BSc Thesis Applied Mathematics

Equilibria in probabilistic location models

Wouter Kregting

Supervisor: Alexander Skopalik

June, 2020

Department of Applied Mathematics
Faculty of Electrical Engineering,
Mathematics and Computer Science

UNIVERSITY OF TWENTE.

Preface

Equilibria in probabilistic location models

Wouter Kregting

June, 2020

Abstract

Location models are usually modelled assuming deterministic client behaviour. Assuming a probabilistic client behaviour changes the actions of the facilities. The exact nature of this probabilistic behaviour determines the extent to which the principle of minimum differentiation is followed.
Keywords: Location modelling, Probabilistic choice function, Hotelling

1 Introduction

Determining the location of a facility can be an important decision for a company. As distance to a facility can be an important factor for the decision of the client, the chosen location can have an impact on the profits of a company. In real life we can see that companies tend to locate their facility close together. This phenomenon is known as the *principle of minimum differentiation* [4], which states that in economy it it beneficial to differ as little as possible from your competitors.

1.1 Related work

Location modelling was introduced by Hotelling [6], and was extended by Downs [3]. In this model they model infinitely many clients evenly spread out on a line and a finite amount of companies that want to open a facility on this line. They choose the location strategically, depending on the other facilities, in order to maximise sales. There have been many different variations on this model. Models where price is also a variable have been considered [1]. In this case Carruthers shows that the outer facilities have an unique position, as the can charge a higher price. There have also been models deviating from the linear model. One example is a circular model [8, 7]. In this case the clients are equally spaced on a circle and the facilities choose their location on the circle. Furthermore there have been papers against Hotelling [2], showing doubt towards the *principle of minimum differentiation.*

1.2 My contribution

The papers mentioned above assume a deterministic behaviour of the clients. In this paper we look at location models where the clients make their choice probabilistically. The model used in this paper is based on the model from the paper: "From Hotelling to Load Balancing: Approximation and the Principle of Minimum Differentiation" [5]. In this model the companies choose the location of their facilities to maximise their load, i.e. attract as much clients as possible. The model assumes clients only base the decision of which facility to visit on a cost function, which is dependent on two variables: distance to

a facility and load of that facility. Each client chooses the facility deterministically, where the client will go to the facility with the lowest costs.
The model used here is a variation on the model mentioned above. Instead of the clients choosing the facility deterministically, they choose the facility probabilistically. They do this according to a probabilistic choice function. Different probabilistic choice functions will be considered, to see what changes affect the behaviour of the facilities. We will look for equilibria in this model. An equilibrium is defined as a location vector for the facilities where none of the facilities can change their location to increase their load. In this case all facilities are satisfied with their location, and do not want to move.

The simulations show that in some cases equilibria can be found. However, in other cases it fails to find an equilibrium. In these cases this is likely due to limitations in the simulation and not necessarily due to the absence of an equilibrium. In the cases that an equilibrium can be determined it can be seen that it depends on the probabilistic choice function whether or not the *principle of minimum differentiation* is followed.

2 Model

The model has a set of n facilities $\mathcal{N} = \{1, \ldots, n\}$, which each choose a location on the interval $S = [0, 1]$. The strategy vector of the facilities will be denoted as $s = (s_1, \ldots, s_n)$, where $s_i \in S$ denotes the chosen location of facility i. In the model we consider a finite set of clients P spaced uniformly over the interval $[0, 1]$. Client j will be at the position $z_j = \frac{j-1}{P-1}$. The load at each facility can now be defined as

$$\ell_i = \frac{1}{P} \sum_{j=1}^{P} p_{i,j}, \tag{1}$$

where $p_{i,j}$ is the probability that client j goes to facility i. This probability will be based on the costs of going to each facility. These costs are

$$\mathcal{C}_{i,j} = (1 - \alpha) \cdot |s_i - z_j| + \alpha \cdot \ell_i, \quad \alpha \in [0, 1]. \tag{2}$$

These costs are then used to calculate the probabilities of going to each facility. This is done by the following equation:

$$\begin{aligned} p_{i,j} &= B \cdot \frac{1}{\mathcal{C}_{i,j}}, \text{ with} \\ B &= \frac{1}{\sum_i \frac{1}{\mathcal{C}_{i,j}}}. \end{aligned} \tag{3}$$

