Agent-Based Spatial Competition

Mark Lijesen and Eveline van Leeuwen
VU University Amsterdam

ABSTRACT

Equilibrium in the Hotelling model of spatial competition is guaranteed if the distribution of consumers is log concave. In the real world, nothing guarantees such a log concave distribution however, rendering the analytical model unable to provide a primer as to what one might expect from empirical applications. We develop an agent-based model of spatial competition that is capable of reproducing the results of the analytical model and also provides meaningful results for some cases where the distribution of consumers is not log concave. Using numerous simulations, on randomly drawn distributions, we derive equilibrium locations and prices and test for uniqueness. Moreover, we check whether the relationships between characteristics of the distribution (e.g. concentration, skewness) and outcomes are consistent with the analytical model.

1. INTRODUCTION

Hotelling's metaphor of spatial competition on Main Street has become one of the most important models in understanding strategic product differentiation. A vast literature has evolved, discussing extensions of the model with respect to the nature of transport costs, distributions of consumers (or preferences), dimensions and so on.

One of the main reasons for the model's popularity lies undoubtedly in its versatility. The model's assumptions can easily be changed, while still allowing for analytical solutions. Finding analytical solutions has its limits however. As Caplin and Nalebuff (1991) show, the existence of a price equilibrium in pure strategies in the Hotelling model can only be guaranteed if the density function of the underlying distribution of consumers is log-concave. This condition is hardly ever met in practice. Moreover, the fact that an equilibrium can't be guaranteed, does not mean it does not exist.

The popularity of Hotelling's model has only recently been translated into empirical applications (see Lijesen, 2010 for a recent overview). Empirical applications relate to situations where many of the simplifying assumptions used in the Hotelling model do not hold. Therefore, authors often relate their findings to generalizations of the theoretical models. The availability of reliable primers for outcomes in not-so-stylized situations would allow for more explicit testing of the Hotelling model.

In this paper, we develop an agent-based version of the Hotelling model, capable of reproducing the results from the analytical model. Moreover, the model can be used to search for solutions, even if they are not guaranteed. Obviously, a solution will not always be found, but first results look promising.

The remainder of the paper is organized as follows. Section 2 describes the model structure, followed by a discussion of the concepts and problems of discrete consumers in section 3. In section 4, we analyze the results of a fairly large number of simulations using randomly drawn distributions. Section 5 zooms in on a subset of these equations, providing more detailed case studies and section 6 concludes.

2. MODEL STRUCTURE

Multi-stage games like the Hotelling game are usually solved analytically by backward induction. Our ag-nt based model is based on decision-rules that are consistent with the concept of Nash equilibrium. The aim of the model is to find the optimal location for a firm to be located, considering the distribution of consumers. The model has been programmed in a Netlogo environment (Wilenski, 1999).

There are two different types of agents, first of all the consumers, secondly the shops:[1]

- The consumers choose a shop based on the distance to it and the price of the product. The consumers minimize their costs, which are calculated by the sum of the distance squared and the price. All consumers buy in total 1 product.
- The shops maximize their profit. The profit is calculated by multiplying the price set by the shop itself by the number of consumers that choose that particular shop.

The model has three main parts: the shop choice of the consumers, and the price optimization and location optimization of the shops (see Figure 1 for a flowchart).

After the initial setup, the consumers choose their favorite shop based on the distance to the shop and its price for the uniform product. The procedure 'shop choice' is run every time one of the shops changes something.

[1] In this model, both types of agents are turtles. Patches (referred to as locations in the remainder of the paper) are used only to calculate the reservation price and communicate this to the shops

