Doctoral Dissertation
Doctoral Program in Urban and Regional Development (33th Cycle)

Consumers, city networks and commercial patterns

An investigation of retail locations in the city of Turin, Italy

By

Roberta Taramino

******

Supervisor(s):
Prof. Luigi Buzzacchi, Supervisor
Prof. Giulio Zotteri, Co-Supervisor

Doctoral Examination Committee:
Prof. Teresa Graziano, Università di Catania
Prof. Mario Paris, Politecnico di Milano
Prof. Johnatan Reynolds, Said Business School
Prof. Giuseppe Stigliano, Libera Università di Lingue e Comunicazione
Prof. Francesca Montagna, Politecnico di Torino

Politecnico di Torino
2021

Declaration

I hereby declare that, the contents and organization of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

Roberta Taramino

2021

* This dissertation is presented in partial fulfillment of the requirements for **Ph.D. degree** in the Graduate School of Politecnico di Torino

Contents

List of Figures

List of Tables

Introduction

Retail firms provide consumers with goods and services for consumption. They operate in an oligopolistic market of differentiated goods, and they deliver either products or services at an observable price. They enter the local market, compete and, possibly, exit the market. The commercial pattern, defined as the provisional distribution of stores in a moment in time, is the result of every single location choice made by retailers who start competing.

For a store owner, entering the local market means dealing with many key decisions. First, store owners have to search for the best combination of store attributes. They should select the products (or services) to sell and their prices, the size of the place where the products (or services) are to be sold and the service level to be offered to consumers consistently.

Second, they have to search for the best location given the selection of store attributes. Indeed, the choice of location is one of the most important decisions in retail since location is not meant to be only a spot on a map. The city cannot be viewed as if it were located on a featureless plan on which all land is of equal quality. Location is the relative position that a retailer can occupy given the proximity to other stores on the one hand and potential consumers on the other hand. Moreover, it denotes the position in the urban network, which is structurally different from one place to another. Some places are more accessible than others; some other places are more visible than others by construction.

The selection of store attributes drives the location choice, of course. A supermarket, which is characterized by a wide assortment of product lines, should be located in a building that is large and easily accessible to consumers by car; parking spaces must be available. A tobacco store, which provides consumers with a single product line, should preferably be located in a building at a place that is densely populated so that it is also perceived as accessible by pedestrian consumers. A high-end clothing store typically chooses a premium location where it is visible to consumers who are shopping and is possibly surrounded by many other top brands stores.

In short, depending on the selected mix of product, price and size, traditional retailers desire to be located in the best place selling the best product. The place can either be the best one because it is easily accessible by consumers or because of relative position with respect to competitors. Indeed, retailers should finally consider the competition that arises from being located near other stores. Spatial competition, indeed, can have two opposite results for the traditional store itself; on the one hand it can cannibalize the demand, on the other hand it can produce positive externalities. Moreover, such a competition can either arise between traditional stores or between stores of different distribution channels.

The thesis includes two empirical tests aiming to investigate commercial patterns as the result of a long process where store owners, having selected a set of store attributes, search for the best location for their store, considering both consumers (and their accessibility to the store) and competitors (and their typology).

The tests are described in chapters 3 and 4 and they are based on the analysis of a database that collects all the commercial licenses that have been issued in Turin (Italy) from 2005 to 2019.

The work is organized as follows. Chapter 1 reviews the literature on retail location choices and urban network measurements. First, it paves the way for understanding typical consumers' behaviors, such as comparison shopping and multi-purpose shopping. Second, it describes some urban network measurements that were used in the empirical tests in the thesis. The empirical tests described in chapters 3 and 4 are the core of the thesis. However, in order to ensure a better understanding of the empirical tests, a comprehensive description of the Turin case study and the database is provided in Chapter 2.

Chapter 1

What drives commercial patterns?

1.1 Commercial patterns, the type of retailer, the type of consumer and the type of good

Retail stores differ according to many variables; they sell different kinds of goods and serve consumers who change their behavior according to different purchasing opportunities. Indeed, *retail stores come in all shapes and sizes, and new retail types keep emerging* (Kotler and Armstrong, 2010: 367). The kinds of services they offer, the price they charge to consumers and the product lines they select are just a few of these variables[1].

Retail stores can be classified according to such variables, and they deal with competitors in two ways. First, they directly compete with retailers of the same category; second, they often cross-compete with retailers that belong to other categories. An example comes from one of the most important differences between retailers: they can sell the goods to the final consumer either in a physical store (namely, brick-and-mortar retailers), online (namely, e-tailers) or both.

Of course, internet retailing is one of the forces that has changed the dynamics of the retail industry (see Weltevreden and Rietbergen, 2009). Clearly, the survival of brick-and-mortar stores might depend on the pressure from internet retailing in the product category they cross-compete with, and all the areas of the city will be subject to the same competitive pressure from online retailers. Moreover, the survival of such physical stores is also threatened by the other brick-and-mortar stores operating in the same market.

[1] The thesis has an empirical core. Hence, in this chapter, we have deepened the classification of retailers according to the only variables provided in the database we used for empirical tests.

When focusing on physical retailers, the main object of the thesis, a reference must be made to the Italian Classification of physical retailers made in the D.lgs. n. 114 since 1998. Finally, after having analyzed all the variables characterizing stores[2], they can be of one of the following types (see Kotler and Armstrong, 2010: 367):

Table 1 - Stores retailers types mentioned above

Class	**Description**	**Sub-classes**
Specialty stores *D.lgs. n. 114/1998 (Art. 4, 1.d,e)*	They carry a narrow product line with a deep assortment. According to the commercial area, they can be as follows: (1) small stores if the commercial area is smaller than 250 sqm; (2) medium stores if the commercial area ranges between 250 sqm and 2500 sqm; (3) large stores if the commercial area is larger than 1500 sqm (or superstores as the category killers that carry a deep assortment in a particular category and have knowledgeable staff).	Stores selling products other than groceries Stores selling home accessories (furniture and furniture accessories) Stores selling cultural and recreational items (books, newspapers, toys) Stores selling electronics, computer appliances and sporting goods Stores selling pets' items and flowers
Convenience stores *D.lgs. n. 114/1998 (Art. 4, 1.d)*	They are typically located near residential areas, are open for long hours seven days a week, and carry a limited line of high-turnover convenience products at slightly higher prices (for example, mini-markets and stores selling fruits and vegetables)	Mini-markets, stores selling frozen products and other stores selling grocery Stores selling grocery and beverage Stores selling tobacco Stores selling medicines and cosmetics

[2] Among the long list of variables characterizing stores, we selected the only ones that are available in the database: the merchandise category the store sells, the commercial area of the store (that can be used as a proxy of the breadth and depth of their product lines) and their location. Finally, the classification we have presented is leaner than the one in the cited book (Kotler and Armstrong, 2010: 367) for two reasons: first, the classification of Italian stores is not as wide as the American ones; second, the available variables in the database we used for the empirical tests allowed us to distinguish only the types of physical stores you see in the list above.