Here B is a multiplication factor to ensure the sum of all probabilities of one client is equal to 1. The term $\frac{1}{\mathcal{C}_{i,j}}$ gives an inverse linear relation between the costs and probabilities. Higher order relations (quadratic, cubic, etc.) between costs and probabilities can also easily be modelled, by changing $\frac{1}{\mathcal{C}_{i,j}}$ with $\frac{1}{(\mathcal{C}_{i,j})^2}$ or $\frac{1}{(\mathcal{C}_{i,j})^3}$ in the probabilistic choice function. Finally an exponential relation between costs and probabilities will also be considered. In this case $\frac{1}{\mathcal{C}_{i,j}}$ will be replaced by $e^{-\mathcal{C}_{i,j}}$

It is possible that for some client j there are some facilities $i_1, \ldots, i_m$ s.t. $\mathcal{C}_{i_k,j} = 0$ with $1 \leq k \leq m$. In this case Equation (3) is undefined. Therefore if this is the case we take $p_{i_k,j} = \frac{1}{m}$ and $p_{i,j} = 0$ for $i \neq i_k$.

Each facility wants as much clients as possible, so they will choose their position to maximise their load.

3 Simulation

The model is simulated in MATLAB.
The equilibria are searched for via a best response dynamic. At the start the facilities are equally spaced on the interval $[0, 1]$. The i-th facility starts at location $\frac{i}{n} - \frac{1}{2n}$. Then for each facility it is checked what their best location would be, assuming the other facilities stay in their current location. The location of each facility is updated and this process is repeated until a strategy vector is reached which has been reached before.

To know which location is optimal for each facility the system of equation described by Equations (1), (2) and (3) must be solved to obtain values for the loads. These equations form a system of equations consisting of $2nP + n$ equations and equally as much variables. Due to Equations (3) this is not a system of equations consisting of only linear equations. I have tried to solve this using the "Solve" function in *Mathematica*. This is possible for really small n and P, but even for $n = 3$ and $P = 5$ this already takes a couple of minutes. This time increases drastically as n and P increase. Because this system of equation needs to be solved multiple times in one simulation and values of $n = 10$ and $P = 1000$ will be used it is not reasonable to do this exactly, as this would take too much time. To still be able to use these equations in the model the variables need to be approximated. This will be done by starting with a guess for the load, and then iteratively improving this guess until some desired accuracy is reached.

The first guess for the load will be $\frac{1}{n}$ for each facility. Using this guess the costs can be calculated using Equation (2). Then using Equation (3) the corresponding probabilities can be calculated. Finally the probabilities can be used to calculate new values for the loads. These new values will be the new guess for the loads. This process is repeated until a predetermined accuracy is reached, i.e. the new guess for the load differs from the old guess less than the predetermined accuracy. This approximation for the load is then used for the rest of the model.

The facilities are modelled on a finite grid. In each step of the model each facility chooses their best location somewhere on this grid. The fineness of this grid will have an influence on the accuracy of the model. This grid also ensures that the MATLAB code will halt, as the code will halt if a strategy vector is reached for the second time. Due to this grid being finite, there is also a finite amount of possible strategy vectors. However, when the code halts, it does not mean that an equilibrium has been found. The final strategy vector is only an equilibrium if the strategy vector has remained unchanged. There is also a possibility that the last strategy vector is the same as the strategy vector some steps earlier. Then there is some cycle of size 2 or larger the model goes through. If this is the case we cannot make a statement about the existence of an equilibrium, only that this method of searching has not yielded one.

4 Results

The model allows for different levels of accuracy by changing the amount of clients, the accuracy of the load and the fineness of the grid for the facilities. In all the following calculations 1000 clients, an accuracy of 10^{-5} for the load approximation and a grid fineness of 10^{-3} have been used. If calculations have been done for multiple α, then a step size of 0.1 is used.

