Statistical Analysis of Dynamic Spatial Competition

by

Wen-Chi Tsai
National ChengChi University

Anirban DasGupta
Purdue University

Technical Report #07-04

Department of Statistics
Purdue University

March 2007

Statistical Analysis of Dynamic Spatial Competition*

Wen-Chi Tsai
National ChengChi University, Taiwan

Anirban DasGupta
Purdue University, USA

Abstract

In this article, we conduct a formal decision theoretic analysis of the problem of spatial competition, including the practically important case of dynamic competition, for both one and two dimensional markets. We give a probabilistic formulation for the dynamic competition problem, where successive vendors choose their geographic locations by using the knowledge about the locations of the already existing vendors. Under this formulation, we consider the problem of determining the optimal location of the first vendor. In addition, we consider the minimax formulation, originally considered by Hotelling only for one dimensional markets, and also the nondynamic setting.

For both one and two dimensional markets, the minimax formulation and the dynamic Bayesian formulation produce the same optimal location for the first vendor under mild conditions when there is one future competitor. Specifically, the minimax results extend Hotelling's classic result to two dimensional markets.

If the number of future competitors gets large, then a very interesting threshold phenomenon occurs. If the number of future competitors exceeds a suitable threshold value n_0, then the optimal location of the first vendor moves towards the boundary from the center of the market, and appears to eventually converge to the boundary of the market. This happens for both the dynamic and the nondynamic case.

Some illustrative examples are also given.

*Research supported by grants from the National Security Agency, USA and the Natural Sciences and Engineering Research Council, Canada.

1 Introduction

In this article, we give, for the first time, a Bayesian and decision theoretical analysis of the problem of optimal location for the first vendor in a competitive market, taking into account information about future vendors. In a classic article, Hotelling (1929) addressed this problem in the simple setting of one future vendor; Hotelling showed that if customers always visited the store closest to them and purchased a constant amount of the merchandise, then any median of the distribution of the customer's location is a minimax location for the first vendor. Due to this historical connection to Hotelling, the problem has popularly come to be known as the Hotelling Beach problem. The goal of this article is to present a series of results and examples that illustrate the very significant effect of the number of future competitors and the effect of the dimension on the optimal location of the first vendor. We present results under both the minimax formulation of Hotelling and the natural Bayes formulation in which the first vendor imposes a prior distribution on the locations of the future competitors. The emphasis is statistical. The problem is not a typical problem of statistical decision theory and yet has practical relevance.

The Hotelling Beach model for spatial competition has been the subject of many studies in urban and spatial economics. If vendors (or firms) sell a certain product at the same price but transportation costs are proportional to the distance between the customer and the firm, then customers buy from the geographically closest firm and the Hotelling beach model applies. Most of the studies in the area of spatial economics have considered the question of eventual stability in competition, when firms relocate from time to time to increase their profit. In contrast, our aim in this article is to study the locational optimization problem, in one as well as two dimensions. We do not consider optimal locations for subsequent competitors; see Steele and Zidek (1980) for some results on that problem.

Throughout the article, S denotes the region in either one or two dimensions where the customers and all vendors are located, n stands for the number of future competitors, and F stands for the CDF of the distribution of Z, the location of a customer. We always assume that F has a density f. We also assume as in Hotelling (1929) that the customer visits the business closest to him, but we do not always assume that the amount purchased is a pure constant. We let the amount purchased be a function $h(d)$ of the customer's distance d from the closest vendor. In general, the minimax results are under a general h and the Bayes results are under a constant h.

A common feature of many of the minimax results when there is only one competitor is that symmetry and unimodality of f plays a very important role. For instance, we show that one can have a general h function and still have the median m of F to be the first vendor's minimax location provided f is symmetric and unimodal around m. This is therefore a generalization of Hotelling's result. However, the minimax location can behave in erratic ways and even the uniqueness of the minimax location becomes false if the unimodality assumption is removed. If two or more competitors are expected to enter the market, then the minimax formulation will no longer work. These results and other examples on the first vendor's minimax location for the case of one-dimensional S are presented and proved in Section 2. The corresponding minimaxity results for a two dimensional market are presented in Section 5.1. We give a nice generalization of Hotelling's one dimensional result. We prove that if f is spherically symmetric and unimodal around some $\underset{\sim}{m}$, then $\underset{\sim}{m}$ is a minimax location for the first vendor even if h is not a constant function.

The Bayes formulation requires the first vendor to assign a prior distribution on the locations of the future rivals. A strictly realistic formulation would have the prior distribution for any competitor depend on the known locations of all the preceding stores. We call this the dynamic setting. In contrast, the formulation in which the locations of the future rivals are assumed to be iid according to some distribution would be called the nondynamic setting. We consider the dynamic and the nondynamic settings for both one dimensional and two dimensional markets. The one dimensional case is considered in Section 3 and Section 4, and the two dimensional case in Section 5.2 and Section 5.3. The main features of these results are as follows. When there is only one future competitor, the first vendor should typically choose the center as his optimal location. Thus the result is similar to that under the minimax formulation. But a very interesting phenomenon happens when the number of future competitors gets large. The optimal location of the first vendor starts to move towards the boundary of the market. This happens, very interestingly, under both the dynamic as well as the nondynamic setting. And very fortunately, for the nondynamic setting, we can pin the optimal location down rather precisely by means of an asymptotic expansion. We refer to this phenomenon as "affinity to the boundary".

The principal contributions of this article can be summarized as follows :

a. we give extensions of Hotelling's minimaxity result to the case of a nonconstant purchase function, as also to two dimensions;

b. we give a formulation of the practically important dynamic competition problem;

c. we show that for the case of one future competitor, the minimax and the dynamic formulation provide similar answers;

d. we show that when the number of competitors becomes large, an affinity to the boundary occurs in both the dynamic and the nondynamic setting;

e. in the nindynamic setting, we are able to pin the optimal location down precisely by an asymptotic expansion;

f. we show this affinity to the boundary occurs for both one dimensional and two dimensional markets;

g. we give many illustrative examples.

2 Minimax Solutions in One Dimension

The Hotelling Beach problem studies optimal locations for the first vendor taking into account the possibility of future rivals. The minimax strategy is usually a nice start when we encounter such a competitive situation involving two or more intelligent players. We start with notation and a description of the formulation of the minimax problem.

2.1 Notation

We begin this section by introducing the mathematical model and the notation. We denote by S the regular domain in which potential customers and businesses are located. The one-dimensional case where S is either a bounded interval or the whole real line will be considered first. Let Z be the possible position of a buyer; we assume Z to be a random variable with cumulative distribution function F and density f. Assume we are the first store to set up a vendor in the region S. The location of our store and the future competitors are represented by x and $y_1, y_2, \ldots, y_n$, respectively, all distinct. Here n is taken as fixed. We assume that a buyer has no preference for any seller and will always visit the closest store. The number of dollars he will spend in that store is assumed to be a function of the distance between him and the store, denoted by $h(d)$.

A summary of the notations is given below for the reader's convenience:

$$\begin{cases} S & : \text{ the market.} \\ Z \sim F \text{ with density } f & : \text{ the position of the customer.} \\ h(d) & : \text{ the amount spent by the customer at a distance } d \text{ from the vendor.} \\ x & : \text{ the location of the first vendor.} \\ y_1, \ldots, y_n & : \text{ the locations of the future competitors.} \end{cases}$$

The sales we expect from one customer when we locate ourselves at x and the future competitors are located at $y_1, \ldots, y_n$ is

$$D(x; y_1, \cdots, y_n) = E\left[h(|Z - x|)\mathbf{1}_{\{|Z-x|<min_{1\leq i\leq n}|Z-y_i|\}}\right].$$

The location x is in our control, but the locations $y_1, \ldots, y_n$ are not. Therefore, the minimax optimal strategy for the first vendor is to select x which maximizes $\inf_{y_1,\cdots,y_n\in S} D(x; y_1, \cdots, y_n)$.

2.2 Minimax Solution with One Competitor

It would be natural to start off with the case when there is only one future competitor. We will denote this sole rival's location by y.

Note that, if this competitor places his vendor to our left, i.e. $y < x$, the customer will visit our store if and only if he is at the right of the middle point $(x+y)/2$. Therefore, our expected sales will be $E[h(|Z - x|)\mathbf{1}_{\{Z>(x+y)/2\}}]$. The domain $\{Z > (x + y)/2\}$ will decrease to the domain $\{Z \geq x\}$ when y tends to x from the left. Hence, clearly, $\inf_{y>x} D(x; y) = E\left[h(|Z - x|)\mathbf{1}_{\{Z\geq x\}}\right]$. Similarly considering y on the right hand side of x, one has $\inf_{y>x} D(x; y) = E\left[h(|Z - x|)\mathbf{1}_{\{Z\leq x\}}\right]$. Combining these we have

$$V(x) \stackrel{def}{=} \inf_y D(x; y) = \min\left(E\left[h(|Z - x|)\mathbf{1}_{\{Z\geq x\}}\right], E\left[h(|Z - x|)\mathbf{1}_{\{Z\leq x\}}\right]\right). \quad (1)$$

For the case $h \equiv 1$, Hotelling (1929) showed that the minimax optimal location for the first vendor is the median of F. We state it for completeness of our presentation.

