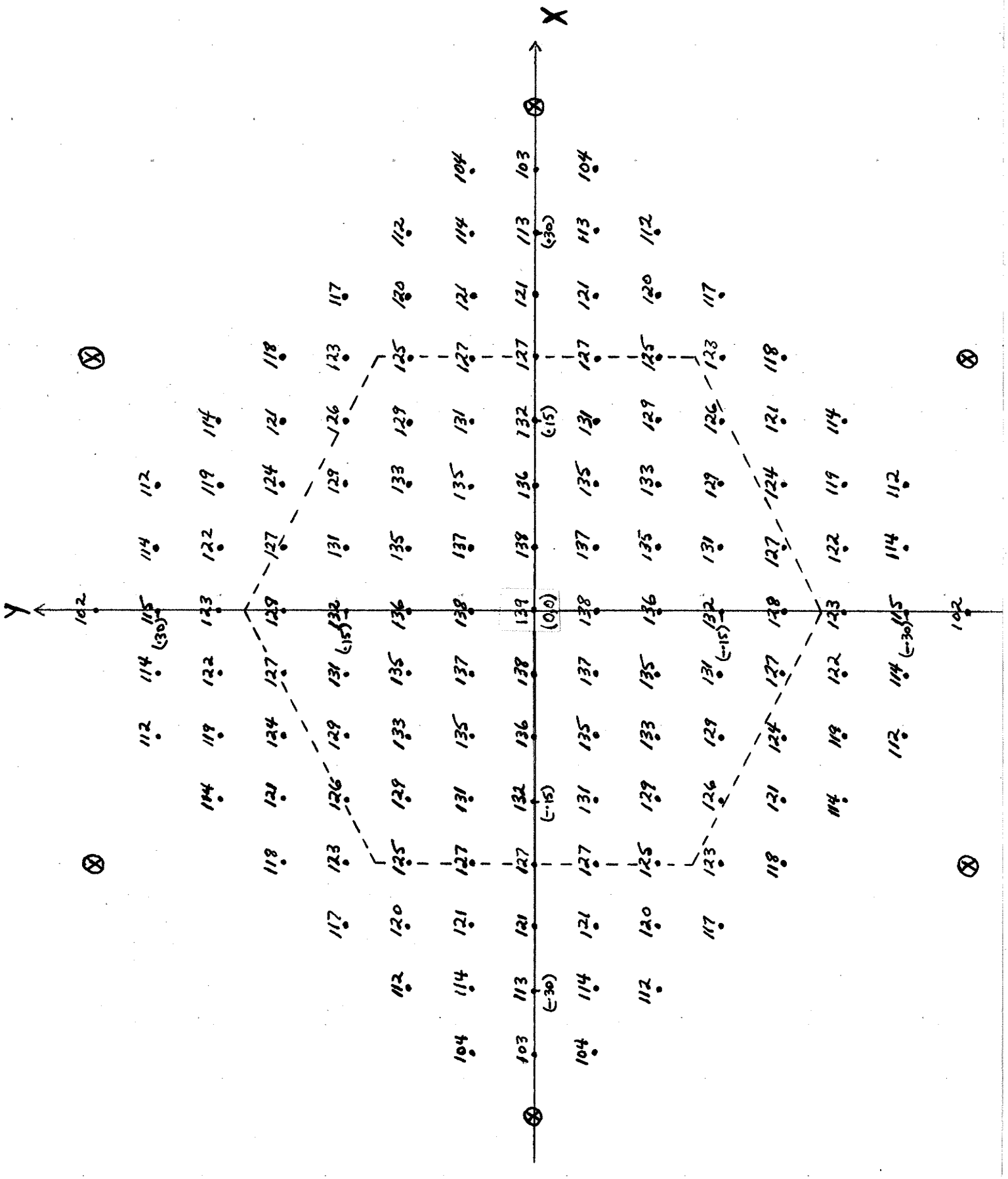


FIGURE 10

Market areas for alternative locations of one firm when all other firms are placed in a rectangular configuration.