4.1 Linear model

The first thing to consider is the linear case. In this case there is an inverse linear relation between costs and probabilities. With $\alpha = 0$ the model is completely based on distance. When the model is completely based on distance equilibria can be found, as can be seen in Figure 1. This is due to the probabilistic choice function. In the deterministic model a facility would lose a client if there was an other facility closer to that client. In the probabilistic model however, the client always has a probability to go to each facility, regardless of their location. If the facility tries to increase the probability of a certain client coming to them, they can do this by, for example, moving to the left. This however, has the consequence that clients to the right of the facility will have a lower probability of coming to them. Therefore it is not very beneficial for a facility to make large changes in their location. This has the result that there is not a clustering behaviour when $\alpha = 0$.

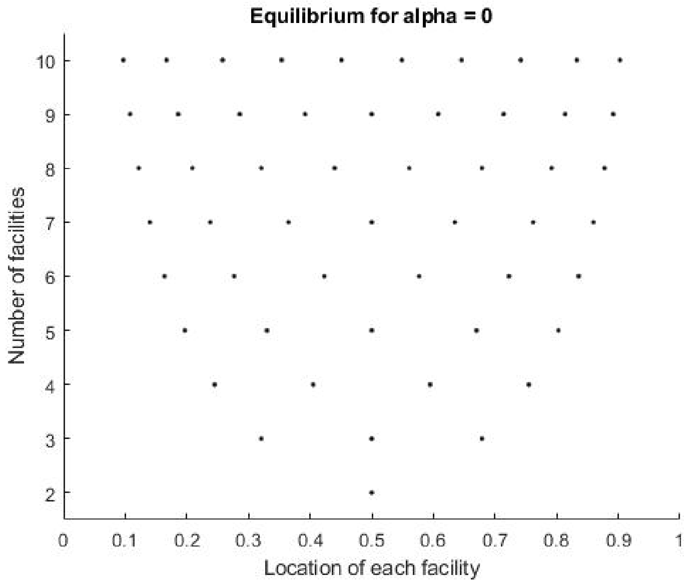
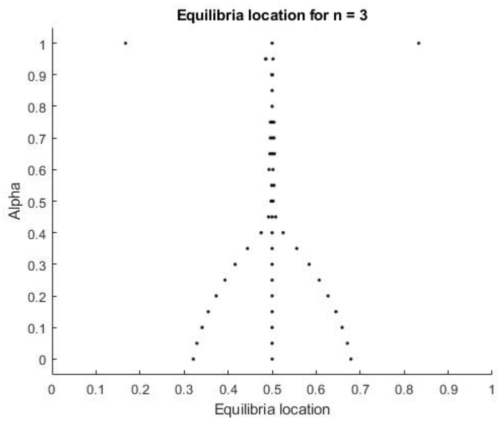


FIGURE 1: Equilibria of the linear model for different n and $\alpha = 0$

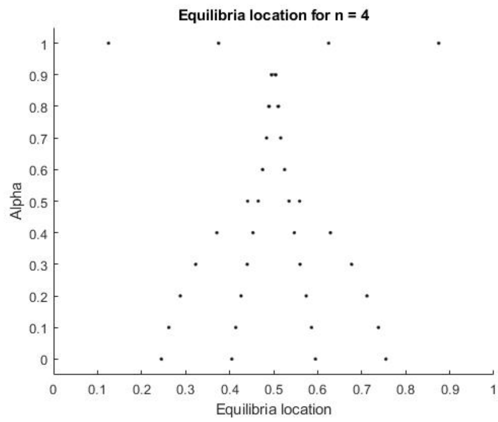
As can be seen in the figure above it looks like there is some pattern in the equilibria for $\alpha = 0$. For example when looking at the leftmost facility it appears that their location in the equilibrium lies close to $\frac{1}{n}$. In the table below the location of the leftmost facility is compared to the value $\frac{1}{n}$. There is still a small difference between the values, but this could be the result of the approximations and limits in accuracy of this model. Therefore I suspect that the equilibrium location of the leftmost facility in the linear model with $\alpha = 0$ is equal to $\frac{1}{n}$.

n	2	3	4	5	6	7	8	9	10
Location of leftmost facility	0.500	0.321	0.245	0.197	0.164	0.140	0.122	0.108	0.097
$1/n$	0.500	0.333	0.250	0.200	0.167	0.143	0.125	0.111	0.100