Supermarkets *D.lgs. n. 114/1998 (Art. 4, 1.e)*	A medium sized store (commercial area ranges between 250 sqm and 2500 sqm), low-cost, low-margin, high-volume, self-service operation designed to serve the consumer's total needs for grocery and household products	Supermarket

In Table 1, the first column reports the names of the basic retail classes in the literature. The second column provides their basic description. In the third column, the names of the Italian stores' sub-categories are presented: the ones available in the database based on which the empirical tests were developed. For example, all the stores selling essential goods such as bread, fruit, meat, tobacco and medicines are grouped under the class of convenience stores. The other stores selling products that are not as essential are grouped under the class of specialty stores. This class includes a large variety of stores that can further be classified given their size.

Keeping in mind this basic distinction between retail stores, the thesis aims at describing what shapes the actual distribution of stores in a city, hereinafter, the commercial patterns. A commercial pattern is the observable distribution of retail stores and depends on three processes (Dennis et al., 2002): entry, location choice and exit. These processes are driven by some economic forces as well as random events. Still, over time, the economic forces tend to overcome the random events and shape the distribution of stores in the city. Indeed, retailers will be more inclined to locate their stores or relocate to (Carree and Thurik, 1996) where they expect to maximize their profits. Additionally, retailers who locate stores in profitable areas are more likely to survive than others.

In order to describe the economic forces shaping the commercial pattern, one must go through the literature about retail location given that the actual distribution of stores is the result of each retail location choice. That literature has a very long history.

In 1929, Hotelling described retailing spatial competition while considering two stores and a linear, bounded market. There, retailers selling homogeneous products agglomerate at the center of the linear market to benefit from a larger market share. He paved the way for the literature about retailers' agglomeration economies (DiPasquale and Wheaton, 1996); the *principle of minimum differentiation* (Boulding, 1966) formalizes homogeneous retailer agglomeration by maximizing consumer utility and allows the co-location of similar stores to enable comparison shopping at shopping sites given a guaranteed reduction in search costs (D'Aspremont et al., 1979; De Palma et al., 1985; Dudey, 1990).

Meanwhile, Christaller (1933) and, subsequently, Losch (1954) developed a two-dimensional model for modeling the spatial distribution of retailers offering identical products; it is known as the Central Place Theory. The model assumes that consumers, who are uniformly distributed along a horizontal line, make single-purpose shopping trips to the nearest store. Finally, each store patronizes a hexagonal catchment area equal and adjacent to one of its competitors' (Eppli and Benjamin, 1994). The Central Place Theory evolved over the years with the contribution of many researchers who gradually relaxed the assumption at the basis of the construction of hexagonal retail markets. One of the most debated is the assumption that consumers always make a single-purpose shopping trip to the nearest shopping center. Just as Hotelling paved the way for the idea of comparison shopping, the Central Place Theory helps scholars to consider multi-purpose shopping trips; they reflect on the probability that consumers who want to purchase a variety of goods may possibly travel to a more distant shopping district, at least for some products (Clark, 1968; Golledge et al., 1966; O'Kelly, 1983; Rushton, 1969).

The long history of the literature on retail location attributes to the consumers two typical behaviors depending on the kind of good they are going to purchase (Eaton and Lipsey, 1979). First, they can minimize the searching cost by performing a comparison between similar stores selling similar goods. Second, they can minimize the transportation cost by performing multi-purpose shopping in shopping districts that offer a large variety of products.

Hence, the kind of behavior the consumer exhibits while purchasing and the type of good the retailer sells compared to one of the competitors already operating in the local market drive location choices (Miller and Finco, 1995) and, hence, commercial patterns.

The thesis proposes a classification of retail stores with respect to incumbents and a consolidated classification of shopping behaviors exhibited by consumers; retailers can either decide to offer to consumers products that are identical, similar or complementary to those sold by competitors, and consumers can either behave as searching cost minimizers or transportation cost maximizers. Table 2 presents the patterns one should expect.

Table 2 - Commercial patterns that emerge by matching consumers' shopping behavior and the product sold by retailers with respect to competitors

	Identical products	Complementary products	Similar products
Minimize searching cost	**Homogeneous distribution** of retailers, each one being equally far from competitors		Similar stores concentrate in **clusters**
Minimize transportation cost		Complementary stores concentrate in **clusters**	

Retail location theories state that two possible commercial patterns can emerge given the different consumers' behaviors and the type of entering retailers with respect to competitors already operating in the same local market (Losch et al., 1954):

- First, retailers disperse all over the territory; the distribution of retailers is spatially uniform (i.e., each one is equally far from competitors) when retailers offer products that are identical to the ones that incumbents sell (Christaller, 1933; DiPasquale and Wheaton, 1996). This is the case of retailers selling standard products that are very similar in price and other product characteristics. The consumer who wants to buy a single (Brown, 1993) standard product typically shops at the retailer whose total delivered price is the lowest. Indeed, he/she wants to minimize the total cost of searching[3] for the good he/she wants to purchase (Wolinsky, 1983).
- Second, retailers agglomerate all over the territory; stores concentrate in clusters (Jensen et al., 2005). Shopping clusters emerge for two reasons; stores can cluster if they sell products that are complementary to that of competitors (Fischer and Harrington, 1996) or if they sell products that are differentiable in price and other good features (Basker et al., 2016; Fujita and Thisse, 2002; Koster et al., 2019; Netz and Taylor, 2002).

[3] Price, transportation cost and inventory cost to store the product at home

In the first case, consumers want to buy multiple goods (Burdett and Malueg, 1981; Carlson and McAfee, 1984) and, hence, decide to minimize the transportation cost[4] only. In this sense, agglomerating also represents a winning strategy for retailers with respect to those customers who are extremely sensitive to transportation cost (Eppli and Benjamin, 1994; Eppli and Shilling, 1995) and go from store to store to pick the best items (namely, cherry-picking; Fox and Hoch, 2005). Indeed, purchasing complementary goods in the same location costs less than purchasing the same goods on many separate trips (Ghosh and Craig, 1983; Mulligan, 1988).

In the second case, stores cluster in order to increase the possibility for consumers to compare similar products (Eaton and Lipsey, 1979; Picone et al., 2009) and provide the consumer with a guaranteed ability to find the best match at the lowest searching cost. In other words, the higher the degree of differentiation between the products sold by competing retailers, the higher is the benefit they gain from clustering. Products can either be different because they belong to different merchandise categories (complementary products) or because they belong to the same merchandise category of goods that are differentiable according to any product feature (partially substitute products).

Product differentiation is at the basis of the next type of classification of stores I'm introducing. Products that have been so far described as homogeneous, complementary and similar to others can also be classified according to merchandise category.

The thesis uses a consolidated classification of products categories defined by Holton in 1958 following the definition provided by the Definition Committee of the American Marketing Association in 1948. Even if such a classification is old, it is the most useful when dealing with commercial patterns.