Proposition 1 (Hotelling) *If $h(\cdot)$ is a constant function, the set of minimax optimals equals the set of medians of F.*

The assumption that h is a constant in Proposition 1 may not be always realistic. For example, we may want to consider the possibility that customers too far away do not make a visit to buy anything. In that case, the choice $h(d) = \mathbf{1}_{\{d \le d_0\}}$ may be better. Fortunately, it is possible to prove a general neat result on the minimax optimal location without requiring that $h(\cdot)$ is a constant. The additional assumption needed for this result is on F; see the proposition below. The examples that follow the proposition show that without this additional assumption, this general result is false.

Proposition 2 *If f is symmetric and unimodal around some m, then for general h, m is a minimax optimal.*

Proof: If we can show that $V_1(x) = E[h(|Z - x|)\mathbf{1}_{\{Z \ge x\}}]$ is a decreasing function of x for $x \ge m$, $V_2(x) = E[h(|Z - x|)\mathbf{1}_{\{Z \le x\}}]$ is an increasing function for $x \le m$, and $V_1(m) = V_2(m)$, then it will follow that m is a minimax optimal.
Consider a nonnegative measure u_x defined by

$$u_x([a, b]) = P\left[Z \in x + ([a, b] \cap [0, \infty))\right],$$

where "+" is the shift operator. Then, we have $V_1(x) = \int h(t)\mathrm{d}u_x(t)$. Let B be any Borel set of $\mathcal{R}$. The assumption that Z is unimodal with mode m implies that $u_x(B)$ is a decreasing function of x for $x \ge m$, and so $V_1(x)$ has the same property.
Similarly, one proves that $V_2(x)$ is increasing for $x \le m$. Furthermore, symmetry of Z easily implies $V_1(m) = V_2(m)$. This, therefore, completes the proof. □

The pleasant features of Proposition 2 disappear when F is not unimodal. Here is an example.

Example 1 Let $f(z) \propto \begin{cases} e^{-2(|z|-2)^2} & \text{if } |z| \le 2 \\ e^{-(|z|-2)^2/32} & \text{if } |z| \ge 2 \end{cases}$ and $h(d) = \begin{cases} e^{-2} & \text{if } 0 \le d \le 2 \\ e^{-d^2/2} & \text{if } d \ge 2 \end{cases}$.

Then Z is bimodal with modes ± 2, and the minimax optimals can be computed to be ± 3.3.

2.3 Minimax Solutions with many competitors

In contrast to the case when there is just one future rival, if two or more competitors exist, then the minimax formulation will no longer work. Let us explain what we mean. Suppose that there are two future competitors, say A and B. A sets up his store at y_1 on our left, and B at y_2 on our right. Now, if A and B move their stores closer and closer to us, the probability that a customer visits us, i.e. $P[Z \in (\frac{y_1+x}{2}, \frac{y_2+x}{2})]$, will decrease to 0. Hence, no matter where we place our vendor to start with, we always have $\inf_{y_1,y_2} D(x; y_1, y_2) = 0$. Therefore, the minimax approach is not interesting at all when $n \geq 2$. This motivates us to look for other reasonable formulations of the problem.

The minimax approach is tantamount to saying that we are not willing to make any assumptions about our future competitors. This conservative approach may be rational in some situations; but surely in certain other situations, we may be willing to make assumptions about our future rivals. We will now see that fortunately this Bayes approach remains meaningful for any $n \geq 1$ and, unlike the minimax approach, does not lead us to an uninteresting dead end if $n \geq 2$. We now move on to the Bayes approach.

3 Dynamic System in One Dimension

In this section, we will first propose a reasonable prior on the locations of the future vendors taking into account the locations of the already existing vendors. A general result on the Bayes solution for the first vendor when there is only one future competitor will be given. It will be proved that, under some conditions on f, the medians of f will be the optimal locations for the first vendor. This proves that Hotelling's result under the minimax formulation can be the dynamic Bayes solution as well. Analytical results in the dynamic set up for two or more future competitors seem to us to be impossible. So we present a numerical analysis of the optimal solutions for the case of two or more future rivals. The numerical results reveal the interesting phenomenon that the Bayes solution for the first vendor starts to move away from the modes/medians of f towards the boundary of the market as the number of the future rivals increases.

Let $g_i(y_i|x, y_0, \ldots, y_{i-1})$ and $G_i(y_i|x, y_0, \ldots, y_{i-1})$, simplified to g_i and G_i, denote, respectively, the density and the distribution of the location of the ith vendor given the already existing stores at $x, y_1, \ldots, y_{i-1}$. An intelligent vendor is likely to prefer those locations at which he can make a larger profit, using the knowledge about the locations of the already existing vendors. Accordingly, we will take

$$g_i(y_i|x, y_1, \ldots, y_{i-1}) \propto E_{Z\sim F}\left[h(|Z-y_i|)\mathbf{1}_{\{|Z-y_i|<min_{0\leq j\leq i-1}|Z-y_j|\}}\right]\mathbf{1}_{\{y_i\in S\}}, \tag{2}$$

where $y_0 = x$. We will see that, under a very weak assumption, namely, $E_{Z\sim F}(|Z|) < \infty$, this density is well-defined, i.e., the integral of the right hand side of (2) over y_i' s is finite. Hence, $D_n(x)$, our expected sales when we build our store at x and believe there will be n future rivals, can be expressed as

$$D_n(x) = E_{Z\sim F,(Y_1|X=x)\sim G_1,\cdots,(Y_n|X=x,Y_1,\cdots Y_{n-1})\sim G_n}\left[h(|Z-x|)\mathbf{1}_{\{|Z-x|<\min_{1\leq i\leq n}|Z-Y_i|\}}\right]. \tag{3}$$

The Bayes solution x_n, the value of x which maximizes the utility function $D_n(x)$, usually has no problem with its existence, e.g., if the market S is compact or closed and h is continuous. Nonetheless, the uniqueness of the Bayes solution is an exception rather than a rule. Note the complexity of the utility function in (3). It is not surprising that analytical solutions are so difficult in this dynamic set up.

For $h(\cdot) \equiv 1$, (2) and (3) are simplified to

$$g_i(y_i|x, y_1, \cdots, y_{i-1}) \propto P_{Z\sim F}\left(|Z-y_i| < \min_{0\leq j\leq i-1}|Z-y_j|\right)\mathbf{1}_{\{y_i\in S\}}, \quad \text{and} \tag{4}$$

$$D_n(x) = P_{Z\sim F,(Y_1|X=x)\sim G_1,\cdots,(Y_n|X=x,Y_1,\cdots Y_{n-1})\sim G_n}\left(|Z-x| < \min_{1\leq i\leq n}|Z-Y_i|\right). \tag{5}$$

We do assume $h(\cdot) \equiv 1$ in the rest of Section 3.

3.1 One Competitor

We start with the case of one competitor. First, we intend to show that g_1 in (4) is well-defined. Before doing that, we need to state a standard fact.

Fact 1 *Suppose $Z \sim F$ and $E(|Z|)$ is finite. Then for any given x, $\int_{-\infty}^{x} F(t)\mathrm{d}t$ and $\int_{x}^{\infty}(1-F(t))\mathrm{d}t$ are finite, and moreover*

$$E(Z) = x - \int_{-\infty}^{x} F(t)\mathrm{d}t + \int_{x}^{\infty}(1-F(t))\mathrm{d}t.$$

Remark: Note that when the market is bounded, the finite first moment of Z is automatically satisfied.

Lemma 1 *Let the customer's location Z be distributed as F with density f. If F has a finite first moment, then $g_1(y_1|x)$ in (4) is well defined.*

Proof: When $n = 1$, $g_1(y_1|x)$ in (4) is simplified to

$$g_1(y_1|x) \propto P_{Z\sim F}\left(|Z - y_1| < |Z - x|\right) = F(\frac{x+y_1}{2})\mathbf{1}_{\{y_1<x\}} + (1 - F(\frac{x+y_1}{2}))\mathbf{1}_{\{y_1>x\}}.$$

Thus, the normalizing constant equals

$$\begin{aligned}&\int_S \left(F(\frac{x+y_1}{2})\mathbf{1}_{\{y_1<x\}} + (1 - F(\frac{x+y_1}{2}))\mathbf{1}_{\{y_1>x\}}\right) \mathrm{d}y_1 \\ &= 2\int_{\frac{S+x}{2}} \left(F(t)\mathbf{1}_{\{t<x\}} + (1 - F(t))\mathbf{1}_{\{t>x\}}\right) \mathrm{d}t \ \le \ 2\int_R \left(F(t)\mathbf{1}_{\{t<x\}} + (1 - F(t))\mathbf{1}_{\{t>x\}}\right) \mathrm{d}t,\end{aligned}$$

which is finite by Fact 1. This proves Lemma 1. □

By a similar argument, after some algebra,

$$\begin{aligned} D_1(x) &= \int_S P_{Z\sim F}(|Z - x| < |Z - y_1|)g_1(y_1|x)\mathrm{d}y_1 \\ &= \frac{\int_{\frac{S+x}{2}} F(t)(1 - F(t))\mathrm{d}t}{\int_{\frac{S+x}{2}} \left(F(t)\mathbf{1}_{\{t<x\}} + (1 - F(t))\mathbf{1}_{\{t>x\}}\right) \mathrm{d}t}. \end{aligned} \tag{6}$$

The following question arises naturally.