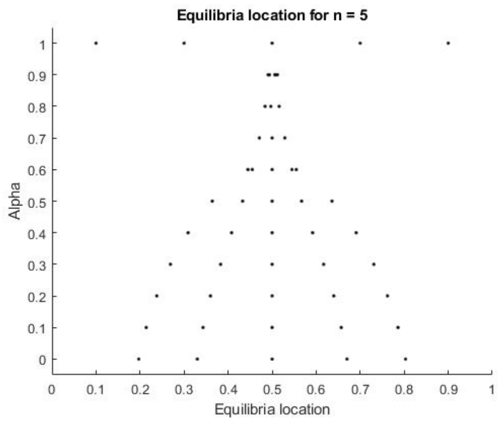
As α increases the probabilistic choice function becomes more dependent on load and less on distance. When $\alpha = 1$ the model is completely dependent on load. In this case the load of every facility is equal to $\frac{1}{n}$, irregardless of the location of the facilities. Therefore every every strategy vector is an equilibrium. What is important to investigate is what happens when α increases from 0 to 1. The model shows, as can be seen in Figure 2, that the facilities will show a clustering behaviour as α increases.



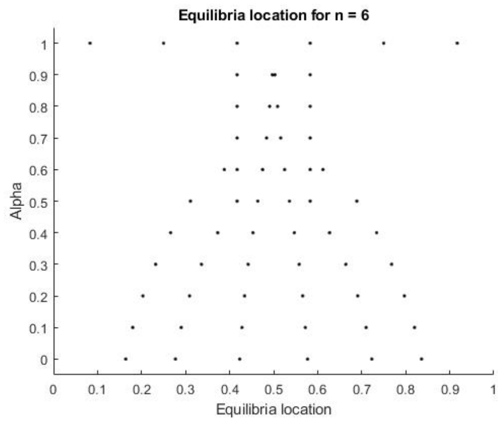
(A) Equilibria locations for $n = 3$



(B) Equilibria locations for $n = 4$



(C) Equilibria locations for $n = 5$



(D) Equilibria locations for $n = 6$

FIGURE 2: Equilibria location for different α and different n

The figures above show a clear clustering behaviour as α increases. This indicates that, as the distance becomes a less important factor for the clients, the facilities have less need to distance themselves from their competitors. In Figure 2a the three facilities seem to cluster closely together at $\alpha \geq 0.5$. The model for $n = 4$ also shows that for the same α values the two leftmost facilities cluster together as well as the two rightmost. However these two groups do keep some distance from each other and converge towards each other as $\alpha \to 1$. Similar behaviour can be seen in Figure 2c, where first the outer facilities cluster together and as $\alpha \to 1$ all facilities cluster together in the middle. Figure 2d shows a different behaviour. The outer facilities do still cluster together, but once they are clustered together they stay in the same location. For high α the two rightmost and the two leftmost facilities cluster together at 0.417 and 0.583 respectively. The middle two facilities still cluster together in the centre as $\alpha \to 1$.

4.2 Quadratic model

The probabilistic choice function can be adapted to give an inverse quadratic relation between the costs and the probabilities. This can be done by changing Equation (3) to

$$p_{i,j} = B \cdot \frac{1}{(\mathcal{C}_{i,j})^2}, \text{ with} \qquad (4)$$
$$B = \frac{1}{\sum_i \frac{1}{(\mathcal{C}_{i,j})^2}}.$$

In Figure 3 the equilibria for different amount of facilities can be seen for $\alpha = 0$. If you compare this figure with Figure 1 of the linear case, then it seems to be that there is little to no difference between the equilibria.

