A Data-Driven Examination of Hotelling's Linear City Model

Xin Liu
University of Pittsburgh
Pittsburgh, Pennsylvania
xil178@pitt.edu

Konstantinos Pelechrinis
University of Pittsburgh
Pittsburgh, Pennsylvania
kpele@pitt.edu

ABSTRACT

In his seminal work *stability in competition*, Hotelling developed a model for identifying the spatial equilibrium for two competing firms such that they maximize their market-share. He considered a linear area of fixed length and he showed that in this setting the two competing firms should be located side-by-side in the middle of the line. Hotelling's study has been then adopted and used to analyze and explain other phenomena in a variety of fields. However, the linear city model is purely theoretical, without any empirical validation. The goal of this study is to explore Hotelling's Law in its original space - i.e., that of firm competition - and identify possible adjustments needed to describe its application/validity in a non-linear city. In particular, we collect data from location-based social networks that include information for the number of customers in a venue and we compare them with the expectations from Hotelling's original law. Overall, we identify that at a large geographic scale there is correlation between the market-share and the inter-venue distance, which is consistent with the Hotelling's Law. However, as we zoom into smaller scales there are deviations from the expectations from Hotelling's law, possibly due to higher sensitivity to the necessary assumptions. Our findings enhance the literature on optimal location placement for a venue and can provide additional insights for owners in regards to the linear city model.

ACM Reference Format:
Xin Liu and Konstantinos Pelechrinis. 2019. A Data-Driven Examination of Hotelling's Linear City Model. In *11th ACM Conference on Web Science (WebSci '19), June 30-July 3, 2019, Boston, MA, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3292522.3326020

1 INTRODUCTION

"Location, location, location". This has been the mantra of real-estate agents when it comes to valuing a property. This is true for local businesses as well. Location is an important driver for the success of a business, that being a restaurant, a cafe, a retail store etc. [27]. Location can impact the success of a business both through convenient access to it, but also through agglomeration of other businesses (even if they are competing in nature). For example, it may be reasonable for a venue v to operate in an area with limited options for the service provided by v in order to achieve oligopoly/monopoly. However, competition can be beneficial in a multitude of ways. For one, it motivates the improvement of the services provided by v [5]. More importantly though, an area with multiple venue options will overall attract more people that are interested in exploring the area and hence, all the venues will potentially enjoy the benefits from the associated network effects [21].

There has been a significant volume of research on the impact of location on the revenue of a local business and the of *optimal* sites for a business. One of the first pieces that dealt with the problem of business location is what is known today as Hotelling's Law or Hotelling's Linear City (HLC) Model [13]. What Hotelling theorized, is that if there are two competing stores v_1 and v_2 on a street and their goal is to maximize their market share, ceteris paribus, their optimal location is right in the middle of the street, next to each other.

More specifically, Hotelling developed a linear city model to identify a spatial equilibrium point for two competing venues that offer identical products aiming at maximizing their market-share. Hence, as shown in Figure 1, for location z on the street segment (i.e., $z \in [0, 1]$) and M total customers uniformly distributed over the linear street segment, there will be $M \cdot z$ customers on the segment $[0, z)$ and $M \cdot (1 - z)$ customers on the segment $(z, 1]$. The utility a customer will get by visiting venue j (i.e., v_j) depends on (i) the unit transportation cost t, (ii) the prices at venue j, p_j and (iii) the gross customer surplus from visiting venue j, s_j. Overall, the utility is expressed by:

$$u_j = s_j - p_j - t \cdot d_j. \quad (1)$$

where d_j is the distance between the customer and v_j. Under the assumption that the market is *covered* (i.e., the excess surplus for every customer is large enough for them to be willing to buy), and that v_1 is at point $z = 0$ and v_2 is at point $z = 1$, the equilibrium point $\hat{z}$ at which a customer is indifferent to the two venues is [18]:

$$\hat{z} = (t + p_2 - p_1 + s_1 - s_2)/(2t). \quad (2)$$

This is also the location where the two venues need to locate their stores. For the case of venues that offer identical services, customers will get the same gross surplus from the two venues (i.e., $s_1 = s_2$), while the prices will be equal (i.e., $p_1 = p_2$), leading to $\hat{z} = 1 - \hat{z} = \frac{1}{2}$. Therefore, HLC states that the Nash equilibrium in this case is for the two venues to be located one by the other on the middle of the line. Despite the rather unrealistic assumptions of HLC (e.g., uniform distribution of customers, linear geography etc.), the model provides very important insights that can impact the site location of businesses. HLC is generalized as Salop's circular city model [18], where N venues (instead of 2 venues) are distributed uniformly over a circle. More generally, HLC implies that businesses will benefit if they are located near their competitors. In particular, given that the businesses offer similar quality product and at similar prices, each should get their fair share of the market.