Question: Hotelling (1929) showed that if h is a constant, then any median of F is a minimax solution. Is the median of F a Bayes solution as well under some general conditions on F in our dynamic set up?

Theorem 1 and Theorem 2 below address this question. Notice that we state the results separately for a bounded market and an unbounded market. Though these two theorems tell us the same story, more assumptions on f, however, are required if the market is bounded.

Theorem 1 *Suppose the market $S = R$, and F has a finite first moment. Then the set of Bayes solutions for the first vendor is exactly the set of medians of F.*

Proof: It is easy to see that the numerator of (6) does not depend on x when $S = R$. Hence, we only have to prove that the denominator of (6) is minimized at any median of F. This can be seen by simply taking a derivative. We omit that calculation. □

Theorem 2 *Suppose the market S is bounded, say $[-1, 1]$, and f is symmetric and unimodal around* 0. *Then* 0 *is the unique Bayes solution of the first vendor.*

Proof: When $S = [-1, 1]$, (6) becomes

$$D_1(x) = \frac{\int_{\frac{-1+x}{2}}^{\frac{1+x}{2}} F(t)(1 - F(t))\mathrm{d}t}{\int_{\frac{-1+x}{2}}^{x} F(t)\mathrm{d}t + \int_{x}^{\frac{1+x}{2}} (1 - F(t))\mathrm{d}t}. \tag{7}$$

Let us denote the numerator and denominator by $V(x)$ and $W(x)$, respectively. If we can show that $V(x)$ has a maximum and $W(x)$ has a unique minimum at 0, then the theorem will be proved.

By the fundamental theorem of calculus, one has

$$V'(x) = \frac{1}{2}\left(F(\frac{1+x}{2})(1 - F(\frac{1+x}{2})) - F(\frac{-1+x}{2})(1 - F(\frac{-1+x}{2}))\right). \tag{8}$$

Using the symmetry of f and the fact that $p(1-p)$ is symmetric and unimodal with mode at $p = 1/2$, it can be shown that $V'(x) \geq 0$ when $-1 \leq x \leq 0$ and $V'(x) \leq 0$ when $0 \leq x \leq 1$, which implies $V(x)$ is maximized at 0.

Now, let us look at $W(x)$. Again, by the fundamental theorem of calculus and the symmetry of f, one has

$$\begin{aligned} W'(x) &= \frac{1}{2}\left(4F(x) - 1 - F(\frac{1+x}{2}) - F(\frac{-1+x}{2})\right) && (9) \\ &= \frac{1}{2}\left(3(F(x) - F(0)) + (F(\frac{1-x}{2}) - F(0)) - (F(\frac{1+x}{2}) - F(x))\right). && (10) \end{aligned}$$

Next, by virtue of the assumption that f is unimodal around 0, F is convex on $[-1, 0]$ and concave on $[0, 1]$; hence one obtains that $W'(x) > 0$ for all $-1 \leq x < 0$ and $W'(x) < 0$ for all $0 < x \leq 1$. This implies $W(x)$ has a unique minimum at 0.

This completes the proof. □

The next example shows that the unimodality assumption of Theorem 2 cannot be relaxed.

Example 2 Let $S = [-1, 1]$ and $f(x) = \frac{3}{2}x^2 \mathbf{1}_{\{|x| \leq 1\}}$. Note that f is symmetric around 0 but not unimodal. When $S = R$, the utility function is

$$D_1(x) = \frac{12}{7(x^4 + 3)},$$

which is maximized at 0. But when $S = [-1, 1]$, the utility function becomes

$$D_1(x) = \frac{1 - \frac{1}{448}(7x^6 + 35x^4 + 21x^2 + 1)}{2(1 + \frac{1}{32}(15x^4 - 6x^2 - 1))},$$

which has a maximum at $x = \pm 0.36$. Hence, 0, the only median of f, is no longer the optimal location for the first vendor.

3.2 With Many Competitors

We now consider the case when there are two or more future competitors. It is again easily shown that $g_i(y_i|x, y_1, \ldots, y_{i-1})$ is well-defined for any $i \geq 2$ if $E(|Z|) < \infty$.

As one can see from (4), the prior on the location of the ith future vendor depends on the locations $x, y_1, \ldots, y_{i-1}$ of the already existing stores. The joint distribution of $(Y_1, \ldots, Y_n)$ is too complicated. This makes any theoretical analysis extremely difficult. In this section, we will give some heuristic arguments and a numerical analysis for the examples where $S = [-1, 1]$, $F = \text{Unif}[-1, 1]$, Beta(2, 2) on $[-1, 1]$, and Beta(4, 4) on $[-1, 1]$, and g_i is as defined in (4) with $n = 1$, 2, 3, and 4. The Bayes solutions and the utility functions are shown in Figure 1. Although the analysis is numerical, the findings are interesting.

We see that the Bayes solutions for the first vendor start to move towards the boundary values ± 1, as the number of the future competitors increases. A possible heuristic explanation is that suppose there are two future competitors to enter the market. If we set our vendor at the median(mode) of f, the first future rival will most likely set up his vendor at our immediate right or left rather than closer to the boundaries; say, he sets up his vendor at our immediate right. Then, the second rival will most likely set up his vendor at our immediate left. This will result in a small amount of sales for the first vendor. Thus it may be not productive for the first vendor to locate himself at the center and indeed he may be better off moving towards the boundary to alleviate the squeezing effect.

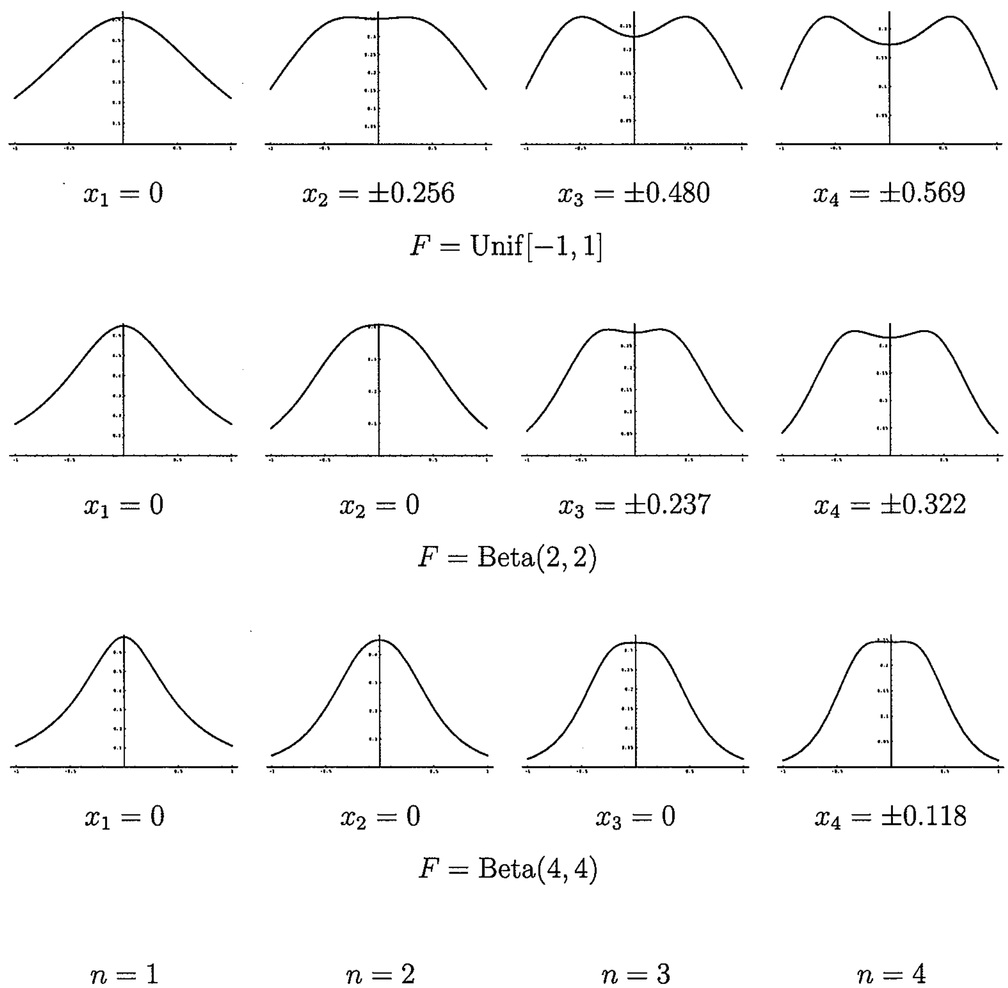


Figure 1: Utility functions and the dynamic Bayes Solutions for the first vendor when $S = [-1, 1]$ and g is as defined in (4).