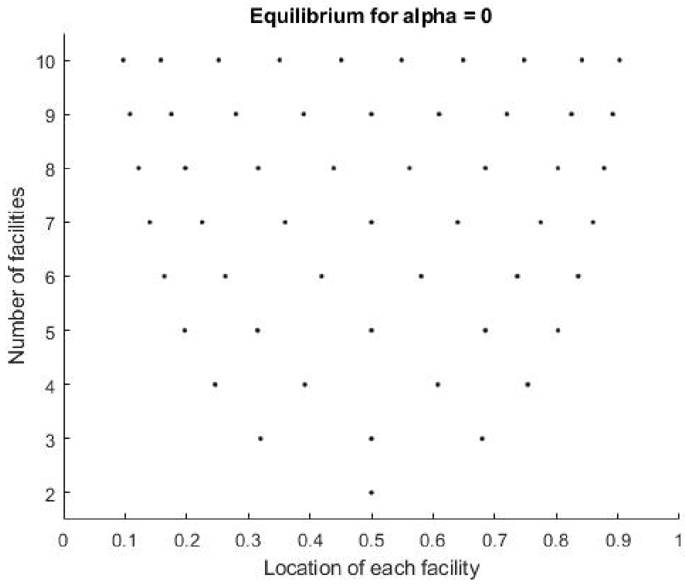


FIGURE 3: Equilibria of the quadratic model for different n and $\alpha = 0$

If we for example take a look at the model with 4 facilities we can indeed see that the difference between the linear and quadratic model is small.

Linear model	0.246	0.392	0.608	0.754
Quadratic model	0.245	0.405	0.595	0.755
Difference	0.001	0.013	0.013	0.001

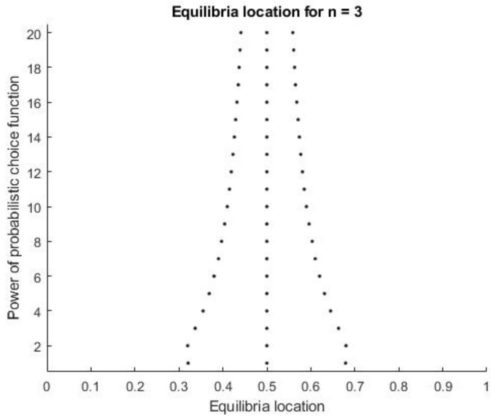
In the linear model a difference of factor 2 between two different costs would result in the respective probabilities differing by the same factor. In the quadratic model this difference in the probabilities is squared, which results in larger differences in probabilities. However comparing Figure 1 with 3 and the table above show that this difference does not change the equilibrium location by much. It is then interesting to look at the behaviour of the equilibria as higher order probability choice functions are used.

It is also interesting to see what happens for different α. Similar as in the linear model, when $\alpha = 1$ the costs do no longer depend on distance, which means that location of each facility in unimportant and every strategy vector is an equilibrium. It is then interesting to see what happens when as $\alpha \to 1$. This however comes with a problem. For $n = 3$, when $\alpha \geq 0.4$ the approximation for the load, as described in Section 3, does not converge to one load vector. Instead it starts to cycle between 2 or more different load vectors. In this case the approximation for the load cannot be used and therefore an equilibrium cannot be determined. For $\alpha < 0.4$ the existence of an equilibrium can still be determined. The model shows however that only for $\alpha \leq 0.2$ an equilibrium exists. For $\alpha = 0.3$ the model starts to cycle between two strategy vectors. In this case the model cannot determine the existence of an equilibrium, as the strategy vector keeps updating. So either there does not exist an equilibrium, or this method is insufficient to find an equilibrium in this model.

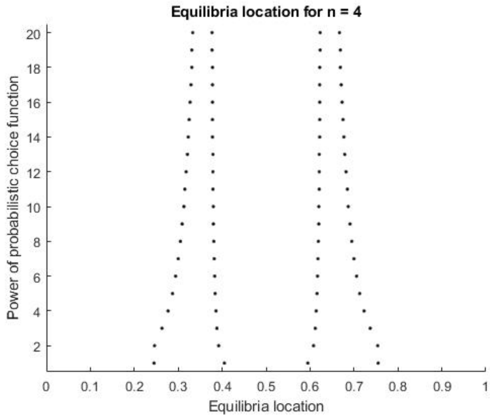
It is possible to search for equilibria in different ways. One way to do this is is by exhaustive search. This method checks all possible strategy vectors. The advantage of this is that, if there is an equilibrium, it will be found. The disadvantage is that this method is extremely time consuming, compared to the best response dynamic. For this reason this method has not been used. Because for $\alpha = 0.3$ no equilibrium has been found, this method of exhaustive search can be used to examine if there really is an absence of an equilibrium. This search will be done with a grid fineness of 0.01, in stead of 0.001 with the best response dynamic, to maintain an acceptable computation time. This does however decrease the accuracy of the model. Due to time constraints this search method was only used for $n = 3$ and $\alpha = 0.3$. This method did find an equilibrium in the location $0.3300, 0.5000$ and 0.6700 for the three facilities. This shows that the best response dynamic is not flawless, as it can miss equilibria. Although this does not guarantee that for higher order probability choice functions, or higher α, equilibria will always exist, it does show that it is reasonable to suspect that those equilibria exist.