WebSci '19, June 30-July 3, 2019, Boston, MA, USA

ACM ISBN 978-1-4503-6202-3/19/06...$15.00
https://doi.org/10.1145/3292522.3326020

This is counter-intuitive at a first glance, since this co-location will significantly increase competition and there can be customer spill-over from v_1 to v_2 and vice-versa. Of course, Hotelling's Linear City model is a purely theoretical law that is based on a set of (idealized) assumptions, such as a uniform mobility model. Under the latter, customers will arrive uniformly on the length of the street and therefore, v_1 will be the closest venue to approximately half of them and v_2 the closest for the rest half. Starbucks have been reported to make site choices based on the mobility patterns and even have stores very close to each other (as HLC would suggest) [1], while clustering observed in gas stations [26], fast food restaurants [25], pharmacies [11] and other retail stores [24] is also an artifact of ideas in Hotelling's Law. Despite these anecdotal stories, to the best of our knowledge, the *validity* of HLC has not been examined through real-world dataset, and especially, in the digital marketing and social media era, where *hyper-locality*, i.e., the ability to obtain information for areas not in our immediate vicinity, can alleviate some of the disadvantages experienced by a remote, potentially isolated, venue.

In this study, we use data from Foursquare, the largest location-based social network to date, and iExit, a service that provides information about Points-of-Interest at highway exits, to explore whether the outcomes expected by HLC hold in the real-world. We focus on a specific setting in order to minimize influence from other factors to the extent possible. In particular, we examine the distribution of customer visits to fast food restaurants that are clustered in highway exists. This mimics the original setting of HLC to a large extent - that is, venues that offer the same service and in similar prices, while also being very close to each other. We analyze the number of customers as a function of the distance between the venues in different scales. Our results indicate that, as we zoom-in to smaller (geographic) scales, the *validity* of HLC is reduced. This can possibly be attributed to many of the Hotelling's Law assumptions that might be harder to hold in the real-world in small scales. The contribution of our work is twofold: we provide evidence for the validity of HLC in a real-world setting and that in small geographic scales some of Hotelling's Laws assumption might not hold.

Related Literature: A large number of studies have analyzed the spatial influences on store competitions and customers' characteristics (e.g., [10, 12, 23]). Furthermore, there is a large volume of studies that aim into identifying optimal locations for business stores and Daskin *et al.* [9] present a detailed taxonomy of the various sub-problems in optimal site identification. In general, a retail store is expected to be more successful if located within a shopping center or a central business district, which provides convenient transportation access and attractiveness [17]. Given also the correlation between retail store density and street network centrality [19, 20], a central location will be preferable. Jensen [14, 15] also considers network effects in interactions between different types of venues, while Aboolian *et al.* [2] develop a spatial interaction model that seeks to simultaneously optimize location and design decisions for a set of new venues. However, the proposed model assumes a purely homogeneous customers base, that is, all customers are identical with respect to their venue preferences and expenditure decisions. Furthermore, the lack of detailed customer volume for the venues, creates issues for estimating the potential total market share for a venue in an area. In a subsequent series of studies [3, 4] the authors assume that customers' demand for a venue v (i.e., the probability of visiting the venue) decreases with the distance from v and increases with the *attractiveness* of it. Based on this assumption, the authors provide a spatial interaction model for locating a set of new facilities that compete for market share. Their models are applied and evaluated on synthetic data, showing the efficiency of the proposed algorithmic solution to the discrete multi-venue competitive interaction optimization problem. However, it is not clear how they will perform in the real-world. Other approaches [6] identify the optimal location for a store by maximizing the number of customers expected to be covered taking mobility patterns into consideration . Bozkaya *et al.* [7] use a genetic algorithm to select a single site among several candidate locations to maximize the market share under budget constraints. The results are verified by a real-world dataset of a supermarket chain in City of Istanbul. However, the methods of selecting candidate locations are not explained. Furhtermore, Karamshuk *et al.* [16] study the predictive power of various geographic and mobility-related features on the popularity of retail stores in New York City using Foursquare data. Even though the authors do not focus particularly on the HLC model, some of the features chosen are relevant to Hotelling's Law (at least at an intuitive level). In particular, the number of competing venues in an area, the density of venues in an area as well as the diversity of venues within an area are features that are examined among others.