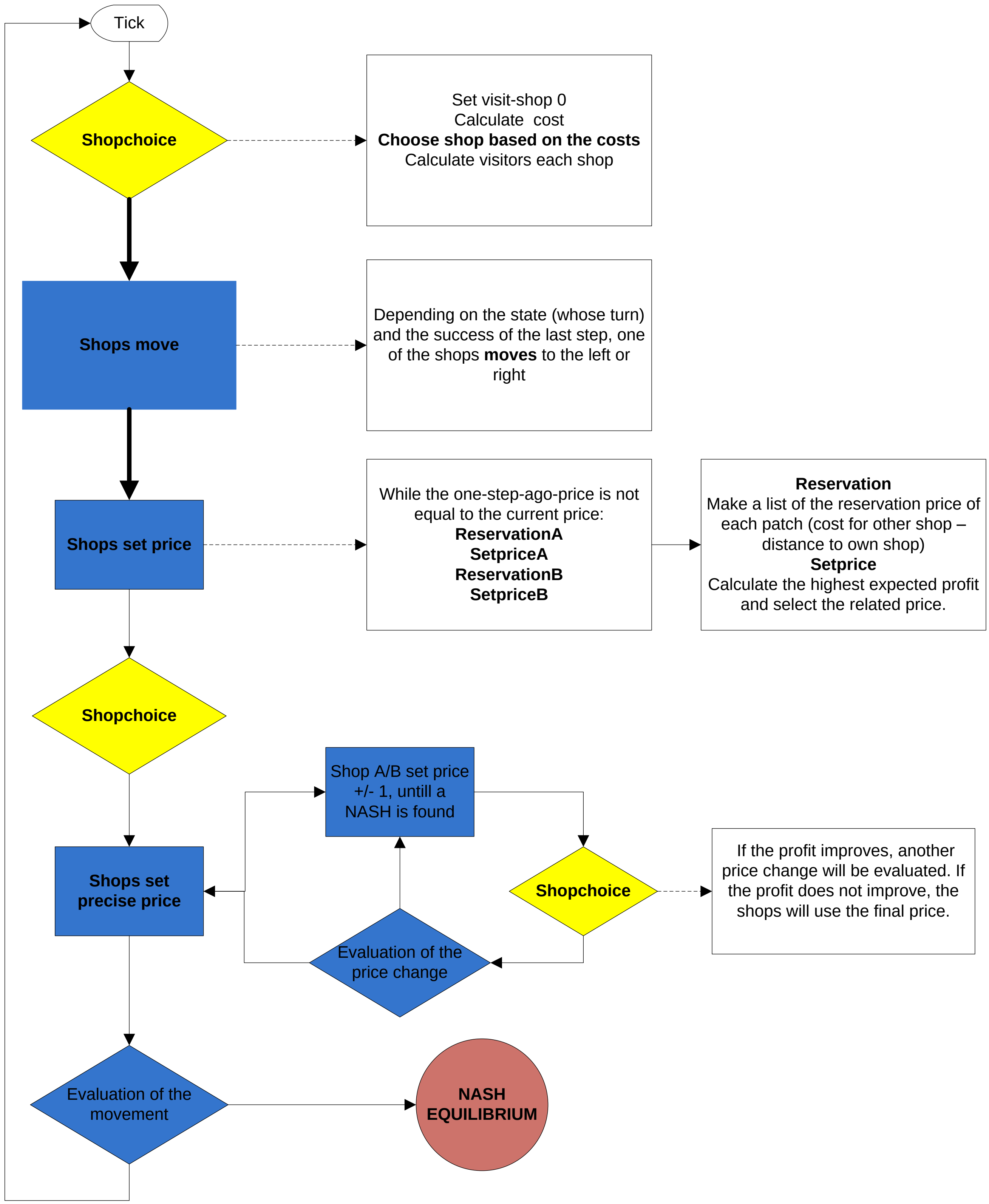

Figure 1. Structure of the model

After this, shop A increases its price, the consumers choose again, and shop A evaluates the effect on its profits. If the new profit is higher than the initial profit, the shop will keep on increasing the price until its profit does not increase anymore. However, if the new profit is not higher than the initial profit, shop A will decrease the price until its profit no longer increases. Once shop A cannot increase its profit any more, shop B will go through the process of increasing or decreasing its price to find the optimal profit in that situation. Both shops repeat the process until neither of them can increase their profits unilaterally and a Nash equilibrium in prices is reached. The model then moves to the location stage.