4.3 Higher order probability choice functions

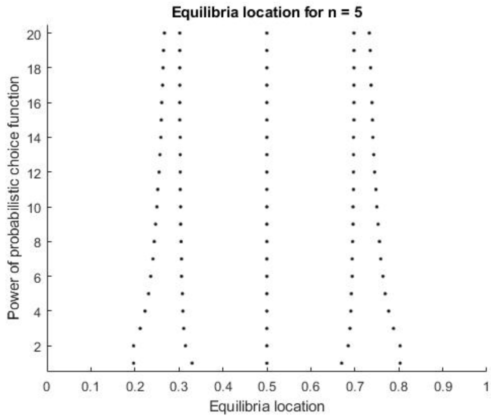
Similar as with the linear and quadratic model, higher order probability choice functions can also be used. As the power of the probabilistic choice function increases a small difference in costs will cause an increasingly larger difference in probabilities. This means that the clients are more likely to go to one single facility as the power increases. In the limit this will mean that the model will converge to a deterministic model where each client goes to the facility with the lowest costs. In the figures below the behaviour of the equilibria as the power of the model increases can be seen. In this case $\alpha = 0$ is used.



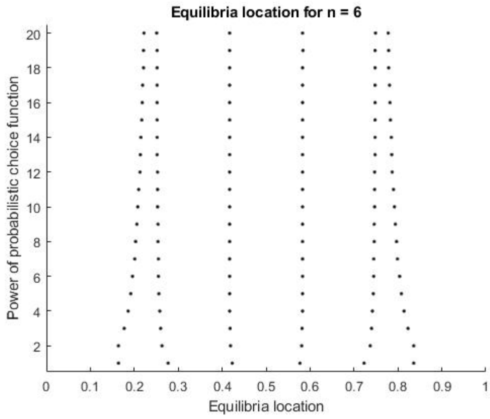
(A) Equilibria locations for $n = 3$



(B) Equilibria locations for $n = 4$



(C) Equilibria locations for $n = 5$



(D) Equilibria locations for $n = 6$

FIGURE 4: Equilibria location for different powers in the probabilistic choice function for different n

Figure 4a shows that for $n = 3$ the facilities start to cluster together in the centre of the interval as the power increases. For more facilities a different behaviour can be seen. The two leftmost facilities are clustering together, as well as the two rightmost facilities. For $n > 4$ the facilities in between these two clusters do not converge to each other or to one of the clusters. They stay evenly spread out between the clusters. This shows that, except for $n = 3$, as the power increases the facilities do not all cluster together towards one spot, but keep some distance to some of the other facilities.

Similar as with the quadratic probability choice function an equilibrium cannot be determined as α increases. For a small increase α, the model starts to cycle between two or more strategy vectors, and for large α the load approximation fails. To determine the existence of equilibria in these conditions other search methods need to be used.

4.4 Exponential model

Just like with the higher order probability choice functions the function can also be adapted to show an exponential relation between costs and probability. In this case Equation (3) can be rewritten as

$$p_{i,j} = B \cdot e^{-\mathcal{C}_{i,j}}, \text{ with} \qquad (5)$$
$$B = \frac{1}{\sum_i e^{-\mathcal{C}_{i,j}}}.$$

This however shows a different behaviour than the polynomial model, as can be seen in the figure below.