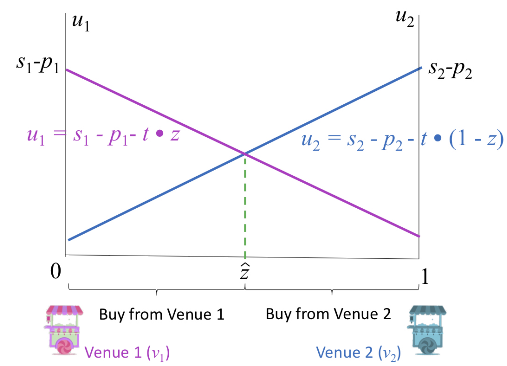

Figure 1: Based on the HLC model, a user that is on the middle of the street ($\hat{z} = 0.5$) is indifferent to the two venues. This is also the location where the two venues need to locate their stores to maximize their customer share.

2 DATASET AND EXPERIMENTAL SETUP

In order to complete our study we need data for customer visitation in local businesses. Furthermore, as aforementioned, we want to focus on cases that *match* - to the extend possible - the assumptions made by the HLC model. For that we focus on a specific type of venues, namely fast food restaurants, and on a particular environment, that is, highway exits. With this setting, our analysis will include venues that offer the same service, at very similar quality and price points.

For our study we focus on the state of Pennsylvania (PA), and we use the iExit API[1] that provides information about points-of-interest at highway exists. Every point of interest is a tuple of the following form: <id, phone, latitude, longitude, address, name, category, rating, price tier, brand name, exit ID>. We collect a total of 1,537 tuples that

[1] https://iexit.readme.io/

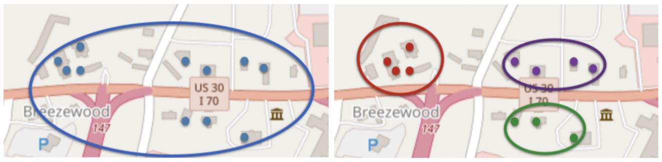

Figure 2: The different scales for our analysis. At a large scale (left) we consider a set of individual venues that are accessible from the same highway exits. At a small scale analysis we consider clusters of venues (right).

correspond to fast foods over the highway network in Pennsylvania over 482 exits. 95.1% of them belong to the lowest price tier, which means that venues in our dataset have very similar prices. We then query Foursquare's public venue API[2] to obtain information about the number of check-ins in each of these venues.

We begin by calculating the average daily check-ins for every venue in our dataset. For this we use the number of days that each venue has been on Foursquare up to the day of data collection (i.e., 19/06/2018). By using the average daily number of check-ins we essentially alleviate problems associated with the fact that older venues might have higher number of total check-ins simply by virtue of being on the system for longer. We also want to filter out venues that did not have a check-in for an extended period of time, which can be a sign of a venue that has been closed, and hence, we remove all venues with average daily check-ins less than 0.1.

According to our earlier discussion when venues - offering similar service quality at similar prices - are co-located, they expect to obtain their fair share of the market. That is, if there is a total of N check-ins in an *area* and k *closely* located venues, each will get approximately (N/k) check-ins. Based on this observation, the objective of our study is to examine whether there are any deviations from this expectation.

Geographic Scale: Intuitively, referring to the original setting for HLC, the geographic scale is related to the length of the linear street and the part of the street covered by businesses and customers. In our case the geographic scale will be initially defined through the area covered from the venues in the vicinity of each highway exit. Based on the data from the iExit API, a venue can be accessed from multiple exits[3]. Venues sharing the same group of highway exits through which they are accessible are all geographically close to each other and they can be considered as co-located at a large-scale. There is a total of 150 such clusters that will form our initial analysis unit and we will refer to them as *large* clusters. In this setting, every venue (i.e., fast food restaurant) within the cluster will be one of the competing venues in the Hotelling's Law. This setting is presented on the left part of Figure 2, where the blue circles correspond to a large cluster of fast food restaurants around a set of highway exits.