In the location stage, shop A will move one step to the right. With the new location, both shops will again change their prices until they reach a Nash equilibrium in prices. Then shop A will compare the (optimal) profit after moving with the old profit and if it increased it will keep on moving in the same direction until the profit does not increase anymore. If the profit after moving is not higher than the old profit, the shop will start moving in the other direction. Once shop A cannot improve its profit by moving, shop B will go through the same relocation procedure. If neither of the shops can increase their profits unilaterally, a Nash equilibrium in locations is reached and the equilibrium outcome is found.

3. DISCRETE CONSUMERS

Consumers in our model are, like in real life, discrete decision making units. This adds realism to the analytical model (that implicitly assumes consumers to be infinitely divisible), but it also introduces new problems.

The most important problem is that a small increase in price may either cause no change in the behavior of the marginal consumer or the extreme change that the marginal consumer switches shops. This implies that shops can increase their price up to the point where the marginal consumer just chooses that store. Since stores take turns in the pricing game, they leapfrog each other in prices, without losing the marginal consumer.

The solution to this problem is to reintroduce some divisibility in the model, by allowing the marginal consumer to spread her expenditure over both stores. To this end, we compute (a scaled version of) the exact location of the indifferent consumer, just as we would do in the analytical model:

$$\widehat{x} = \frac{p_B - p_A \;\; + x_B^2 - x_A^2}{2 \;\; x_B - x_A} \quad (1)$$

The expenditure of the consumer located at (the integer of) $\widehat{x}$ are divided between the shops accordingly. This way, we ensure that even small price increases lead to a decrease in demand, allowing the model to optimize the trade-off between them

One other issue with discrete consumers, is that solutions of the location stage are limited to discrete locations as well. This is especially problematic if the number of consumers (or locations) is small. To shed some light on this issue, we have performed simulations with the model while varying the length of Main Street.

At this point, we note that our model allows firms to locate outside the inhabited part of Main Street. Tabuchi and Thisse (1995) show that, if firms are allowed to locate outside Main Street (defined as locations from 0 to 1), they will locate at -0.25, 1.25 if preferences are distributed uniformly. We allow firms to locate between -0.5 and 1.5, implying that the total length of Main Street is twice the length of its inhabited part. We adjust the scale to accommodate discrete consumers. Apart from scaling, we have that The smallest size of Mains Street for which the model can be solved is 8 locations.

We solved the model for all lengths of Main Street from 8 to 400 locations (in steps of 4), using a uniform distribution of consumers (i.e. each location holds the same number of consumers). We then rescaled the outcomes and compared them to the outcome of the analytical model. Figure 2 below plots the relative deviations in outcomes against the total length of Main Street.

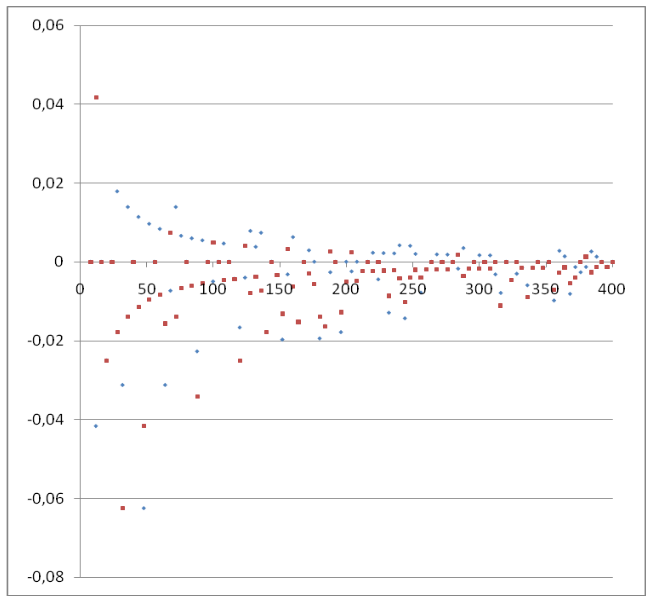

Figure 2. Deviations in outcomes versus total length of Main Street (uniform distribution of consumers).