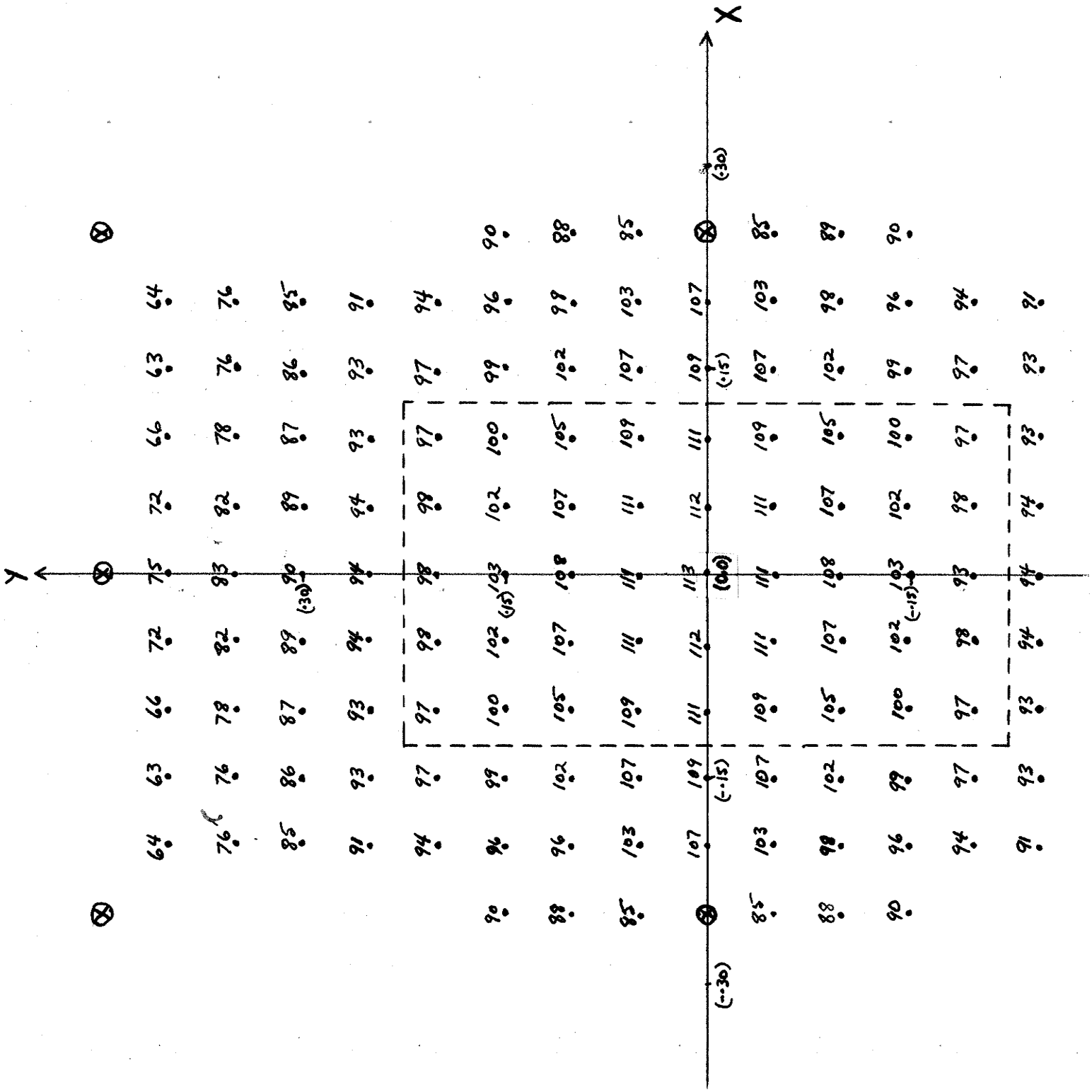


FIGURE 11

The Determination of the Market Area for a Firm
Located at P_0 on the Y Axis