In order to examine how the expectations from HLC are met as we *zoom in* to smaller scales, we further divide the large clusters to smaller sub-clusters and explore the market share in these smaller scales. We identify sub-clusters based on their density using HDBSCAN with haversine distance [8]. We identify a total of 256 sub-clusters, which we will refer to them as *small* clusters. Focusing on these smaller clusters we are going calculate the market share of each venue within this smaller scale and compare it with the expectations from HLC. We will refer to this case as the small scale setting (right part of Figure 2). To obtain an idea of the actual length scales the different settings refer to we calculate the maximum pairwise distance in each case for all the clusters. The average of these maximum pairwise distances are 2.44 miles and 0.51 miles for large and small scale respectively.

Recall that the HLC model suggests that it is at the best interest of competing venues to locate as close to each other as possible. In order to capture the degree to which this is satisfied in each setting, we calculate the average pairwise haversine distance of venues within the area of interest. In particular, in the case of the large scale this corresponds to the average pairwise distance of all the venues in a large cluster. Formally, with $d(i, j)$ being the distance between venues i and j, that belong to cluster A, the average pairwise distance of venues d_A, in cluster A, is given by:

$$d_A = [\sum_{i,j=1, i \neq j}^{N} d(i,j)]/N_d \tag{3}$$

where N is the number of venues in cluster A and $N_d = N(N-1)/2$ is the number of venue pairs within A. For the case of the small scale setting the average pairwise distance is calculated in the same way, using only the venues within the corresponding small cluster.

Market Share Fairness: Every set of venues A (let us assume a large-scale cluster WLOG) can be described through a vector $C_A = [c_1, c_2, \ldots, c_N]$, where c_i is the average daily check-ins in venue i of cluster A. According to the HLC model, vector C_A should exhibit *fairness*, that is, every venue in A obtains their fair share of the market. To quantify the market share fairness f_A in the set of venues A we are going to use the coefficient of variation of C_A [22]:

$$f_A = \text{std}(C_A)/\text{mean}(C_A) \tag{4}$$

When the total market within cluster A is allocated fairly across the venues in A, f_A will be 0. Hence, a smaller value of f_A corresponds to more fair allocation of the market share within the cluster venues.

3 RESULTS

In order for our analysis to provide evidence supporting (or not) the HLC model, we should examine the correlation between d_A and f_A. Based on HLC, this correlation needs to be positive, i.e., smaller average pairwise distance for the venues in A should lead to more fairness in the market share, which corresponds to smaller f_A. Of course, simply examining the correlation between these two variables (e.g., through the Pearson correlation coefficient) is not appropriate since there are other factors that can affect the market share fairness for a set of venues even if our dataset consists of venues with similar quality of service (all are fast foods venues) and similar pricing. Hence, in order to control for these factors we will start by building a regression model where our dependent variable will be the market share fairness f_A and our independent variables will include various covariates - additional to d_A - that can have an impact on f_A. In particular, we include the following three variables in our model:

Venue reputation: Even though one expects that fast food restaurants offer a similar quality of service, the *reputation* of specific brands might impact the market share they get. To get an estimate for the reputation of a venue we use the average Foursquare rating of all the brand's venues in the 10 largest cities in PA. We then use the

[2] https://developer.foursquare.com/

[3] These exits typically correspond to different directions on the same highway, or exits that are located close enough to provide accessibility to the same venues.