The overall pattern in figure 2 clearly resembles that of errors due to rounding; errors (both positive and negative) are larger for small sizes and decrease rapidly with size. Additionally, we find some observations with relatively large negative deviations, that also decrease rapidly. These deviations follow from the discrete nature of the model as well, with shops being unable to make the final steps toward the predicted optimal location.

Overall, the model performs fairly well; the vast majority of deviations from the analytical model is smaller than 2 percent. In the remainder of the paper, we will use a version with a total length of Main Street of 80 locations, as this model predicts the analytical outcome for the uniform distribution perfectly and is not too large, so it keeps computation time limited.

4. SIMULATIONS USING RANDOMLY DISTRIBUTED PREFERENCES

This section discusses our analysis of location equilibria following from 10.000 simulations with the model described in the preceding sections. In these simulation, each location was assigned a random number of consumers, drawn from a Poisson distribution with λ=10. This results in a randomly generated distribution of locations, which is highly unlikely to be log concave. In fact, none of the 10.000 distributions generated this way was log concave. All simulations were run from one single starting position; (12,68). A more detailed analysis for a smaller number of cases is presented in section 5.

Using a single starting position was done to limit simulation time. It has three drawbacks however. First of all, by using a single starting position, we underestimate the number of runs that has a solution. Secondly, the simulations that solve are biased towards outcomes that are closer to the starting positions. Finally, the outcomes from the simulations may represent suboptimal equilibria. The first two drawbacks are unlikely to influence the results in the remainder of this section, but the third may, calling for some caution in drawing conclusions from our results.[2] We use four indicators to describe the distributions used in the simulations. Each of these indicators is discussed below.

Average density is defined as the average of the number of consumers located at all inhabited locations. Since the number of consumers at each individual location is drawn from a Poisson distribution with λ=10, we would expect this value to be close to 10 on average.

Weighted position is defined as the average position of all locations, weighted by their number of consumers (calculated as the productsum of the numbers of consumers and positions divided by the sum of densities). This indicator reflects whether consumers are situated relatively more to the right side (high value) or the left side (low value) of Main Street. For a uniform distribution (in fact, for any symmetric distribution), the value of this indicator would be 39.5.

The *Hirschmann Herfindahl Index* is a commonly used concentration measure, calculated by the sum of squared shares (of densities in this case). The higher the indicator, the larger the difference between peaks and valleys in the distribution. For a uniform distribution, the value of this indicator would be 250, which is also the theoretical minimum, given the fact that we have 40 inhabited locations.

Spread indicates how far consumers are from the centre, weighted by their number of consumers (calculated as the productsum of densities and distance to the centre divided by the sum of densities). The higher the value of this indicator, the further consumers are from the centre of Main Street. For a uniform distribution, the value of this indicator would be 10.

Out of the 10.000 runs, 442 resulted in a Nash equilibrium in prices and locations, for this starting location. That may seem like a low success rate, but one has to bear in mind that the distributions of preferences used are very whimsical and that we only used one starting location. As we will see in section 5, increasing the number of starting locations, increases the success rate too. Table 1 below provides descriptive statistics on the distribution measures discussed above.

Table 1. Descriptive statistics of distribution measures (n=442)

Indicator	Mean	St.dev	Min	Max
Average density	10.02	0.52	8.48	11.55
Weighted position	39.51	0.59	37.71	41.27
HHI	274	5.64	261	294
Spread	9.88	0.28	8.81	10.59

The overall image from table 1 suggests that the means are close to the expected values, but the sample provides sufficient variation for further analysis.

To get a basic idea of the validity of the results we find, we run two regressions. First, we regress the distance between equilibrium locations found in our simulations against the distribution measures presented above. Based on Neven (1986), one would expect that the measure 'spread' positively impacts the distance between equilibrium locations; if consumers locate more towards the fringes, firms will also move outward.

In the second regression, we relate the average position (or the midpoint between positions) against the same distribution measures. Again, based on Neven (1986), we would expect a positive impact from the measure 'Weighted position'; if consumers locate further to the right, firms will do so as well. Table 2 presents the results for both regressions, using OLS with robust standard errors.