Holton classified goods as convenience, shopping and specialty ones. He stated, "*[T]he essence of the distinction between convenience goods and shopping goods may lie in the gain resulting from price and quality comparisons relative to the searching costs. For convenience goods this ratio is low, but for shopping goods the probable gain is large enough to call forth more extensive searching. Specialty goods seem to overlap both of the other categories and are distinguished only by the limited size of the market demand for the goods*" (Holton, 1958:1).

[4] The cost for moving from home to the store and back

In particular, from the consumers' point of view, the following should be noted:

- Convenience goods are not differentiable; they are "*purchased frequently, immediately and with a minimum effort*" (American Marketing Association, 1948). Convenience goods are those goods for which "*the probable gain from making price and quality comparisons among alternative sellers is thought to be small relative to the consumer's appraisal of the searching costs in terms of time, money and effort*" (Holton, 1958:2).
- Shopping goods are strongly differentiable; they are "*those goods which the consumer, in the process of selection and purchase, characteristically compares on such bases as suitability, quality, price and style*" (American Marketing Association, 1948). Hence, they are *"goods for which the probable gain from making price and quality comparison among alternative sellers is thought to be large relative to the consumer's appraisal of the searching cost in terms of time, money and effort*" (Holton, 1958:2).
- Specialty goods are "*those consumers' goods in which a significant group of buyers characteristically insist and for which they are willing to make a special purchasing effort*" (American Marketing Association, 1948).

Retail stores can be one of the following:

- They can be convenience stores if they sell products that are purchased by consumers who only want to minimize the time spent searching for the goods they want to purchase; they are not interested in comparing products before purchasing. This happens because convenience goods are standard in price and quality by nature. Hence, the effort eventually spent in comparing products is not gratified from a "better buy" (Holton, 1958:2) given that convenience products are the same in all the stores. Hence, stores selling convenience goods prefer to be located far from one another in order not to cannibalize the market. They prefer to be located as near as possible to consumers, satisfying their desire to minimize the searching cost.
- Specialty stores selling shopping goods sell products that are purchased by consumers who want to compare products before purchasing. Those kinds of goods are differentiated, indeed; they can be different among alternative stores in terms of price and other product characteristics. Herein lies the "best buy" option: a consumer who wants to purchase a shopping good aspires to buy the "best product" according to his needs and taste. Hence, stores selling shopping goods prefer to be located near one another in order to maximize the possibility for consumers to compare products and, hence,

increase the probability of them finding their "best" option by visiting many stores. So, all the stores together attract a higher number of shopping goods consumers (Konishi, 2005).

- Specialty stores selling specialty goods are stores that are reached by consumers who are interested in a particular brand or product and who are willing to make a special purchasing effort and cover a special cost to complete the purchase; for example, they can make a special trip to solely shop at that place. Hence, stores can be located wherever they want; they are "destinations" for consumers[5].

The thesis, however, takes into account only shopping and convenience goods as merchandise categories. Indeed, the list of stores selling specialty goods must consider some stores' sensitive information about the retailer and brand, which the database used for the empirical test does not account for.

Finally, the commercial pattern one should expect given consumers' behavior and merchandise categories sold by incumbents is the following:

Table 3 - Homogeneous, complementary and heterogeneous products given consumer behaviour and merchandise categories sold by retailers

		Incumbent stores	
		Specialty store selling shopping goods	Convenience store
New entrant store	Specialty store selling shopping goods	**Heterogeneous products**	**Complementary products**
	Convenience store	**Complementary products**	**Homogeneous products** if the new entrant sells the same merchandise category as the incumbent **Complementary products** if the new entrant sells another merchandise category with respect to the incumbent

[5] Namely, destination stores.

Stores selling shopping goods prefer to be clustered near each other in order to increase the possibility for consumers to compare products given that the products they sell are heterogeneous.

Stores selling convenience goods can be distributed in two ways; they are typically located far away from stores selling the same merchandise category and near stores selling another convenience good to provide consumers with the widest variety of goods. Additionally, stores selling convenience (shopping) goods benefit from being agglomerated with stores selling shopping (convenience) goods; together, they allow consumers to cherry-pick all the products they want in their basket in a single shopping trip. Hence, the transportation cost is minimized.

Table 4 - The commercial pattern we should expect given consumer behaviour and merchandise categories sold by retailers

		Incumbent stores	
		Specialty stores selling shopping goods	Convenience stores
New entrant store	Specialty stores selling shopping goods	**Agglomeration** search cost is minimized	**Complementary products** transportation cost is minimized
	Convenience stores	**Complementary products** transportation cost is minimized	**Homogeneous products** search cost is minimized **Complementary products** transportation cost is minimized

Table 4 presents the final summary of the literature review and describes the commercial patterns driven by the category of merchandise the store sells with respect to competitors and the possible consumer behaviors discussed in the literature on retail location.

1.2 Commercial patterns and the urban network

Commercial patterns should also be noted while considering that a city is not linear and "*can not be viewed as if it was located on a featureless plan, on which all land is of equal quality*" (Alonso, 1964). Due to the seminal work of Jane Jacobs (Jacobs, 1969), urban planners have started acknowledging the need to understand the social and economic interactions in the built environments (Andres Sevtsuk, 2014b). They all agree that locations are not meant to only be a spot on a map.

Buildings, indeed, are connected by routes and relate to each other with an intensity that increases with proximity. Hence, the more a street that passes in front of a building is used, the more a location is visible, and the more it is close to other amenities, the more the building happens to be attractive.

A considerable body of research from urban planners suggested methods for measuring the quality of a location and its position in the complex urban network (for example, Casalaina and Horst, 1967; Hidalgo et al., 2020; Horst, n.d.; Levin, 1964; March and Steadman, 1971; Miller and Wu, 2000; Okabe and Sugihara, 2012; Peponis et al., 2007, 2008; Porta et al., 2006; Sevtsuk and Kalvo, 2018; Sevtsuk and Mekonnen, 2012; Xie and Levinson, 2007; Yoshimura et al., 2020). Indeed, analyzing the complexity of the urban realm has become a matter of interest for urban planners in the last seventy years (Batty, 2005).

All the network-based models of the built environment "*encode [the] explicit relationship between the elements of the network, documenting, for instance, how streets are connected to one another, how long the travel times between different districts, buildings, or rooms are, or how many people commute between them*" (Sevtsuk, 2014b, p.4).

Dealing with retailers makes this reading of a city as heterogeneous and interconnected even more interesting. Indeed, as described earlier, retailing is a business activity that has to do with consumers; it is important for a store owner who wants to maximize his own profit to locate the store in locations that are perceived as subjectively attractive by potential consumers. Urban planning has often dealt with the distribution of retail activities in the urban network, seeing the city as the location where interconnections happen and attract consumers and competitors (Buzzacchi et al., 2020; Dolega et al., 2016; Kang, 2016; Ozuduru and Varol, 2011; Porta et al., 2009; Andres Sevtsuk, 2014a; Andres Sevtsuk and Kalvo, 2018; Yoshimura et al., 2020).