We can observe from the plots a threshold value of n for x_n to move away from 0 towards the boundary values ± 1 and the threshold relates to the degree of the sharpness of f at 0 and the rate at which f vanishes at ± 1. The sharper f is at 0 and the faster f vanishes at ± 1, the larger the threshold will be. However, what seems to be most interesting is the phenomenon that the dynamic Bayes solution x_n always seems to move towards the boundary when n is large enough, regardless of how fast the customer density drops to zero at the boundary points. This phenomenon of movement to the boundary when there are a threshold number of future competitors is qualitatively interesting.

4 Nondynamic System in One Dimension

In the last section, the distribution imposed on the location of the ith rival depends on the locations of the already existing stores. This complicates the problem very much from a theoretical viewpoint, especially when there is more than one future competitor. So it is interesting that if we adopt a nondynamic setting, a number of additional analytical results can be obtained, and most interestingly, some of these results are similar to the ones in the dynamic setting. Suppose the future rivals decide on their locations independently, according to some distribution G which has a density g; symbolically, $Y_1, \ldots, Y_n \overset{i.i.d.}{\sim} G$. Then the first vendor's $D_n(x)$ can be expressed as

$$D_n(x) = P_{Z \sim F, Y_1, \cdots, Y_n \overset{i.i.d.}{\sim} G}\left(|Z - x| < \min_{1 \leq i \leq n} |Z - Y_i| \right). \tag{11}$$

Again, the Bayes solution x_n has no problem regarding its existence but not the uniqueness. This is especially so when n is moderately large. In fact, a very interesting threshold phenomenon again holds. x_n, under this nondynamic setting, behaves like in the dynamic system. The median (mode) of f is the Bayes solution for the first vendor only when n is small, and x_n starts to move towards the boundary of the market when n gets larger. In addition, we can pin down x_n very precisely via an asymptotic expansion.

4.1 Bayes Solution with Many Competitors

We consider only the case where there exist two or more future competitors. The case of one competitor can be seen in Tsai, DasGupta, and Zidek (1999). We take S to be a bounded interval $[a, b]$; without loss of generality, we may choose $[a, b]$ to be $[-1, 1]$. Let also G be the uniform distribution on $[-1, 1]$. The case of a general G would be briefly commented on.

Note that a customer located at z on our left, i.e. $z < x$, will visit us if and only if there are no competitors located between $2z - x$ and x and, similarly, if the customer is on our right. Consequently, the general formula for the utility function in (11) becomes

$$D_n(x) = \int_{-\infty}^{x} (G(2z - x) + 1 - G(x))^n f(z) \mathrm{d}z + \int_{x}^{\infty} (G(x) + 1 - G(2z - x))^n f(z) \mathrm{d}z. \tag{12}$$

For $S = [-1, 1]$ and $G = U[-1, 1]$, (12) simplifies to

$$\begin{aligned} D_n(x) &= (\frac{1-x}{2})^n F(\frac{x-1}{2}) + \int_0^{\frac{1+x}{2}} (1-t)^n f(x-t)\mathrm{d}t \\ &\quad + \int_0^{\frac{1-x}{2}} (1-t)^n f(x+t)\mathrm{d}t + (\frac{1+x}{2})^n (1 - F(\frac{1+x}{2})) \qquad (13) \\ &= (I)_1 + (I)_2 + (I)_3 + (I)_4. \qquad (14) \end{aligned}$$

One can observe from (13) that $D_n(x)$ is continuous and bounded for $x \in [-1, 1]$. Consequently, a Bayes solution exists.

The actual analysis in the remainder of this section will get rather technical. Thus, it would be helpful to have a preview of the principal phenomena that arise from the results. The preview would be useful to appreciate how the case of many competitors differs from the case of one competitor and to provide an explanation for the reason the differences arise. We now give a brief preview.

4.1.1 Preview

The key difference between the cases of a small n and a large n is that for large n (for some F's), the first vendor's optimal strategy is to move away from the center and to choose a location near the boundary. We will show that if F is uniform, then the first vendor should set up his store at 0 only if $n \leq 4$, and move to locations $\pm x_n$ near the boundary if $n \geq 5$ ($n = 5$ thus being a threshold), where x_n is a root of a complicated equation; but there is a nice approximation, namely, $x_n = 1 - \frac{4}{n} + o(\frac{1}{n})$, $n \to \infty$. Heuristically, one can guess that the Bayes solution will move from zero starting at $n = 5$ by noting that the second derivative of the utility function satisfies $D_n''(0) = \frac{n(n-4)}{2^{n+1}}$ which is positive for $n \geq 5$ and, from symmetry, $D_n'(0) = 0$ for any n. Figure 2 illustrates the shape of $D_n(x)$.

For general F, the final results are similar, although at first they seem counterintuitive. For instance, an intuition might suggest that if F has a unique mode m in the interior of S, then the first vendor ought to be located at that unique mode. For large n, this is not exactly correct. If the unique mode m is a very pronounced one, then indeed the first vendor will stay there. But if the density f at a boundary point is only a bit less than the density at that interior mode, then for large n, the first vendor should move closer to that boundary point. We make precise what the meaning of "only a bit less" is. For example, if f is unimodal with a mode at 0, and, $f(1) \geq .9366 f(0)$, then for large n, the first vendor should abandon the center and seek a location x_n near the boundary 1, where, fortunately, the approximation $x_n = 1 - \frac{4}{n} + o(\frac{1}{n})$, $n \to \infty$, is valid for general F, not just when F is uniform.

Intuitively, if n is large, then the first vendor would be sandwiched between competitors on his left as well as on his right if he were to select a central location. By moving closer to the boundary, he can essentially eliminate competition from one side, say his right. If there are still a good number of customers on his right, they are all necessarily his clients, and, in this way, he is better off moving towards the boundary than staying at the central location.

We describe precisely this verbal exposition. If $\{x_n\}_{n=1}^{\infty}$ is any sequence of optimal locations of the first vendor, then every accumulation point of $\{x_n\}$ must be ± 1 or an interior mode of f. Hence, asymptotically, the first vendor's task is simple: consider only those interior locations (if any) that are modes of f, and also consider the two points $1 - \frac{4}{n}$ and $-1 + \frac{4}{n}$. If he limits his search to just these locations, he would be approximately correct.

We will present the special case F = uniform first. There are two reasons for doing this. Firstly, it is an important special case and, secondly, there is an explicit threshold result for the uniform case.

We should remind the reader at this point that we have already assumed G to be uniform on $[-1, 1]$ for the entire Section 4.1.

4.1.2 The case of Uniform F

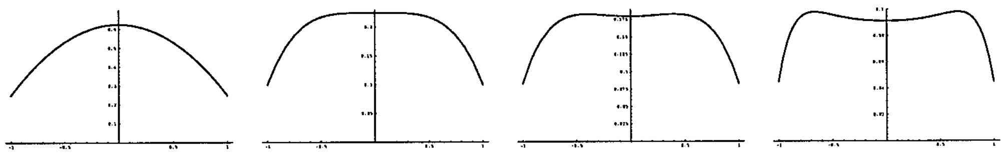
Theorem 3 *Suppose that F and G are both uniform on $[-1, 1]$. Then,*

(a) *when $n \leq 4$, there is a unique Bayes solution $x_n = 0$,*

(b) *when $n \geq 5$, there are exactly two Bayes solutions, $-x_n$ and x_n, for some $0 < x_n < 1$, and*

(c) $x_n = 1 - \frac{4}{n} + o(\frac{1}{n})$, *as $n \to \infty$.*

Some plots of the utility functions for F and G uniform on $[-1, 1]$ are shown in Figure 2 for illustration. Note the similarity to Figure 1 in the dynamic case.



$n = 1$ $\quad$ $n = 4$ $\quad$ $n = 5$ $\quad$ $n = 10$.

Figure 2: Plots of $D_n(x)$ when F and G are both uniform distributions on $[-1, 1]$.

Proof: Step 1: When $F = U[-1, 1]$, $D_n(x)$ can be simplified to

$$D_n(x) = \frac{1}{2(n+1)} \left[2 - (n+2)\left((\frac{1+x}{2})^{n+1} + (\frac{1-x}{2})^{n+1}\right) + (n+1)\left((\frac{1+x}{2})^{n} + (\frac{1-x}{2})^{n}\right)\right]. \tag{15}$$

Consequently,

$$D_n'(x) = \frac{1}{4}\left[-(n+2)\left((\frac{1+x}{2})^{n} - (\frac{1-x}{2})^{n}\right) + n\left((\frac{1+x}{2})^{n-1} - (\frac{1-x}{2})^{n-1}\right)\right]. \tag{16}$$

Since $D_n(x)$ is symmetric, we only need to check the sign of $D_n'(x)$ for x in $(0, 1)$ to find the Bayes solutions.