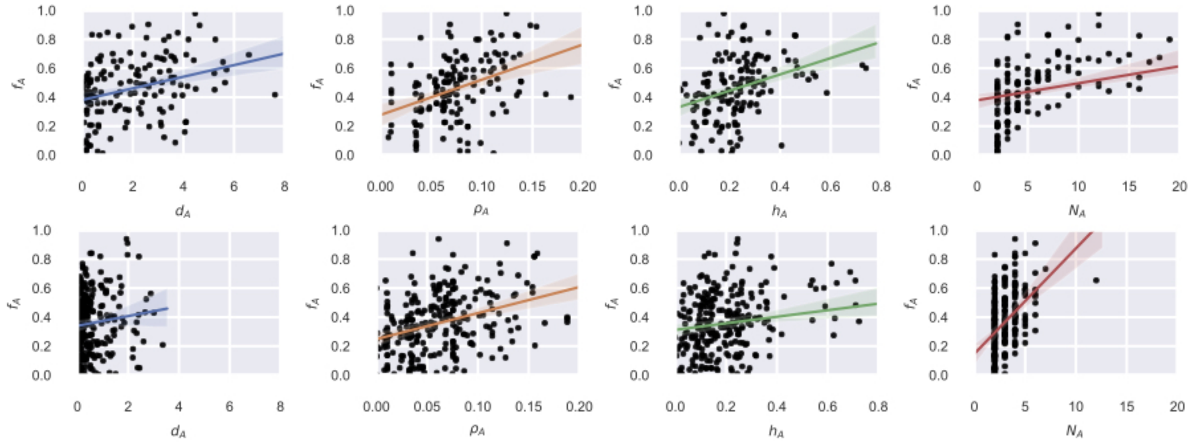

Figure 3: The relationship between pairwise venue distance d_A and market share fairness is strong in the large scale environment (top row) as compared to the small scale environment (bottom row).

coefficient of variation for the reputation ρ_A of the venues within a cluster A as an independent variable in our regression.

Hours of operations: If a venue within a cluster has significantly different hours of operations (e.g., shorter hours of operations), then this will potentially affect the market share it obtains. Hence, we collected hours of operation for every venue in our dataset and calculated the coefficient of variation for the weekly hours of operations h_A for the venues in each cluster A.

Number of venues: Figure 3 presents the correlations between the independent variables and the market share fairness for both scales examined, while Table 1 presents the results from our regression models. As we can see, even after controlling for other covariates, the distance between the venues is still significantly and positively correlated with the market share fairness in the large scale setting. However, in the small scale the relationship is less strong and not significant. Consequently, it has very limited explanatory power. While in the large scale setting the average venue pairwise distance explains about 11% of the total variance in the market share fairness, in small scale it merely explains 1% of it.

Part for this difference could be attributed to the much smaller variability of the pairwise distance in the small scale setting (as it is evident from the x-axis range in Figure 3). In particular, the variance of d_L for the large scale clusters L is $\sigma^2_{d_L} = 3$, while for the small scale clusters S is $\sigma^2_{d_S} = 0.45$. With a small variability in the regressor it is extremely difficulty to identify any meaningful relationship even if one exists. However, apart from that one of the key ideas behind the HLC model is that the distance to the venue is an important factor in the decision making process. For the small scale setting, since all venues are extremely close to each others the market share fairness can be very sensitive to other parameters that we have assumed are similar among venues in our setting (e.g., pricing and service quality), while factors such as the venue reputation are more important for the customer's decision (thus, if there is a larger skew in the reputation of the cluster's venues this translates to a skew at the market share). In contrast, at the large scale, while venues are relatively close to each other as well, they are also reachable from many different highway exits. This means that specific venues might be preferable to others purely based on the *direction* of arrival in the large cluster, leading to a more fair share of the market. I.e., the relative co-location of venues attracts drivers from many different exits but then their relatively *larger* pairwise distance - as compared to that between venues in the small clusters - can be the deciding factor for the customer's choice. We visualize this idea in Figure 4, where on the top we have two venues belonging to the same large cluster, while on the bottom we have two venues belonging to the same small cluster. In the former case, the venues are accessible from different exits and the circled areas include the ingress points from the highway. Customers within these areas will prefer the corresponding venue (similar to the HLC model on a line). However, in the small scale cluster, these areas have significant overlap, which means that now customers from this area might use other criteria to choose between these venues.

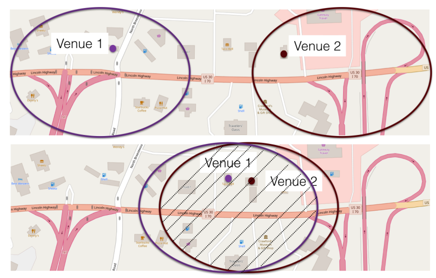

Figure 4: In a large scale cluster (top), two venues will be accessible from different exits leading to a more fair allocation of the customers, while in a small scale cluster (bottom) there will be significant overlap in the *service areas* of the venues and hence, the market share can be extremely sensitive to other factors.