In order to empirically relate consumers' behaviors and the location decision made by store owners, one can rely on some measurement of the urban network that will return some measurements of locations' characteristics in terms of accessibility.

In this chapter, the Multiple Centrality Assessment[6] presented by Porta et al. in 2006 is described; it is the method for analyzing an urban network that we have selected and that was used for the empirical analysis in the thesis.

Conceptually, the MCA is based on *centrality;* the more a location is *central*, the more attractive it is. Centrality has an old history that starts from the idea of Bavelas (Bavelas, 1948, 1950; Leavitt, 1951; Shaw, 1964; Shimbel, 1953); he stated that *the more central a place is, the more power and control it has on other locations*.

Empirically, the MCA is based on two basic elements of the built environment: edges and nodes. The edges are the routes followed by travelers, and nodes are the edges' intersections (Andres Sevtsuk, 2014b).

Figure 1 - Edges and nodes of the built environment

[6] Hereby, MCA

Ultimately, *centrality* has a triple nature as comprehensively described in the work of Porta et al. in 2006. So, a node (i.e., location) can be defined as central when it has the following characteristics:

- It is located in areas that are dense due to the presence of other nodes, which is known as the closeness centrality;
- It is located in a street with heavy traffic, which is known as the betweenness centrality;
- It is located in a place that is highly visible, which is known as the straightness centrality.

Following the works of Sevtsuk (2014) and Porta et al. (2009), the thesis applies centrality to retailing both as a concept and as an empirical measurement.
The first adjustment one needs when dealing with *centrality* in retailing has to do with the definition of nodes. In Figure 1, stores are not located in the red nodes (i.e., the location where streets cross) but along the street itself on the ground floor of a building.

Regarding retailers, the object of the thesis, computing the *centrality* measurements revealed that the nodes for retailers are the buildings that are the potential location for stores, and they are projected to the street network.

Figure 2 - Edges and nodes for the built environment in retailing

Finally, given the triple nature of centrality associated with retailing, a building can be defined as central when it has the following characteristics:

- It is located in areas that are dense due to the presence of other nodes, which is known as closeness centrality. In general, retailers prefer to be located in places that are well attended by consumers, and it is more probable for buildings to be located in areas densely populated by other buildings.
  The closeness centrality for a general building is measured as it follows, and it basically measures the density of buildings (B) around building *i*:

$$Closeness_i(r) = \sum_{d_{ij}<r} B_j$$

  where the sum is extended to (B_j), which is the number of buildings located in the urban landscape, provided that the distance (d_{ji}) between the reference building i and the building j is lower than the radius *r*.

- It is located in a street with heavy traffic, which is known as betweenness centrality. Locations with heavy traffic are, of course, central in the eyes of each store owner. Buildings can get footfall either because they are close to other buildings[7] or because they are located between other buildings and, hence, are able to utilize the traffic generated by the buildings on their sides. In general, the betweenness indexes (*Betwenness*$_i$) consider all the routes from a generic building *j* and generic building *k*, passing next to a specific building *i* (see Porta et al., 2009; Sevtsuk, 2014):

$$Betweenness_i = \sum_j \sum_k \alpha^i_{jk} W_{jk}$$

  So, α^i_{jk} is equal to 1 if the shortest path from *j* to *k* passes through *i*, and it is 0 otherwise; W_{jk} is the (possibly estimated) number of individuals taking the route from *j* to *k*.

- It is located directly in line with other buildings, which is known as straightness centrality. A building that is located on straight roads benefits from the visibility of the building itself from all the other buildings located along the same straight line.

[7] See the closeness centrality.

This idea can be captured by the following index proposed by Porta et al. (2009):

$$Straightness_i(r) = \frac{1}{B(r)} \sum_{\substack{j \\ d_{ij} < r}} \frac{d_{ij}}{D_{ij}}$$

where the sum is extended to every one of the $B(r)$ buildings whose (Euclidean) distance d_{ij} from the building i is smaller than r. D_{ij} is the actual distance on the road network between building *i* and building j.

Figure 3 - The closeness centrality for all the building in Turin (r=600)

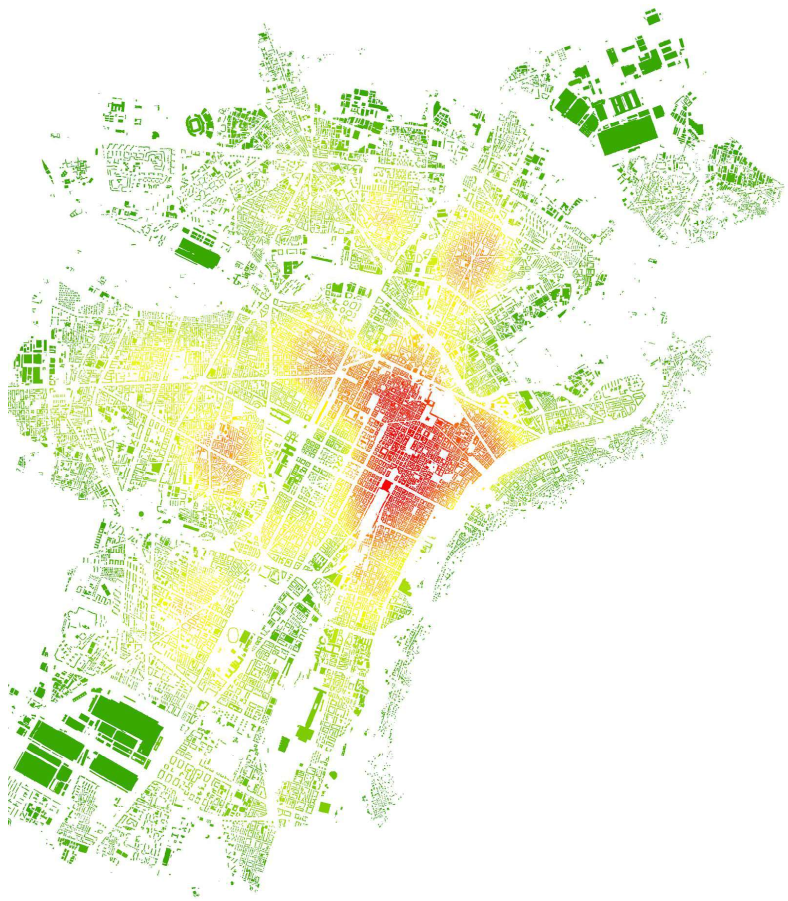


Figure 4 - The betweenness centrality for all the buildings in Turin (r=600)