Define $a_n(x) = (\frac{1+x}{2})^n - (\frac{1-x}{2})^n$ for $n \geq 1$ and $b_n(x) = a_n(x)/a_{n-1}(x)$ for $n \geq 2$. The following recurrence equations for $a_n(x)$ and $b_n(x)$ can be easily derived:

$$a_n(x) = a_{n-1}(x) - \frac{1-x^2}{4} a_{n-2}(x) \quad \text{and} \quad b_n(x) = 1 - \frac{1-x^2}{4}\frac{1}{b_{n-1}(x)}, \quad \text{for all } n \geq 3. \tag{17}$$

Also note that

$$D_n'(x) >, = \text{ or } < 0 \quad \text{if and only if} \quad b_n(x) <, = \text{ or } > \frac{n}{n+2}, \quad \text{respectively.} \tag{18}$$

By a straight application of L'Hospital's rule and some algebra, the following claims can be obtained:

Claim (A): For each fixed $n \geq 3$, $b_n(x)$ is a strictly increasing function of x in $(0, 1)$.

Claim (B): $\lim_{x \to 0} b_n(x) = \frac{n}{2(n-1)}$ and $\lim_{x \to 1} b_n(x) = 1$.

Claim (C): For each fixed $x \in (0, 1)$, $\{b_n(x)\}_{n=2}^{\infty}$ is a strictly decreasing sequence of n.

Step2: (Proof of part (a) of Theorem 3). To prove that 0 is the unique Bayes solution, it is enough to show that $D_n'(x) < 0$ for all $x \in (0, 1)$. By (18), this is equivalent to

$b_n(x) > \frac{n}{n+2}$ for all $x \in (0,1)$. Because this result needs to be verified for only four values of n, one can show it by direct computation. We omit that verification.

Step 3: (Proof of part (b) of Theorem 3). If we can show that there exists $0 < x_n < 1$ such that

$$b_n(x_n) = \frac{n}{n+2},\ b_n(x) < \frac{n}{n+2} \text{ for } 0 < x < x_n, \text{ and } b_n(x) > \frac{n}{n+2} \text{ for } x_n < x < 1, \tag{19}$$

it will follow from (18) that $D_n(x)$ strictly increases in the interval $(0, x_n)$ and strictly decreases in the interval $(x_n, 1)$. This implies x_n and $-x_n$ are the only two Bayes solutions. The existence of such x_n follows from claims (A), (B), and (C) stated above.

Step 4: (Proof of part (c) of Theorem 3). For $n \geq 5$, plugging $x = x_n$ into (17), we have

$$b_n(x_n) = \frac{n}{n+2} = 1 - \frac{1-x_n^2}{4}\frac{1}{b_{n-1}(x_n)} \iff (n+2)(1-x_n) = \frac{8b_{n-1}(x_n)}{1+x_n} \tag{20}$$

Furthermore, by claims (A), (B), and (C) given above, we have $\lim_{n\to\infty} b_{n-1}(x_n) = 1$. Hence, from (20), $x_n = 1 - \frac{4}{n} + o(\frac{1}{n})$ as was claimed, and this completes the proof of part (c). □

In reality, the number of the future competitors is usually unknown. Therefore, we may want to treat n as a parameter and impose a prior on it. Suppose for example that $n \sim Poisson(\lambda)$, i.e. $P(n = k) = e^{-\lambda}\lambda^k/k!, k = 0, 1, \ldots$. Note that under this model, it is possible that there will be no future competitors. Obviously $D_0(x) \equiv 1$. Hence, the utility function for the first vendor becomes

$$\begin{aligned} D_\lambda(x) &= \sum_{n=0}^{\infty} e^{-\lambda}\frac{\lambda^n}{n!}D_n(x) \\ &= \frac{1}{\lambda} + \frac{1}{2}\left(\frac{1-x}{2}e^{-\frac{1-x}{2}\lambda} + \frac{1+x}{2}e^{-\frac{1+x}{2}\lambda}\right) - \frac{1}{2\lambda}\left(e^{-\frac{1-x}{2}\lambda} + e^{-\frac{1+x}{2}\lambda}\right). \end{aligned} \tag{21}$$

Notice that $D_\lambda(x)$ is an even function of x. Therefore, if x_λ is a Bayes solution for the first vendor, so is $-x_\lambda$. Interestingly, even under this model, there is a threshold result:

Theorem 4 *Suppose that F and G are both uniform on $[-1, 1]$, and n is Poisson distributed with parameter λ. Then we have*

(a) *when $\lambda \leq 6$, the Bayes solution is unique and it is 0,*

(b) *when $\lambda > 6$, there are exactly two Bayes solutions, x_λ and $-x_\lambda$, where $0 < x_\lambda < 1$, and*

(c) $x_\lambda = 1 - \frac{4}{\lambda} + o(\frac{1}{\lambda})$, *as $\lambda \to \infty$.*

Note: It is interesting that the threshold value for λ is 6, a whole number.

Proof: Refer to Tsai et al. (1999) for the proof. □

5 Two-Dimensional Markets

Instead of a linear market, it may be more realistic to consider a two-dimensional market. We will continue to use the notation introduced in Section 2.1 for one-dimensional markets. Also, we still assume that a buyer visits the closest store. As before, we would like to investigate these two-dimensional markets from three perspectives: minimax and the dynamic as well as the nondynamic set up.

5.1 Minimax Solutions in Two Dimensions

As we did in the case of a linear market, first let us consider the case of only one future rival. Let us denote the future rival's location by $\underset{\sim}{y}$. The expected sales for the first vendor then becomes

$$D(\underset{\sim}{x}, \underset{\sim}{y}) = E(h(\|\underset{\sim}{Z} - \underset{\sim}{x}\|)\mathbf{1}_{\{\|\underset{\sim}{Z}-\underset{\sim}{x}\|<\|\underset{\sim}{Z}-\underset{\sim}{y}\|\}}) = E(h(\|\underset{\sim}{Z} - \underset{\sim}{x}\|)\mathbf{1}_{\{\underset{\sim}{Z}\in H((\underset{\sim}{x}+\underset{\sim}{y})/2, \underset{\sim}{x}-\underset{\sim}{y})\}}), \tag{22}$$

where $H(\underset{\sim}{p}, \underset{\sim}{n})$ is the open half space containing the point $\underset{\sim}{p} + \underset{\sim}{n}$ with its boundary line passing through the point $\underset{\sim}{p}$ and perpendicular to the vector $\underset{\sim}{n}$, i.e., $H(\underset{\sim}{p}, \underset{\sim}{n}) = \{\underset{\sim}{z} : \underset{\sim}{n} \cdot (\underset{\sim}{z} - \underset{\sim}{p}) > 0\}$. Let $e_\theta = (\cos\theta, \sin\theta)$. By a simple argument, one can see that the minimax solutions are points maximizing

$$V(\underset{\sim}{x}) \overset{def}{=} \min_{\underset{\sim}{y}\neq\underset{\sim}{x}} D(\underset{\sim}{x}, \underset{\sim}{y}) = \min_{\theta} E(h(\|\underset{\sim}{Z} - \underset{\sim}{x}\|)\mathbf{1}_{\{\underset{\sim}{Z}\in H(\underset{\sim}{x}, e_\theta)\}}) = \min_{\theta} V_\theta(\underset{\sim}{x}), \tag{23}$$

$$\text{where} \quad V_\theta(\underset{\sim}{x}) \overset{def}{=} E(h(\|\underset{\sim}{Z} - \underset{\sim}{x}\|)\mathbf{1}_{\{\underset{\sim}{Z}\in H(\underset{\sim}{x}, e_\theta)\}}). \tag{24}$$

When $h(\cdot) \equiv$ a constant, say 1, from (23), the minimax solutions are those points which maximize $V(\underset{\sim}{x}) = \min_\theta P(Z \in H(\underset{\sim}{x}, e_\theta))$. Points which maximize $\min_\theta P_{\underset{\sim}{Z}\sim F}(\underset{\sim}{Z} \in H(\underset{\sim}{x}, e_\theta))$ are, in fact, defined as the halfspace medians for a bivariate distribution F.

Indeed, the halfspace median, a generalization of the univariate median to higher dimensions due to Tukey (1975) and Donoho (1982), is motivated by this minimaxity result (see Small (1990) for details). Therefore, we conclude a result similar to Proposition 1: if $h(\cdot)$ is a constant function, the set of minimax optimals for the first vendor equals the set of halfspace medians of F.

If we intend to discard the assumption that $h(\cdot) \equiv$ a constant and still want to get a clean result, more conditions on F will be needed. The following is a general result analogous to Proposition 2. It is about the best possible clean extension of Hotelling's result to the case of two dimensions.

Proposition 3 *If f is spherically symmetric and unimodal around some $\underset{\sim}{m}$, then for a general h, $\underset{\sim}{m}$ is a minimax optimal.*

Proof: Without loss of generality, we can take $\underset{\sim}{m}$ to be the origin.