The expected effects are clearly visible from both regressions, confirming Neven's theoretical result that firms follow consumer, a result that is also found in many empirical studies (Lijesen, 2010). The order of magnitude seems to be somewhat off though. Based on the analytical model, one would expect the parameter reflecting the impact of spread on the distance between firms to equal 6. Note however, that this expectation is based on a uniform distribution of consumers. Moreover, the parameter for distance is not significantly different from 6.

[2] Note to the referee: we are currently building a larger database of detailed simulation runs. and will update the analysis in this section in a later version

Table 2. Regression results (n=442)

	Distance		Average position	
	Coeff	**t-value[a]**	**Coeff**	**t-value[a]**
Constant	-3.54	-.014	34.06	5.84***
Average density	1.28	2.29**	-0.20	-1.58
Weighted position	0.06	0.13	0.25	2.22**
HHI	-0.08	-1.67*	-0.01	-1.03
Spread	6.96	6.49***	0.17	0.79
R^2		0.13		0.02
F		12.95***		2.22*

* Significant at the 10%-level
** Significant at the 5%-level
*** Significant at the 1%-level
[a] t-values are based on robust standard errors

The parameter for the impact of the weighted position of consumers on the average position of firms seems to be off a bit more. Irrespective of the type of distribution, one would expect that if every consumer moves one step to the right, both shops will move one step to the right. Therefore, one would expect the parameter to be close to unity, whereas we find a value of 0.25. The fact that some of our outcomes represent suboptimal locations may be an important cause for this bias, as we would expect our fixed starting position to dampen the effect and hence yield a lower value in the estimation. This may also explain why the fit of the 'average position' model is low. [3]

Some of the other indicators, which we merely entered to correct for unexpected influences, also seem to have a significant effect. Although one can think of economic reasons why the average density would impact the incentives for strategic product differentiation, we have to acknowledge that the mechanisms behind this are not part of our model and can therefore not explain why we find a statistically significant impact of the average density on the distance between two firms.

5. CASE STUDIES; UNIQUENESS AND STARTING POSITIONS

In section 4, the distributions that result in a solution for starting positions 12, 68 were analyzed. However, it could well be possible that amongst the simulated solutions, suboptimal locations can be found, or that those that did not result in a solution, could have solved when other starting positions had been used. Therefore, in this section, case-studies are shown of distributions for which we have run the model for all possible starting positions. By doing this, it should be possible to distinguish between optimal and suboptimal locations, and it should be possible to say something about what happens when different starting positions are chosen.

Based on a small number of runs simulating all possible combinations of starting points (1600 runs per distribution), we know that looking at only every second possible combination gives a good enough idea of possible locations and the related starting positions. Therefore, for 30 different distributions, we have run the model 400 times, for every second combination of starting points between 0-40 for shop A and for 40-80 for shop B.

From the 30 case studies we have run, 16 did not solve at all. This means that the model never (for none of the 400 starting positions) resulted in a (sub) optimal location for the shops. In general, this happened because no Nash equilibrium could be found. In a few cases because the shops wanted to move outside the world (beyond 0 or 80).

Figure 3 shows at the left side, for four different distributions, the (sub) optimal locations found for shop A and B, and to the right the starting positions that resulted in one of the locations. Starting positions within the same series have a joint solution.

When using an uniform distribution, from every starting position the same optimal location is found (see Figure 3a). This is the classical way of representing Hotelling's theory. However, when using a random distribution, the results are clearly different.

From the 14 case- studies that did result in a solution, it appears that there is a wide variety of the number of solutions and their locations, as well as a wide variety of starting positions that result in a solution.

In general, instead of one, several (sub) optimal locations can be found. In only one case, one optimal location was found, however, that could only be reached from one starting position (see Figure 3b). In four cases two solutions were found, which could be reached from many different starting positions (see Figure 3c).

Furthermore, not surprisingly, the case-studies show that starting points near a (sub) optimal location more often result in a solution. In addition, it also seems that starting points of between 80 and 60 for shop B relatively often result in a solution.