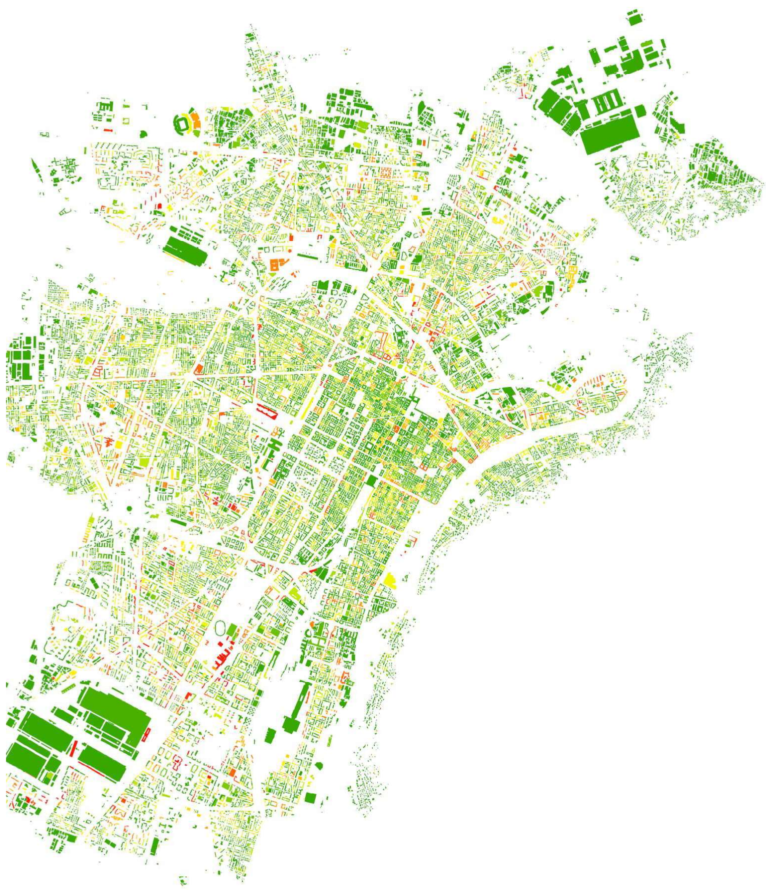
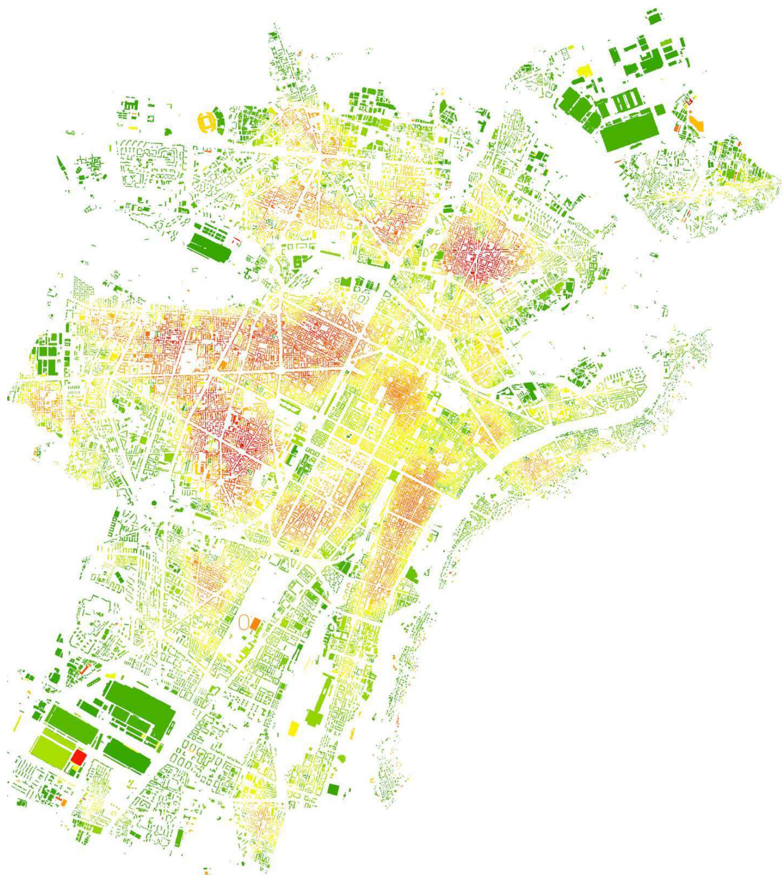


Figure 5 - The straightness centrality for all the buildings in Turin (r=600)



Chapter 2

Case study location and data

The thesis focuses on the dynamics related to the retail stores. The database this thesis deals with collects all the retail trade activities carried out locally in the city of Turin, Italy. In the database, retailers can either be: (1) companies that professionally buy products to suppliers and sell products to consumers in a fixed location or (2) companies that serve consumers with prepared food and beverage, in bars, restaurants or pubs. Actually, the database collects all the commercial licenses that have been issued in Turin from 2004 to 2019; given that a commercial license[8] is a special permit for a company to sell a merchandise category to consumers.

For the empirical tests in the thesis, however, a strong selection has been put in place: since the final aim is to describe how retailers as economic activities distribute in space, the database was cleaned out of all of that activities that do not

[8] The transposition of the Bolkestein directive introduces the freedom of establishment and the freedom to provide services to consumers assuming that new retailers can open without quotes or territorial limits, subject only to some constraint basically related to the standard of health, workers and environmental protection. In this sense, the opening and the closure of retailers can be considered as business decisions that reveal market opportunities and dynamics.

have one or more shopping windows overlooking the urban environment and, hence, contributing to its vitality.

In the database, each license is associated with many information. Among the others, licenses are associated to a merchandise category, a location, a commercial surface, an opening date and a closure date and a VAT number in an anonymous form. The merchandise categories are expressed in the form of 89 five-digit ATECO codes. In the thesis, sometimes, merchandise categories are aggregated in families given some common characteristics. Moreover, the location-based licenses can be aggregated in municipalities, census zones, in buildings depending on the interest in the specific test performed.

2.1 The economic situation in the retail sector

Apart from the specific interest on Turin, retailing activities have a clear entrepreneurial nature. Moreover, the supply side, according to some market principle, faces the preferences and the willingness to pay of consumers. In this sense it is important to describe the economic situation of the Italian companies in the sector for better analysing the dynamics in the Turin city.

The selection the thesis made of retailers (being either (1) companies that professionally buy products to suppliers and sell products to consumers in one fixed location or (2) companies that serve consumers with prepared food and beverage, in bars, restaurants or pubs) is unusual and requires to select references from many institutions (ISTAT, MISE, professional associations). Therefore, the following description provides information that are not always coherent in time span, completeness and comparability. However, the emerging picture can be considered coherent and significant.

In the following, I present an analysis of the retail sector (as defined before) in Italy. It shows the evolution of supply in time, the evolution of demand in time and the firm movements along merchandise categories and distribution forms and channels.

2.1.1 The demand side

The retail sector in Italy deals with consumption choices as well as the willingness to pay for consumers. Table present the trend in household expenditure in Italy from 2000 to 2018.