Step 1: We claim $V_\theta(re_\theta)$ is a decreasing function of r in $[0, \infty)$ for all θ; recall that the formula $V_\theta(\cdot)$ is as defined in (24). It is obvious that

$$V_\theta(re_\theta) = \int_{\underset{\sim}{z} \in H(re_\theta, e_\theta)} h(\|\underset{\sim}{z} - re_\theta\|) f(\underset{\sim}{z}) \mathrm{d}\underset{\sim}{z} = \int_{\underset{\sim}{t} \in H(0, e_\theta)} h(\|\underset{\sim}{t}\|) f(\underset{\sim}{t} + re_\theta) \mathrm{d}\underset{\sim}{t}. \tag{25}$$

Now observe $\|\underset{\sim}{t} + re_\theta\|^2 = \|\underset{\sim}{t}\|^2 + r^2 + 2r\underset{\sim}{t} \cdot e_\theta$, which is an increasing function of r for $\underset{\sim}{t} \in H(\underset{\sim}{0}, e_\theta)$. Therefore, the assumption that f is spherically symmetric and unimodal around the origin implies that for any fixed θ, $f(\underset{\sim}{t} + re_\theta)$ is a decreasing function of r for $\underset{\sim}{t} \in H(\underset{\sim}{0}, e_\theta)$. Applying this to (25), we prove this claim.

Step 2: We claim there exists a constant c such that $V_\theta(\underset{\sim}{0}) = c$ for all θ. This is obvious; since f is spherically symmetric around the origin, it is evident that $V_\theta(\underset{\sim}{0})$ does not depend on θ.

Step 3: Now we shall show $V(\underset{\sim}{x}) \leq V(\underset{\sim}{0})$ for all $\underset{\sim}{x}$ in R^2. For any $\underset{\sim}{x}$ in R^2, there exist $r_0 > 0$ and $0 \leq \theta_0 < 2\pi$ such that $\underset{\sim}{x} = r_0 e_{\theta_0}$. Therefore one has $V(\underset{\sim}{x}) = \min_\theta V_\theta(r_0 e_{\theta_0}) \leq V_{\theta_0}(r_0 e_{\theta_0})$, which is smaller than $V_{\theta_0}(\underset{\sim}{0})$ by Step 1. On the other hand, Step 2 implies $V(\underset{\sim}{0}) = \min_\theta V_\theta(\underset{\sim}{0}) = c = V_{\theta_0}(\underset{\sim}{0})$ for all θ_0. Therefore, we have $V(\underset{\sim}{x}) \leq V(\underset{\sim}{0})$ for all $\underset{\sim}{x}$. This completes the proof of Proposition 3. □

For reasons identical to the case of one dimension, the minimax formulation is uninteresting if the number of future rivals is more than 1. So again, we will now proceed to consideration of the Bayes approach. Theorem 3 below is a clean general result for the dynamic case.

5.2 Dynamic System in Two Dimensions

Throughout this section we will take $h(\cdot) \equiv 1$.

As we mentioned in Section 3, the dynamic model assumes that the future competitors decide where to set up their vendors based on the available information on the locations of the existing stores. We assume each competitor situates his business at location $\underset{\sim}{y}$ with a likelihood proportional to the profit that he will make if he sets up his vendor at $\underset{\sim}{y}$. Namely, as in Section 3,

$$g_i(\underset{\sim}{y}_i|\underset{\sim}{x}, \underset{\sim}{y}_1, \cdots, \underset{\sim}{y}_{i-1}) \propto P_{\underset{\sim}{Z}\sim F}\left(||\underset{\sim}{Z} - \underset{\sim}{y}_i|| < \min_{0\leq j\leq i-1} ||\underset{\sim}{Z} - \underset{\sim}{y}_j||\right) \mathbf{1}_{\{\underset{\sim}{y}_i \in S\}}, \tag{26}$$

where $\underset{\sim}{y}_0 = \underset{\sim}{x}$.

Similar to Proposition 3 in the last section but with less assumptions on f, the halfspace median will again be an optimal for the first vendor in the dynamic system with one future rival when f is spherically symmetric. Before proving this statement, we have to show that the density in (26) is well-defined.

Fact 2 *If $||\underset{\sim}{Z}||$ has a finite second moment, then we have*

$$\int_{R^2} P_{\underset{\sim}{Z}\sim F}\left(||\underset{\sim}{Z} - \underset{\sim}{y}_i|| < \min_{0\leq j\leq i-1} ||\underset{\sim}{Z} - \underset{\sim}{y}_j||\right) \mathrm{d}\underset{\sim}{y}_i < \infty.$$

Thus, $g_i(\underset{\sim}{y}_i|\underset{\sim}{x}, \underset{\sim}{y}_1, \cdots, \underset{\sim}{y}_{i-1})$ in (26) is well-defined.

Notice that for a one-dimensional market, the corresponding sufficient condition is a finite first moment of Z, while in a two-dimensional market, we seem to need a finite second moment.

Proof: Since the integrand is bounded by $P_{\underset{\sim}{Z}\sim F}\left(||\underset{\sim}{Z} - \underset{\sim}{y}_n|| < ||\underset{\sim}{Z} - \underset{\sim}{x}||\right)$, it is enough to give the proof for the case $i = 1$. Moreover, after performing a location transformation, we may assume $\underset{\sim}{x} = \underset{\sim}{0}$.

Now note that $P_{\underset{\sim}{Z}\sim F}\left(||\underset{\sim}{Z} - \underset{\sim}{y}|| < ||\underset{\sim}{Z}||\right) \leq P_{\underset{\sim}{Z}\sim F}\left(||\underset{\sim}{Z}|| > \frac{||\underset{\sim}{y}||}{2}\right)$. Therefore, we have

$$\int_{R^2} P_{\underset{\sim}{Z}\sim F}\left(||\underset{\sim}{Z} - \underset{\sim}{y}|| < ||\underset{\sim}{Z}||\right) \mathrm{d}\underset{\sim}{y} \leq 8\pi \int_0^\infty P_{\underset{\sim}{Z}\sim F}\left(||\underset{\sim}{Z}|| > r\right) r\mathrm{d}r,$$

which is finite since $E(||\underset{\sim}{Z}||^2) < \infty$. $\square$

Next, some formulae for $D(\underline{x})$, which will be used to prove Theorem 5, are presented below.

Lemma 2 *Suppose f is spherically symmetric around $\underline{0}$ with a finite second moment, and the only one future competitor sets up his vendor according to the density given in (26). Then the expected profit for the first vendor when he sets up his business at $\underline{x}$ is given by*

$$D(\underline{x}) = \frac{\int_{\underline{y}\in S} p(\frac{d}{2})(1-p(\frac{d}{2}))\mathrm{d}\underline{y}}{\int_{\underline{y}\in S} p(\frac{d}{2})\mathrm{d}\underline{y}}, \tag{27}$$

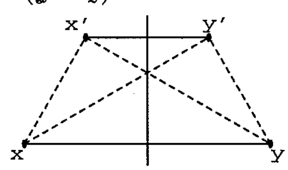
where $d = \frac{||\underline{x}||^2-||\underline{y}||^2}{||\underline{y}-\underline{x}||}$ and $p(d) = P_{\underline{Z}\sim F}(Z_1 \le d)$.

Furthermore, we can also express $D(\underline{x})$ in the alternative form:

$$D(\underline{x}) = \frac{\int_{\underline{y}'\in\underline{x}+S} p(\frac{||\underline{y}'||}{2})(1-p(\frac{||\underline{y}'||}{2}))\,\frac{|(\underline{y}'-2\underline{x})\cdot\underline{y}'|}{||\underline{y}'||^2}\mathrm{d}\underline{y}'}{\int_{\underline{y}'\in\underline{x}+D(0,||\underline{x}||)} p(\frac{||\underline{y}'||}{2})\,\frac{|(\underline{y}'-2\underline{x})\cdot\underline{y}'|}{||\underline{y}'||^2}\mathrm{d}\underline{y}' + \int_{\underline{y}'\in\underline{x}+(S\setminus D(0,||\underline{x}||))}(1-p(\frac{||\underline{y}'||}{2}))\,\frac{|(\underline{y}'-2\underline{x})\cdot\underline{y}'|}{||\underline{y}'||^2}\mathrm{d}\underline{y}'}, \tag{28}$$

where $D(\underline{c}, r) = \{\underline{x} : ||\underline{x}-\underline{c}|| \le r\}$. *(note that $\underline{y}'$ in formula (28) is not to be understood as a transpose of $\underline{y}$).*