Unfortunately, it is not always possible to distinguish the optimal locations from the (sub) optimal ones. From the 14 cases, 9 result in a clear optimal location in which the profits for both shops are the highest. In five cases, no optimal location could be pointed out. In those five cases, often a relatively high number of sub optimal locations is found with large distances between each other, such as shown in Figure 3d.

[3] Note to the referee: we are currently building a larger database of detailed simulation runs, which will allow us to run and present regressions based on global optima alone.

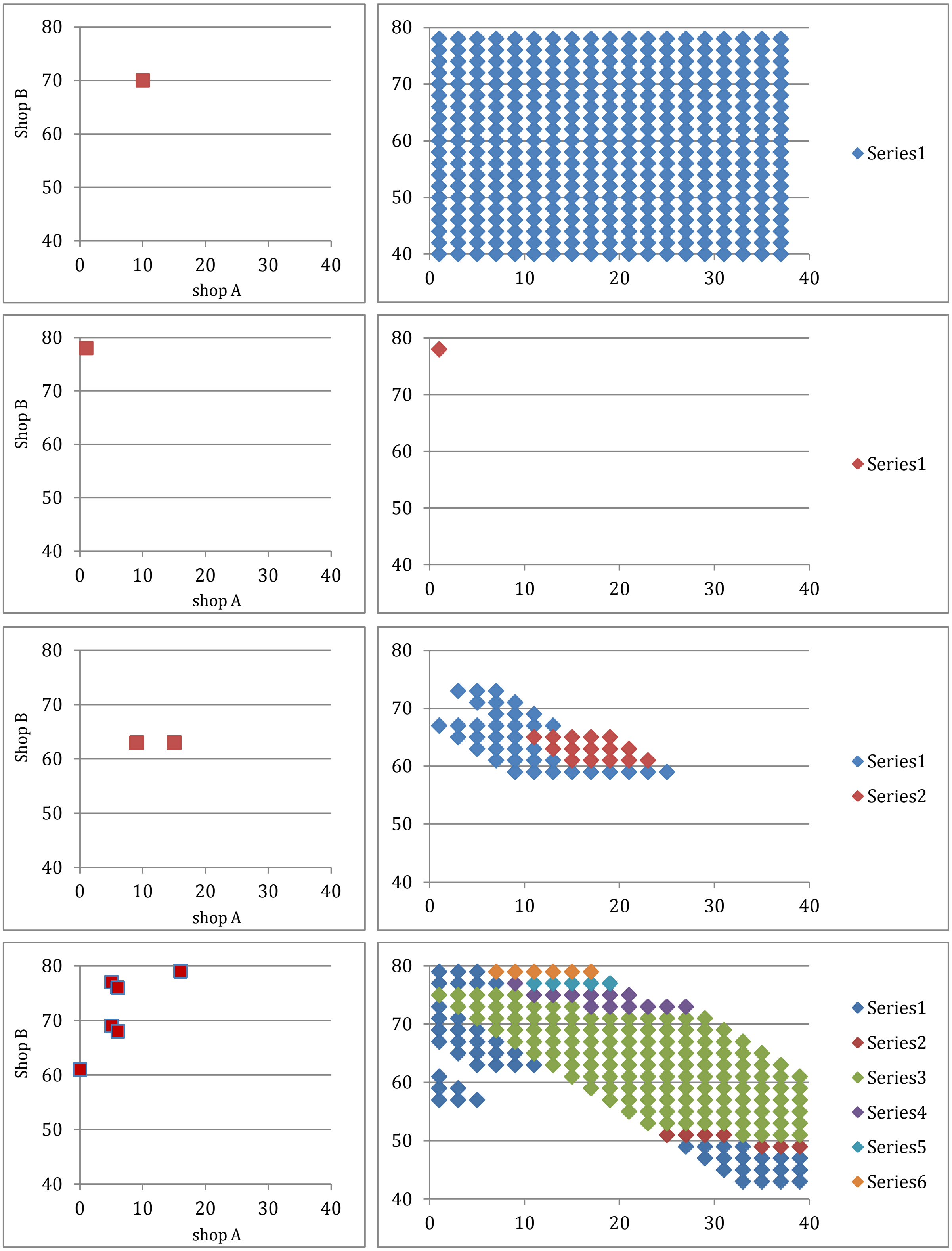

Figure 3 a-d: the combinations of locations found (to the left), while using certain starting positions (shown to the right).