Table 5 - Households average expenditure per macro-category

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Foodstuffs	393.00 €	400.00 €	412.90 €	436.40 €	441.00 €	443.50 €	455.30 €	453.00 €	462.30 €	445.70 €
	16,9%	17,2%	17,6%	17,7%	17,3%	17,2%	17,3%	17,1%	17,5%	17,2%
Non-food items	644.70 €	652.90 €	633.00 €	664.10 €	678.10 €	676.30 €	679.00 €	666.30 €	653.10 €	634.90 €
	27,7%	28,0%	27,0%	27,0%	26,6%	26,3%	25,8%	25,2%	24,7%	24,5%
Total	**1,037.70 €**	**1,052.90 €**	**1,045.90 €**	**1,100.50 €**	**1,119.10 €**	**1,119.80 €**	**1,134.30 €**	**1,119.30 €**	**1,115.40 €**	**1,080.60 €**

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Foodstuffs	448.90 €	459.60 €	447.10 €	439.40 €	436.10 €	441.50 €	448.00 €	457.10 €	461.70 €	-
	17,2%	17,4%	17,5%	17,8%	17,5%	17,7%	17,7%	17,8%	18,0%	-
Non-food items	629.50 €	621.60 €	581.30 €	554.50 €	572.20 €	573.00 €	581.80 €	586.00 €	587.90 €	-
	24,2%	23,5%	22,8%	22,4%	23,0%	22,9%	23,0%	22,9%	22,9%	-
Total	**1,078.40 €**	**1,081.20 €**	**1,028.40 €**	**993.90 €**	**1,008.30 €**	**1,014.50 €**	**1,029.80 €**	**1,043.10 €**	**1,049.60 €**	-

2.1.2 The supply side

To extend the boarders of the analysis for analyzing the supply side, even if only with descriptive purposes, is a must, also. There, one can analyze either the stock of retailers or the flow or retailers, proposing a deepening of the entry and exit dynamics in the Italian retail sector, instead.

First of all, the Annuario Statistico Italiano in 2018 stated that at the end of 2016 the retail sector for retailers selling products to consumers, counted 495.505 companies with 1.650.655 employees (on average, 3,33 employees per company). It also stated that the food and beverage retail sector (combined with the hospitality services industry, a sample that is higher than the one of interest) counted 323.563 companies and 1.378.600 employees (on average, 4,26 employees per company).

Table x presents the total number of local units of retailers selling products and prepared food and beverage between 2010 and 2018, classified according to the kind of product they sell.

Table 6 - Number of retail activities in Italy [2010-2018][9]

Merchandise category	2018	2010[d]	Annual growth rate
Supermarkets	17779	14998	2.1%
Minimarkets, fronzen products and other stores selling grocery	71342	80677	-1.5%
Stores selling other than grocery	14857	13893	0.8%
Grocery and beverage	91875	95014	-0.4%
Tobacco	34262	30645	1.4%
Home accessories (furniture and furniture accessories)	97181	115155	-2.1%
Cultural and recreational items (books, newspapers, toys)	58855	70493	-2.2%
Clothing, shoes, clothing accessories	157999	178991	-1.5%

[9] Source: MISE, Rapporto sul sistema distributivo (2018 and others)

Medicines and cosmetics	51067	47866	0.8%
Electronics, computer appliances and sporting goods	42459	44287	-0.5%
Pets, flowers	83487	80875	0.4%
Food and drink serving, without kitchen (bar, pubs)	167677	149992	1.4%
Restaurants	204934	171787	2.2%
Total	**1093774**	**1094673**	0.0%

One can notice that, even if the number of units is stable and the annual growth rate is only weakly positive, the size of some merchandise sectors has strongly increased and the one of some others has strongly decreased.

A strong increase is measured for large-scale retailers (supermarkets and hypermarkets, large warehouses), tobacco stores, stores preparing food and beverage and, more marginally, food stores. Clothing stores, shoes stores, stores selling cultural and leisure product stores, households products retailers are greatly reduced.

2.2 The economic situation in the retail sector in Turin

However, the thesis focuses on Turin. Hence, a deep analysis of the dynamics of the retail sector in Turin is mandatory. In order to give some context about the relevance of that sector and the huge amount of jobs it provides, the following tables report the number of operating spaces and the number of employees in the city of interest.

One can conclude that the retailing sector in Turin is relevant; in 2011, it employed 42.660 workers and 13.795 spaces. Among the 13.795 spaces, 164 employed more than 20 workers each.

Table 7 - The distribution of employees and industries in the retail sector in Turin [2011]

(a) Number of employees

	Units with 0-2 employees	Units with 3-19 employees	Units with 20+ employees	**Total**
Supermarkets	47	1.351	3.819	5.217
Minimarkets, fronzen products and other stores selling grocery	352	752	-	1.104
Stores selling other than grocery	69	279	336	684
Grocery and beverage	1.351	1.328	-	2.679
Tobacco	443	554	-	997
Home accessories (furniture and furniture accessories)	1.19	1.155	690	3.035
Cultural and recreational items (books, newspapers, toys)	830	505	75	1.41
Clothing, shoes, clothing accessories	2.043	3.2	531	5.774
Medicines and cosmetics	408	2.17	49	2.627
Electronics, computer appliances and sporting goods	386	740	322	1.448
Pets, flowers	1.224	846	102	2.172
Total	**8.343**	**12.88**	**5.924**	**27.147**
Food and drink serving, without kitchen (bar, pubs)	2.238	2.695	339	5.272
Restaurants	1.943	6.843	1.455	10.241
Total	**4.181**	**9.538**	**1.794**	**15.513**

(b) Number of units

	Units with 0-2 employees	Units with 3-19 addetti	Units with 20+ addetti	**Total**
Supermarkets	40	154	58	252
Minimarkets, fronzen products and other stores selling grocery	275	130	-	405
Stores selling other than grocery	54	36	7	97
Grocery and beverage	987	306	-	1.293
Tobacco	266	151	-	417
Home accessories (furniture and furniture accessories)	923	244	9	1.176
Cultural and recreational items (books, newspapers, toys)	629	109	3	741
Clothing, shoes, clothing accessories	1.588	623	17	2.228
Medicines and cosmetics	302	383	2	687
Electronics, computer appliances and sporting goods	304	131	9	444
Pets, flowers	969	170	4	1.135
Total	**6.337**	**2.437**	**109**	**8.875**
Food and drink serving, without kitchen (bar, pubs)	1.6	617	13	2.23
Restaurants	1.427	1.221	42	2.69
Total	**3.027**	**1.838**	**55**	**4.92**