Proof: Let $\underline{x}' = (0,0)$ and for any given $\underline{x}$ and $\underline{y}$, $\underline{y}' = \frac{||\underline{y}||^2-||\underline{x}||^2}{||\underline{y}-\underline{x}||^2}\cdot(\underline{y}-\underline{x})$. Notice that two vendors located at $\underline{x}$ and $\underline{y}$ are visited by the same set of customers when we "relocate" the vendors to $\underline{x}'$ and $\underline{y}'$ (refer to the picture on the right). With careful algebra, one has:



$$||\underline{Z}-\underline{y}|| < ||\underline{Z}-\underline{x}|| \iff \begin{cases} ||\underline{Z}-\underline{y}'|| < ||\underline{Z}-\underline{x}'||, & \text{if } ||\underline{x}'-\underline{x}|| < ||\underline{x}'-\underline{y}|| \\ ||\underline{Z}-\underline{y}'|| > ||\underline{Z}-\underline{x}'||, & \text{if } ||\underline{x}'-\underline{x}|| > ||\underline{x}'-\underline{y}|| \end{cases}.$$

Therefore, by the assumption of spherical symmetry of f, we have

$$P_{\underline{Z}\sim F}(||\underline{Z}-\underline{y}|| < ||\underline{Z}-\underline{x}||) = \begin{cases} 1-p(\frac{||\underline{y}'||}{2}), & \text{if } \underline{y} \notin D(\underline{0}, ||\underline{x}||) \\ p(\frac{||\underline{y}'||}{2}), & \text{if } \underline{y} \in D(\underline{0}, ||\underline{x}||) \end{cases}.$$

Then (27) follows immediately from the facts that $D(\underline{x}) = \frac{\int_S P(||\underline{Z}-\underline{x}||<||\underline{Z}-\underline{y}||)P(||\underline{Z}-\underline{y}||<||\underline{Z}-\underline{x}||)\mathrm{d}\underline{y}}{\int_S P(||\underline{Z}-\underline{y}||<||\underline{Z}-\underline{x}||)\mathrm{d}\underline{y}}$, $||\underline{y}'|| = |\frac{||\underline{y}||^2-||\underline{x}||^2}{||\underline{y}-\underline{x}||}|$, and $p(-d) = 1-p(d)$.

As for (28), it can be derived by the change of variable $\underline{y} \to \underline{y}'$ with Jacobian equal to $\frac{|(\underline{y}-2\underline{x})\cdot\underline{y}'|}{||\underline{y}||^2}$, and by the fact that $\underline{y} \in D(\underline{0}, r)$ if and only if $\underline{y}' \in \underline{x} + D(\underline{0}, r)$ for each $r \ge 0$. This completes the proof of Lemma 2. □

Now we give a characterization of the optimal location of the first vendor under the dynamic Bayes set up when there is only one future rival.

Theorem 5 *Suppose the market $S = R^2$. Assume $\underset{\sim}{Z} \sim F$ and F is spherically symmetric around $\underset{\sim}{0}$ with a finite second moment, and the only one future competitor sets up his vendor according to the distribution defined in (26). Then $\underset{\sim}{0}$ is an optimal location for the first vendor.*

Remark 1 Note that this theorem addresses the market $S = R^2$ only. Recall the results in Section 3. When a one-dimensional market is the whole real line, the median is the optimum location for the first vendor. We do not need assumptions on the shape of f. However, if the one-dimensional market is bounded, the median is optimal only if f is symmetric and unimodal around the median. On the other hand, for a two-dimensional market, it seems that the halfspace median (which is just the point of symmetry) is always an optimum choice for the first vendor if f is spherically symmetric, without the unimodality assumption, regardless of whether the market is bounded or not. Numerical analysis shows that $\underset{\sim}{0}$ is the optimal location for the first vendor for the special examples $S = D(\underset{\sim}{0}, 1)$ and $f(r) \propto 1, r, \frac{1}{r}, r^2, r^{10}, r^{20}$, or $r(1-r)$. Note that in the special case if $f(r) \propto 1/r$ $(0 \le r \le 1)$, F is spherically symmetric, but not unimodal.

Proof of Theorem 5: Since f is spherically symmetric around $\underset{\sim}{0}$, $D(\underset{\sim}{x})$ depends only on $||\underset{\sim}{x}||$. Without loss of generality, we may assume $\underset{\sim}{x} = (r_x, 0)$, where $r_x \ge 0$. Moreover, when the market $S = R^2$, the domain $\underset{\sim}{x} + S$ is again the whole plane R^2. Therefore, (28) becomes, on algebra,

$$D(r_x) = \int_{(r,\theta)} p(\frac{r}{2})(1-p(\frac{r}{2}))|r - 2r_x\cos\theta| \mathrm{d}\theta \mathrm{d}r \Big/ \int_{(r,\theta)} \Big(p(\frac{r}{2})\mathbf{1}_{\{(r\cos\theta, r\sin\theta)\in D((r_x,0),r_x)\}} + (1-p(\frac{r}{2}))\mathbf{1}_{\{(r\cos\theta, r\sin\theta)\notin D((r_x,0),r_x)\}}\Big) |r - 2r_x\cos\theta| \mathrm{d}\theta \mathrm{d}r. \tag{29}$$

The proof will now require a very careful analysis. The following facts are useful in simplifying the integral in (29). We will not provide their proofs, which require only algebra and calculus.

(A) For every fixed $r < 2r_x$, we have $r - 2r_x\cos\theta \ge 0$ if and only if $\theta_0 \le \theta \le 2\pi - \theta_0$, where $\theta_0 = \arccos(r/2r_x)$; as a consequence, $\int_0^{2\pi} |r - 2r_x\cos\theta| \mathrm{d}\theta = (2\pi - 4\theta_0)r + 8r_x\sin\theta_0$.

(B) For every fixed $r \geq 2r_x$, we have $r - 2r_x\cos\theta \geq 0$ and, hence, $\int_0^{2\pi}|r - 2r_x\cos\theta|\mathrm{d}\theta = 2\pi r$.

(C) For every fixed $r < 2r_x$, we have $(r\cos\theta, r\sin\theta) \notin D((r_x, 0), r_x)$ if and only if $\theta_0 \leq \theta \leq 2\pi - \theta_0$. Consequently, $\int_{(r\cos\theta, r\sin\theta)\in D((r_x,0),r_x)}|r - 2r_x\cos\theta|\mathrm{d}\theta = -2\theta_0 r + 4r_x\sin\theta_0 \geq 0$ and $\int_{(r\cos\theta, r\sin\theta)\notin D((r_x,0),r_x)}|r - 2r_x\cos\theta|\mathrm{d}\theta = (2\pi - 2\theta_0)r + 4r_x\sin\theta_0 \geq 0$.

(D) For every fixed $r \geq 2r_x$, we have $(r\cos\theta, r\sin\theta) \in D((r_x, 0), r_x)$ for all θ and, therefore, $\int_{(r\cos\theta, r\sin\theta)\in D((r_x,0),r_x)}|r - 2r_x\cos\theta|\mathrm{d}\theta = 2\pi r$.

By the facts listed above, (29) becomes

$$
\begin{aligned}
D(r_x) &= \Big(\int_0^{2r_x} p(\frac{r}{2})(1 - p(\frac{r}{2}))((2\pi - 4\theta_0)r + 8r_x\sin\theta_0)\mathrm{d}r + \int_{2r_x}^{\infty} p(\frac{r}{2})(1 - p(\frac{r}{2}))2\pi r\mathrm{d}r\Big) \\
&\quad \Big/ \Big(\int_0^{2r_x} p(\frac{r}{2})(-2\theta_0 r + 4r_x\sin\theta_0)\mathrm{d}r + \int_0^{2r_x}(1 - p(\frac{r}{2}))((2\pi - 2\theta_0)r + 4r_x\sin\theta_0)\mathrm{d}r \\
&\qquad + \int_{2r_x}^{\infty}(1 - p(\frac{r}{2}))2\pi r\mathrm{d}r\Big) \\
&= \frac{\int_0^{\infty} p(\frac{r}{2})(1 - p(\frac{r}{2}))2\pi r\mathrm{d}r + \int_0^{2r_x} p(\frac{r}{2})(1 - p(\frac{r}{2}))(-4\theta_0 r + 8r_x\sin\theta_0)\mathrm{d}r}{\int_0^{\infty}(1 - p(\frac{r}{2}))2\pi r\mathrm{d}r + \int_0^{2r_x}(-2\theta_0 r + 4r_x\sin\theta_0)\mathrm{d}r}. \qquad (30)
\end{aligned}
$$

Now we will show that $D(0) \geq D(r_x)$ for every $r_x > 0$ which will prove the theorem. This is equivalent to the following inequality:

$$
\begin{aligned}
&\int_0^{\infty} p(\frac{r}{2})(1 - p(\frac{r}{2}))2\pi r\mathrm{d}r \cdot \int_0^{2r_x}(-2\theta_0 r + 4r_x\sin\theta_0)\mathrm{d}r \\
&\quad \geq \int_0^{\infty}(1 - p(\frac{r}{2}))2\pi r\mathrm{d}r \cdot \int_0^{2r_x} p(\frac{r}{2})(1 - p(\frac{r}{2}))(-4\theta_0 r + 8r_x\sin\theta_0)\mathrm{d}r. \qquad (31)
\end{aligned}
$$

By definition of $p(\cdot)$, we have $p(\frac{r}{2}) \geq \frac{1}{2}$ for every $r \geq 0$ which immediately implies that

$$\int_0^{\infty} p(\frac{r}{2})(1 - p(\frac{r}{2}))2\pi r\mathrm{d}r \geq \frac{1}{2}\int_0^{\infty}(1 - p(\frac{r}{2}))2\pi r\mathrm{d}r.$$

On the other hand, since $1 \geq 4p(\frac{r}{2})(1 - p(\frac{r}{2}))$ and $-2\theta_0 r + 4r_x\sin\theta_0 \geq 0$, one has

$$\int_0^{2r_x}(-2\theta_0 r + 4r_x\sin\theta_0)\mathrm{d}r \geq 2\int_0^{2r_x} p(\frac{r}{2})(1 - p(\frac{r}{2}))(-4\theta_0 + 8r_x\sin\theta_0)\mathrm{d}r.$$

Therefore, inequality (31) holds and, hence, $\underset{\sim}{0}$ is optimal for the first vendor. □

When there are two or more competitors, the search for the optimal location for the first vendor is a difficult task. We do not have any results for this case worth reporting, and are conducting some numerical analysis.