Overall, we can say our results support the HLC model at a high level. In particular, at a large scale, the co-location of venues provides agglomeration effects, while at the same time the slight distance between the venues allows them to attract customers arriving from different directions and hence, obtain a market share closer to the fair allocation. As we move to smaller geographic scales, the strength of HLC model reduces, and this might be due to the fact that in very small scales, given the negligible distance between venues (especially compared to the overall distance traveled on the highway), the market

	large scale		small scale	
variable	**d_A only**	**all features**	**d_A only**	**all features**
intercept	0.381***	0.166***	0.340***	0.111***
	(0.025)	(0.043)	(0.017)	(0.030)
d_A	0.0402***	0.0170*	0.0337*	0.0092
	(0.010)	(0.010)	(0.019)	(0.017)
ρ_A		1.7273***		1.0881***
		(0.432)		(0.299)
h_A		0.3870***		0.0524
		(0.115)		(0.067)
N_A		0.0060***		0.0565***
		(0.002)		(0.010)
N	150	150	256	256
R^2	0.108	0.345	0.012	0.250

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 1: Our regression model results.

share can be very sensitive to HLC assumptions (e.g., quality of service and pricing). Thus, our results do not necessarily say that Hotelling's Law does not hold in small scales, but (possibly) that the model is more sensitive to its assumptions in this scale.

4 DISCUSSIONS AND CONCLUSIONS

In this paper our objective has been to explore the *validity* of Hotelling's Law in the real-world. We identified a specific setting that satisfies the main assumptions of HLC, namely, similar service quality and pricing. Under HLC venues will maximize their market share (and in fact they will get their fair share of the market) when they are located close to each other. Our results indicate that for large scale clusters, when the average pairwise distance between the venues is smaller, the market share for each venue is closer to its fair share, supporting the idea that clustering of competing venues can have significant agglomeration effects for a business. When focusing on smaller scale clusters, this relationship is not significant anymore, potentially due to the higher sensitivity of the model to its assumptions.

Of course, our study exhibits limitations. Firstly, even though mobile and web-platforms like Foursquare offer a convenient way to obtain data that can verify theoretical constructs (like the HLC model), we are aware of possible biases that exist in the use of platforms such as Foursquare and the underlying assumption is that check-in volume is representative of actual *popularity*/customers of a venue, which might not be true. Furthermore, despite the evidence provided here in support of HLC, the setting analyzed is not very representative of a dense urban environment. For example, in the highway setting we focused on in this study, the transportation cost can be neglected given the type of trips that pass through these highways/exits. In an urban environment, especially a dense one, this can be crucial. Moreover, we focused specifically on fast food restaurants that tend to exhibit similar characteristics in terms of quality, pricing and outlook to customers. However, in an urban environment venues that seemingly offer the same service (e.g., restaurants) can exhibit significantly diverse characteristics, and clustering might have small impact on the market share a business obtains. In addition, the sensitivity to the HLC's model assumptions can be higher in a dense urban setting, similar to the small scale setting in this study.

REFERENCES

[1] 2012. Starbucks' Store Location Approach Examined. https://www.10news.com/news/starbucks-store-location-approach-examined.

[2] Robert Aboolian, Oded Berman, and Dmitry Krass. 2007. Competitive facility location and design problem. *European Journal of Operational Research* 182, 1 (2007), 40–62.

[3] Robert Aboolian, Oded Berman, and Dmitry Krass. 2007. Competitive facility location model with concave demand. *European Journal of Operational Research* 181, 2 (2007), 598–619.

[4] Robert Aboolian, Oded Berman, and Dmitry Krass. 2009. Efficient solution approaches for a discrete multi-facility competitive interaction model. *Annals of Operations Research* 167, 1 (2009), 297–306.

[5] Rajiv D. Banker, Inder Khosla, and Kingshuk K. Sinha. 1998. Quality and Competition. *Management Science* 44, 9 (1998), 1179–1192. http://www.jstor.org/stable/2634708

[6] Oded Berman and Dmitry Krass. 2002. The generalized maximal covering location problem. *Computers & Operations Research* 29, 6 (2002), 563–581.

[7] Burcin Bozkaya, Seda Yanik, and Selim Balcisoy. 2010. A GIS-based optimization framework for competitive multi-facility location-routing problem. *Networks and Spatial Economics* 10, 3 (2010), 297–320.

[8] Ricardo J. G. B. Campello, Davoud Moulavi, and Joerg Sander. 2013. Density-Based Clustering Based on Hierarchical Density Estimates. In *Advances in Knowledge Discovery and Data Mining*, Jian Pei, Vincent S. Tseng, Longbing Cao, Hiroshi Motoda, and Guandong Xu (Eds.). 160–172.