2.2.1 The demand side in Turin

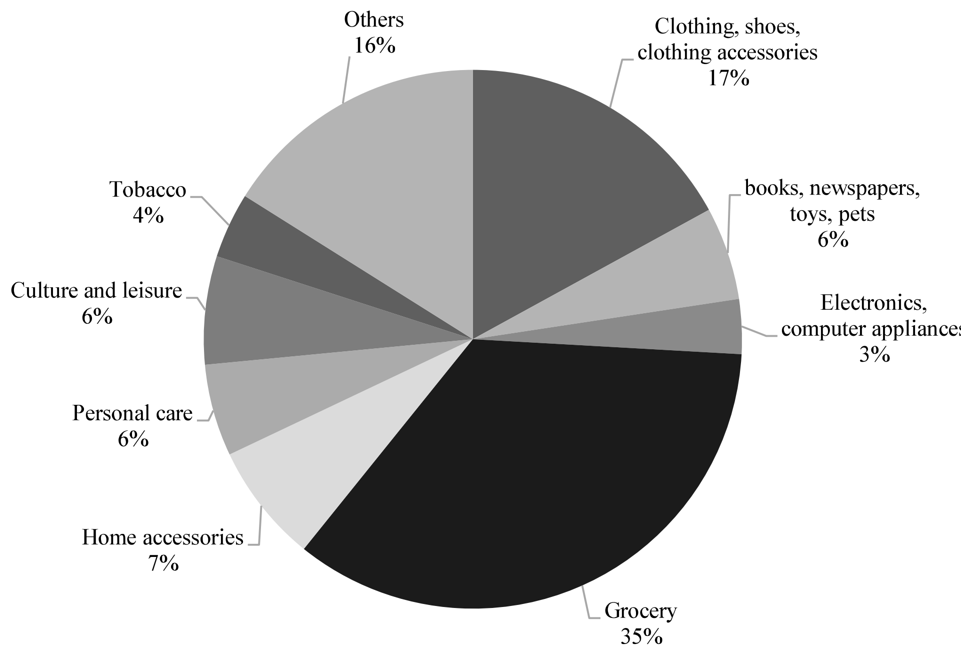
Again, consumers' willingness to pay can not be omitted since consumers' spending habits and tastes drive the entire industry success.

Some information about the distribution of consumption expenditure of residents are offered by Esri Demographic Data – MBR.

On a sample basis, expenditure data are classified both according to merchandise categories and districts.

In 2017, pro-capita consumption is equal to 8.744€. It is divided as in the Figure below.

Figure 6 - Total pro-capite expenditure per merchandise category [2017][10]



The huge amount of the total expenditure (more than one third) is used for grocery, one third for consuming clothing and clothing accessorizes and one third for consuming electronics and computer appliances.

Moreover, the total expenditure is heterogeneous among districts; the high-income districts are 1.300€ above the average (Crocetta), the ones with lower income are (up to) 1.000 below it (Falchera).

[10] Data source: Escri Demographics in Arcgis Business Analyst

Figure 7 - Total pro-capite consumption expenditure in Turin [2017] per district

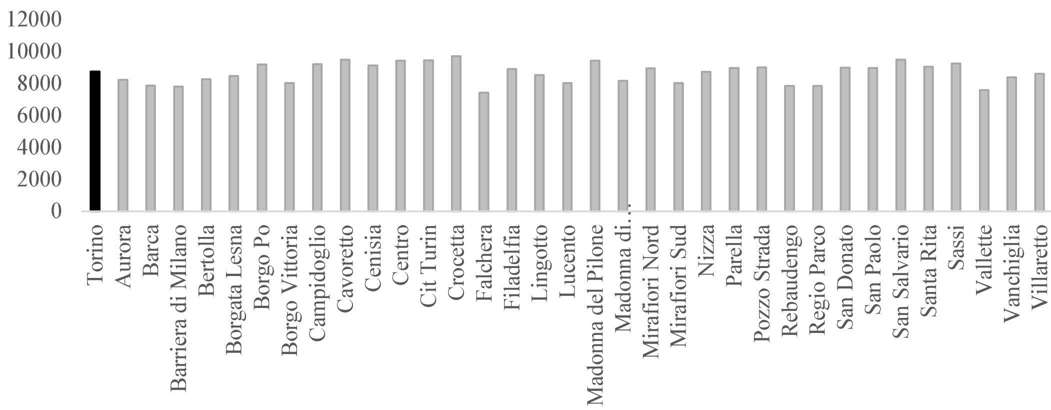
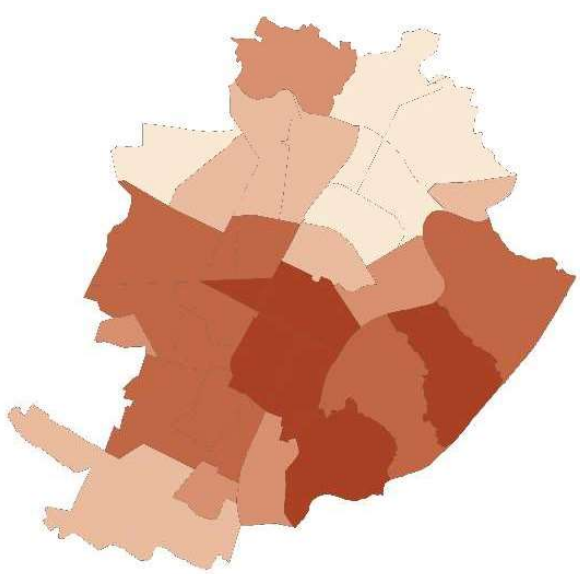


Figure 8 The distribution of total pro-capite expenditure per district, Turin [2017]



Finally, the subdivision of the total expenditure per merchandise categories in districts changes as the level of total expenditure per district changes. An example is proposed in Figures where the *clothing and clothing accessorizes* expenditure per district is mapped in comparison with the same information about the expenditure in *food* and *tobacco* products (essential goods); one can finally observe that the total consumption in essential goods is higher in lower-income districts.

To conclude, the analysis of the willingness to pay of residents and, hence, their demand for consumption goods is heterogeneously distributed in Turin, in 2017. Hence, this evidence should be reasonably reflected in an heterogeneous distribution of the supply (retail stores) in the city.

Figure 9 - Consumption ependiture of residents in Turin per district in clothing and clothing accessorizes normalized with respect to the total expenditure [2017]

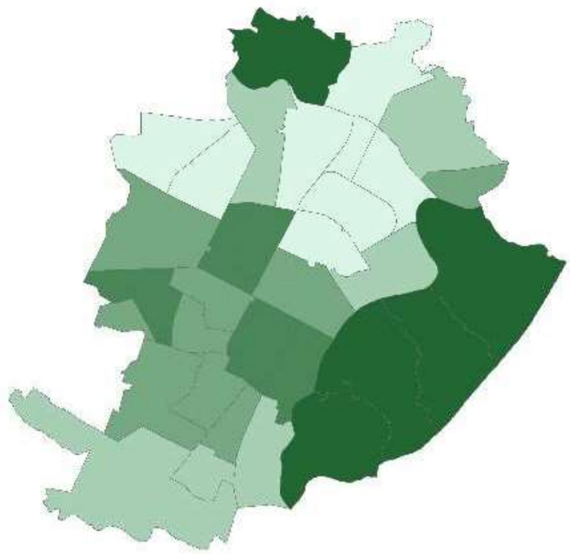
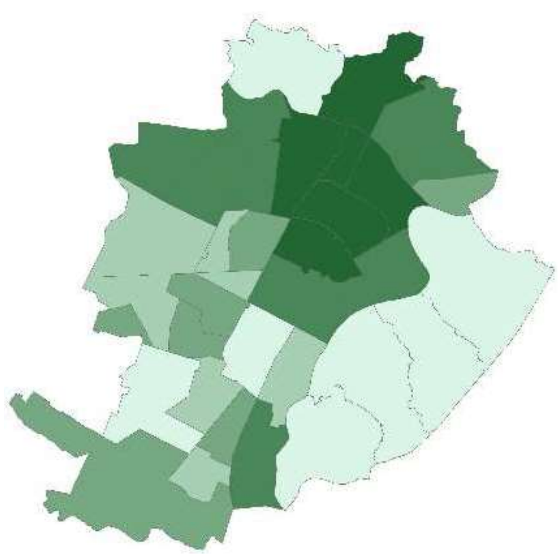