Finally, a diagonal pattern can be distinguished in many of the case-studies. This could point out that symmetric starting positions have a higher probability to result in a solution, even with non- symmetric, random distributions of consumers.

These 30 case-studies have given us a hint of the wide variety of solutions and the importance of checking more starting positions. However, it is clear that more simulations should be run to better understand how different distributions result in more or less solutions, including one optimal location.

6. CONCLUSIONS

The aim of this research was to develop an agent based model of spatial competition that is capable of reproducing the results of the analytical Hotelling model and also provides meaningful results for cases where the distribution of consumers is not log concave.

Our agent based model is based on decision-rules that are consistent with the concept of Nash equilibrium. The aim of the model is to find the optimal location for a firm to be located, considering the distribution of consumers. The model has been programmed in a Netlogo environment.

After having solved the issue with discrete consumers, we focused on random distributions of consumers, using location equilibria following from 10.000 simulations of which 442 found a (sub) optimal location for shops A and B. In these simulation, using a world size of 80 (i.e. 40 inhabited locations), each location was assigned a random number of consumers, drawn from a Poisson distribution with λ=10. This resulted in a randomly generated distribution of location. s, of which none was log concave.

We then performed two OLS regressions: one that regresses the distance between equilibrium locations found in our simulations against several distribution measures, and one that regresses the average position (or the midpoint between positions) against the same distribution measures.

The regressions confirm theoretical findings that firms tend to follow consumers in choosing locations; if the bulk of the consumers is located more to the left, firms will also locate to the left. If consumers concentrate around the centre, firms will also move towards the centre and vice versa.

After this, shop A increases its price, the consumers choose again, and shop A evaluates the effect on its profits. If the new profit is higher than the initial profit, the shop will keep on increasing the price until its profit does not increase anymore. However, if the new profit is not higher than the initial profit, shop A will decrease the price until its profit no longer increases. Once shop A cannot increase its profit any more, shop B will go through the process of increasing or decreasing its price to find the optimal profit in that situation. Both shops repeat the process until neither of them can increase their profits unilaterally and a Nash equilibrium in prices is reached. The model then moves to the location stage. Our contribution to the literature lies in the fact that our model does not require the distribution of consumers (or preferences) to be log concave. This would for instance be highly relevant for distributions of preferences with multiple peaks. Moreover, the model could generate primes for empirical studies of location choice.

After this quantitative assessment of the solutions of 442 distributions found while using the same starting position, we explored the effect of choosing different starting positions. Therefore, for 30 distributions, we run the model for 400 starting positions. In 14 cases the model found more or less (sub) optimal locations. In general, more locations were found for the same distribution. In 9 of the 14 cases, a clear optimal location could be found, with for both shops the highest profits, among the total set of locations. The other six cases, with often a relatively high number of possible locations, quite far from each other, certain solutions are more profitable for shop A and other for shop B.

The wide variety of results from the case-studies illustrate the need for building a larger database of detailed simulation runs.

7. REFERENCES

[1] Caplin, A. and B. Nalebuff, 1991, Aggregation and imperfect competition: on the existence of equilibrium, *Econometrica,* **59**, pp. 25–59

[2] Lijesen, M.G., 2010, Empirical applications of spatial competition; an interpretative literature review, VU-FEWEB Research Memorandum 2010-6. http://dare.ubvu.vu.nl/bitstream/1871/15800/2/2010-6.pdf

[3] Neven, D.J., 1986, On Hotelling's competition with non-uniform customer distributions, *Economics Letters*, **21**, 121-6

[4] Tabushi, T. and J.-F. Thisse, 1995, Asymmetric equilibria in spatial competition, *International Journal of Industrial Organization*,**13** , 213-227

[5] Wilensky, U. (1999). NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.