5.3 Nondynamic System in Two Dimensions

Analogous to one dimensional markets, the Bayes solutions for the first vendor in a planar market can possibly move towards the boundary of the market when the number of future competitors n is large. We will give some nice examples first and then state a general result.

Example 3 Let $S = [-1,1] \times [-1,1]$. Take F and G to be uniform distributions in S. Figure 3 illustrates the shape of the utility functions for some selected values of n. Some Bayes optimals $\underset{\sim}{x}_n$ are also calculated and they are $\underset{\sim}{x}_1 = \ldots = \underset{\sim}{x}_5 = (0,0), \underset{\sim}{x}_6 = (\pm.11, \pm.11), \underset{\sim}{x}_7 = (\pm.27, \pm.27), \underset{\sim}{x}_{10} = (\pm.41, \pm.41), \underset{\sim}{x}_{20} = (\pm.58, \pm.58)$, and $\underset{\sim}{x}_{30} = (\pm.65, \pm.65)$. From these observations, we believe that when $n \leq 5$, the origin will be the unique Bayes solution, and when $n \geq 6$, there will be four Bayes solutions, one in each quadrant, and they will move towards the four corner points of S as n goes to infinity. If this statement is true, we again get a threshold result for a two dimensional market. The threshold value will be 6, which is larger than the threshold value for the one-dimensional market which was 5. One possible explanation for this larger threshold value is that, due to an increase in the number of dimensions, the competitors themselves will spread out, making the central location relatively safer for the first vendor.

The next example is similar to Example 3 except that the shape of the market is now circular.

Example 4 Let S be the unit sphere in R^2. Suppose again that both of F and G are uniform distributions in S. Then, after some calculations, one has $\underset{\sim}{x}_1 = \ldots = \underset{\sim}{x}_5 = (0,0)$, $\underset{\sim}{x}_6 = .20(\cos\theta, \sin\theta)$, $\underset{\sim}{x}_{10} = .48(\cos\theta, \sin\theta)$, $\underset{\sim}{x}_{20} = .64(\cos\theta, \sin\theta)$, and $\underset{\sim}{x}_{30} = .71(\cos\theta, \sin\theta)$, where θ is arbitrary. We can see that the threshold value is 6, the same as the threshold value for a rectangular market. It seems that the threshold value does not depend significantly on the shape of S.

Now, let us look at an example where the customers prefer to stay near the center and yet the first vendor's Bayes solutions move towards the boundary for large n.

Example 5 Take S to be the unit sphere in R^2. Suppose the position of a customer has the density $f(\underset{\sim}{z}) = \frac{10}{9\pi}(1 - \frac{\|\underset{\sim}{z}\|^2}{5})$, and the future rivals are uniformly located in S, i.e., G

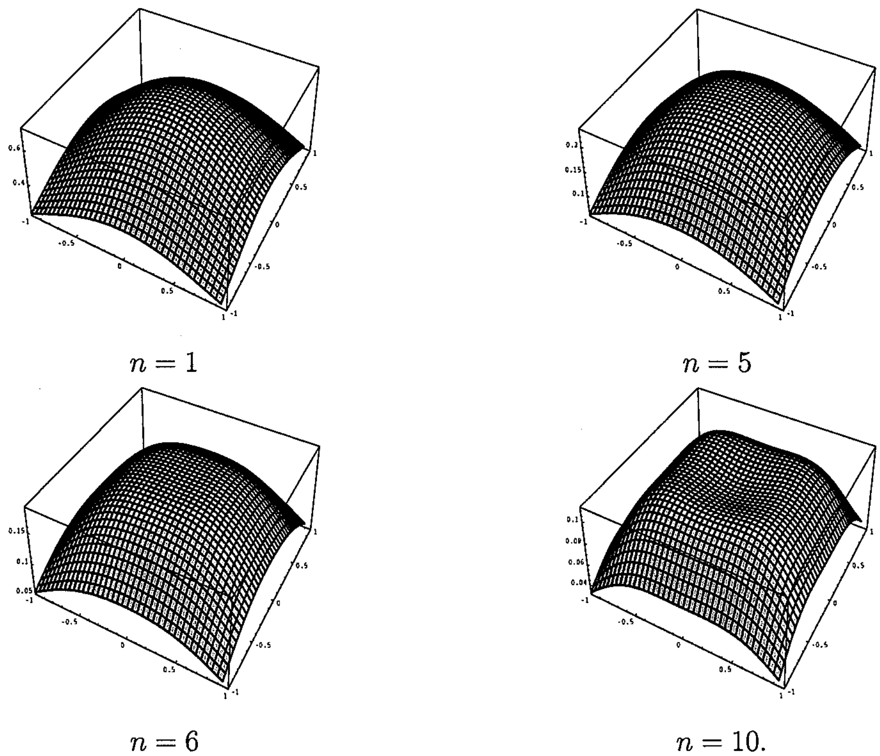


Figure 3: Plots of $D_n(x)$ when F and G are both uniform distributions on $[-1,1]\times[-1,1]$.

is still uniform. Then we have, for example, $\underset{\sim}{x}_1 = \ldots = \underset{\sim}{x}_6 = (0,0), \underset{\sim}{x}_7 = .23(\cos\theta, \sin\theta)$, and $\underset{\sim}{x}_{10} = .41(\cos\theta, \sin\theta)$, where θ is arbitrary. We now see that the threshold value seems to be $n = 7$. Also note that $\underset{\sim}{0}$ is the mode of f with value $f(\underset{\sim}{0}) = \frac{10}{9\pi}$, and f takes value 0 on all boundary points.

The general result given next says that an accumulation point of any sequence of Bayes solutions must be either an interior mode of f or a boundary point of S, as was the case for one dimension.

Theorem 6 *Let S be a bounded region in R^2 such that for each point $\underset{\sim}{x}$ in S, there exists a sequence of interior points of S convergent to $\underset{\sim}{x}$. Suppose that f is continuous and G is the uniform distribution in S. Let $\{\underset{\sim}{x}_n\}_{n=1}^{\infty}$ be any sequence of Bayes solutions for the first vendor. Then any accumulation point of $\{\underset{\sim}{x}_n\}_{n=1}^{\infty}$ is either an interior mode of f or a boundary point of S.*

Proof: Refer to Tsai et al. (1999). □

6 Summary

We have studied the Hotelling Beach model for spatial competition in one- and two-dimensional markets in the spirit of a statistical decision problem. We ask how the first chooser of a location for his store should act keeping future rivals in mind. We show that general results are possible under reasonable assumptions, typically symmetry and unimodality.

The number of future rivals is seen to play a very decisive role. In particular, we show that even if customers prefer to stay near a central location, the first chooser may prefer to be located near the boundary if there are a large number of future rivals. In particular, this eventual affinity to the boundary is true in both the dynamic and the nondynamic setup, and in the nondynamic case, we can pin things down via asymptotic expansions for the optimal location. We also show that when $n = 1$, the minimax and the dynamic Bayesian formulation lead to the same solution under simple conditions.

References

Donoho, D. (1982). Breakdown properties of multivariate location estimators. Ph.D. Qualifying Paper, Department of Statistics, Harvard University.

Hotelling, H. (1929). Stability in competition. *Econom. J.*, 39:41–57.

Small, C. (1990). A survey of multidimensional medians. *International Statistical Review*, 58:263–277.

Steele, J. and Zidek, J. (1980). Optimal strategies for second guessers. *J. Am. Statist. Assoc.*, 75:596–601.

Tsai, W.-C., DasGupta, A., and Zidek, J. V. (1999). Decision theory for spatial competition. Technical Report 99-22, Department of Statistics, Purdue University.

Tukey, J. (1975). Mathematics and the picturing of data. In *Proc. International Congress of Mathematicians*, volume 2, pages 523–31, Vancouver.

Corresponding author: Anirban DasGupta

Department of Statistics

Purdue University

West Lafayette, IN 47907-1399

dasgupta@stat.purdue.edu