[9] Mark S Daskin. 2008. What you should know about location modeling. *Naval Research Logistics (NRL)* 55, 4 (2008), 283–294.

[10] Alon Eizenberg, Saul Lach, and Merav Yiftach. 2015. Retail prices in a city: an empirical analysis. *Maurice Falk Institute for Economic Research in Israel. Discussion paper series.* 7 (2015), 0_1.

[11] Susan Feyder. [n. d.]. Nation's biggest pharmacies sidle right up to each other. Retrieved November 1, 2018 from http://www.startribune.com/nation-s-biggest-pharmacies-sidle-right-up-to-each-other/176188911/

[12] Rachel Griffith, Ephraim Leibtag, Andrew Leicester, and Aviv Nevo. 2009. Consumer shopping behavior: how much do consumers save? *Journal of Economic Perspectives* 23, 2 (2009), 99–120.

[13] Harold Hotelling. 1990. Stability in competition. In *The Collected Economics Articles of Harold Hotelling*. Springer, 50–63.

[14] Pablo Jensen. 2006. Network-based predictions of retail store commercial categories and optimal locations. *Physical Review E* 74, 3 (2006), 035101.

[15] Pablo Jensen. 2009. Analyzing the localization of retail stores with complex systems tools. In *International Symposium on Intelligent Data Analysis*. Springer, 10–20.

[16] Dmytro Karamshuk, Anastasios Noulas, Salvatore Scellato, Vincenzo Nicosia, and Cecilia Mascolo. 2013. Geo-spotting: mining online location-based services for optimal retail store placement. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 793–801.

[17] Alexander Kubis and Maria Hartmann. 2007. Analysis of location of large-area shopping centres. A probabilistic Gravity Model for the Halle–Leipzig area. *Jahrbuch für Regionalwissenschaft* 27, 1 (2007), 43–57.

[18] Andreu Mas-Colell, Michael Dennis Whinston, Jerry R Green, et al. 1995. *Microeconomic theory*. Vol. 1. Oxford university press New York.

[19] Sergio Porta, Vito Latora, Fahui Wang, Salvador Rueda, Emanuele Strano, Salvatore Scellato, Alessio Cardillo, Eugenio Belli, Francisco Cardenas, Berta Cormenzana, et al. 2012. Street centrality and the location of economic activities in Barcelona. *Urban Studies* 49, 7 (2012), 1471–1488.

[20] Sergio Porta, Emanuele Strano, Valentino Iacoviello, Roberto Messora, Vito Latora, Alessio Cardillo, Fahui Wang, and Salvatore Scellato. 2009. Street centrality and densities of retail and services in Bologna, Italy. *Environment and Planning B: Planning and design* 36, 3 (2009), 450–465.

[21] Stuart S Rosenthal and William C Strange. 2004. Evidence on the nature and sources of agglomeration economies. In *Handbook of regional and urban economics*. Vol. 4. Elsevier, 2119–2171.

[22] Michael Sheret. 1984. Note on Methodology: The Coefficient of Variation. *Comparative Education Review* 28, 3 (1984), 467–476. http://www.jstor.org/stable/1187327

[23] Howard Smith and Donald Hay. 2005. Streets, malls, and supermarkets. *Journal of Economics & Management Strategy* 14, 1 (2005), 29–59.

[24] Ken Steif. [n. d.]. Why Do Certain Retail Stores Cluster Together? Retrieved November 1, 2018 from https://www.planetizen.com/node/65765

[25] Presh Talwalkar. [n. d.]. Why are McDonaldś and Burger King usually located near each other? Fast food location game theory. Retrieved November 1, 2018 from https://mindyourdecisions.com/blog/2012/10/23/why-are-mcdonalds-and-burger-king-usually-located-near-each-other-fast-food-location-game-theory/

[26] Presh Talwalkar. 2014. *The joy of game theory: An introduction to strategic thinking.* CreateSpace Independent Publishing Platform.

[27] Kerry D. Vendell and Charles C. Carter. 1993. Retail Store Location and Market Analysis: A Review of the Research. *Journal of Real Estate Literature* 1, 1 (1993), 13–45. http://www.jstor.org/stable/44103233