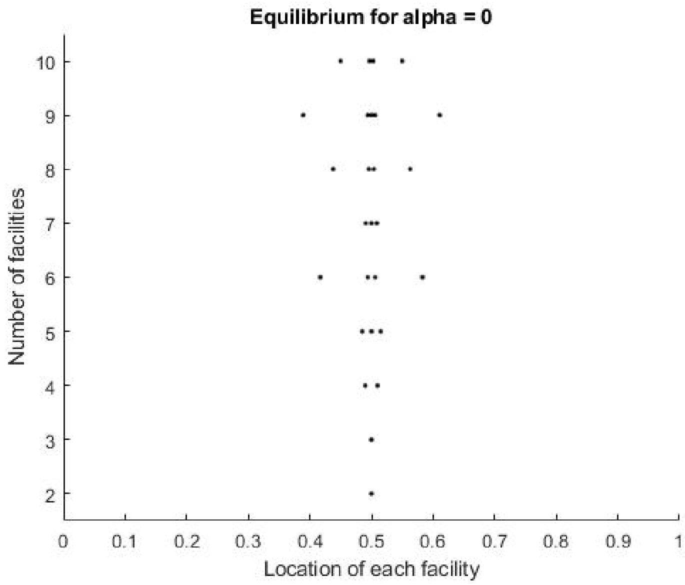


FIGURE 5: Equilibria of the exponential model for different n and $\alpha = 0$

Unlike the other functions the exponential function causes the facilities to cluster. This different behaviour can be explained by the costs being smaller than 1. In this model the costs are always a value in the interval $[0, 1]$. In the linear model this means that $\frac{1}{\mathcal{C}_{i,j}}$ has a value on the interval $[1, \infty]$. This has the result that the probabilities can differ from each other by any factor. This is the same in any higher order model. In the exponential model however this is different. Because the costs lie in the interval $[0, 1]$, the value $e^{-\mathcal{C}_{i,j}}$ lies in the interval $[\frac{1}{e}, 1]$. This has the result that the probabilities can at most differ by a factor of e from each other. This smaller difference between the different probabilities is most likely the cause of the clustering behaviour in the model.

In Figure 5 a surprising behaviour of the outer facilities can be seen, as they do not always cluster together with the rest of the facilities. I do not have a solid explanation for this behaviour and can only guess that this is the result of inaccuracies in the model and calculations.

5 Conclusion

The model shows different behaviour, depending on which variables are changed. For both a low power of the probabilistic choice function and a low α the model shows facilities prefer to distance themselves from other facilities. This goes against the *principle of minimum differentiation.* When distance becomes a less important factor, clustering behaviour can be seen. When increasing the power of the probabilistic choice function some clustering behaviour can also be seen, but this is not similar to the linear model with large α. Only the outer facilities cluster together. Apart from that the facilities still distance themselves from each other. With an exponential probabilistic choice function clustering behaviour can again be seen. This shows that in a model with a probabilistic choice function the exact parameters of this function highly influence how much the *principle of minimum differentiation* is met.

The existence of equilibria could not be determined in models with higher order probabilistic choice function and high α, using best response dynamic. An exhaustive search showed that in one case an equilibrium sill exists. This shows that, although best response dynamic may not always find an equilibrium, it does not guarantee the absence of an equilibrium. Determining the existence of equilibria must then be done via different methods.

References

[1] Norman Carruthers. Location choice when price is also a decision variable. *The Annals of Regional Science*, 15(1):29–42, 3 1981.

[2] C. d'Aspremont, J. Jaskold Gabszewicz, and J.-F. Thisse. On Hotelling's "Stability in Competition". *Econometrica*, 47(5):1145, 9 1979.

[3] Anthony Downs. An Economic Theory of Political Action in a Democracy. *Journal of Political Economy*, 65(2):135–150, 4 1957.

[4] B. Curtis Eaton and Richard G. Lipsey. The Principle of Minimum Differentiation Reconsidered: Some New Developments in the Theory of Spatial Competition. *The Review of Economic Studies*, 42(1):27, 1 1975.

[5] Matthias Feldotto, Louise Molitor, Pascal Lenzner, and Alexander Skopalik. From hotelling to load balancing: Approximation and the principle of minimum differentiation. In *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS*, volume 4, pages 1949–1951. International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), 2019.

[6] Harold Hotelling. Stability in Competition. *The Economic Journal*, 39(153):41, 3 1929.

[7] Amoz Kats. More on Hotelling's stability in competition. *International Journal of Industrial Organization*, 13(1):89–93, 3 1995.

[8] Steven C. Salop. Monopolistic Competition with Outside Goods. *The Bell Journal of Economics*, 10(1):141, 1979.