Figure 10 - Consumption ependiture of residents in Turin per district in food and tobacco normalized with respect to the total expenditure [2017]



2.2.2 The supply side in Turin

Finally, the distribution of stores offering consumption goods and prepared food and beverage is analysed. This analysis can be conducted using different lens. Indeed, the supply in the retail sector is in the hand of *industries* that offer a special assortment of products in a *store* to consumers. In order to perform their commercial activity, industries must ask to the Municipality for a *commercial license* that admit the retail store opening. It collects some information about the assortment the industry sells to consumers and its commercial space, i.e. the merchandise category and its location.

Hence, the analysis can be conducted counting either the number of commercial licenses (showing the distribution of merchandise categories in the city) or the number of commercial spaces (showing the distribution of stores in the city). Both the analysis are interesting also because the different analysis can provide matching information. Indeed, for example, in a retail store can co-exist more than one commercial license and an industry can operate with more than commercial space in the city.

In the following sub-chapters, the descriptive analysis:

- Disaggregates data according to merchandise categories and districts
- Aggregates merchandise categories in macro-categories (nine or fifteen)

Data about commercial licenses issued by the Municipality, that can be used to describe the distribution of retail activities, are collected on a 2010-2019[11] time span.

[11] Data about 2019 are collected up to October 31st.

Table 8 - Amount of commercial licenses issued in Turin per merchandise category [2010, 2018, 2019][12]

Merchandise category	2019	2018	2010	Annual growth rate 2018-2019	Annual growth rate 2010-2018
Supermarkets	286	274	177	4.4%	5.0%
Minimarkets, fronzen products and other stores selling grocery	1005	1005	794	0.0%	2.7%
Stores selling other than grocery	203	212	192	-4.2%	1.1%
Grocery and beverage	2050	2093	1848	-2.1%	1.4%
Tobacco	898	891	630	0.8%	3.9%
Home accessories (furniture and furniture accessories)	1492	1533	1509	-2.7%	0.2%
Cultural and recreational items (books, newspapers, toys)	1087	1143	1194	-4.9%	-0.5%
Clothing, shoes, clothing accessories	2478	2597	2626	-5.6%	-0.1%
Medicines and cosmetics	977	982	803	-0.5%	2.3%
Electronics, computer appliances and sporting goods	737	775	557	-4.9%	3.7%
Pets, flowers	1617	1683	1393	-3.9%	2.1%
Food and drink serving, without kitchen (bar, pubs)	2978	3036	2717	-1.9%	1.2%
Restaurants	3329	3339	1986	-0.3%	5.9%
Totale	**19137**	**19563**	**16426**	**-2.2%**	**1.5%**

Overall, the number of licenses has increased from 2010 to 2019. The only categories decreasing according to that number are clothing and clothing accessorizes and cultural items. Numbers are consistent with the national ones presented before; again, large-scale retailers (supermarkets and hypermarkets) and prepared food and beverage retailers are growing faster. The first ones increased by 62% from 2010 to 2019, the second ones by 34% in the same years. However, note

[12] Source: Municipality of Turin, data elaborated by the author

that in 2019 many merchandise category present a negative annual growth. Annual growth rates of restaurants, supermarkets and hypermarkets, tobacco stores and electronics and appliances are noteworthy; they are much higher than the others.

The absolute number of commercial licenses issued in Turin and its dynamics gives a demographic overview of the retail sector, that is the focus of the thesis. However, the introduction of spatial variables is essential to understand the density and the easiness to buy different merchandise categories for residents.

In Turin, as described in Table, the spatial availability of commercial licenses measures 185 licenses per squared km. The same density, however, changes when district changes; Villaretto has 2 licenses per squared km and Centro has 1.028 licenses per squared km. It changes according to the category, also.

In the last columns of Table 9, are presented the districts that are below the 10th percentile and above the 90th percentile in terms of density of a category. Hence, it is possible to identify the residential districts with few or very few retail activities (Bertolla, Cavoretto, Sassi, Villaretto, …) and the ones that are intense in terms of commercial activity (Centro, Cit Turin, San Salvario when looking at the overall density and, for example, Barriera di Milano if considering only some categories).

Table 9 - The density of commercial licenses [lic/sq.km] in Turin per category and district [2019]

Merchandise category	Torino [density]	Min [density]	Max [density]	10th percentile [density, districts]	90th percentile [density, districts]
Supermarkets	2.2	0	7.95	Bertolla, Borgata Lesna, Cavoretto, Villaretto	Barriera, di, Milano, Crocetta, Pozzo Strada, San Donato
Minimarkets, fronzen products and other stores selling grocery	7.71	0	29.31	Bertolla, Cavoretto, Falchera, Villaretto	Aurora, Barriera, di, Milano, Centro, San Salvario
Stores selling other than grocery	1.57	0	9.68	Bertolla, Borgo Po, Cavoretto, Lingotto, Madonna Del Pilone, Pozzo Strada, Regio Parco, Villaretto	Aurora, Barriera, di, Milano, Cenisia, Centro
Grocery and beverage	16.38	0.18	62.03	Bertolla, Falchera, Sassi, Villaretto	Campidoglio, Cenisia, Centro, San Salvario
Tobacco	6.89	0	26.69	Bertolla, Cavoretto, Sassi, Villaretto	Cenisia, Centro, Cit Turin, San Donato

Home accessories (furniture and furniture accessories)	11.56	0.18	54.18	Bertolla, Borgo Po, Sassi, Villaretto	Borgo San Paolo, Cenisia, Centro, San Salvario
Cultural and recreational items (books, newspapers, toys)	8.35	0	46.32	Barca, Bertolla, Sassi, Villaretto	Cenisia, Centro, Cit Turin, San Salvario
Clothing, shoes, clothing accessories	19.1	0	208.33	Barca, Bertolla, Sassi, Villaretto	Centro, Cit Turin, Crocetta, Santa Rita
Medicines and cosmetics	7.6	0	37.95	Bertolla, Cavoretto, Sassi, Villaretto	Centro, Crocetta, San Donato, Santa Rita
Electronics, computer appliances and sporting goods	5.66	0	19.89	Bertolla, Cavoretto, Sassi, Villaretto	Cenisia, Centro, San Donato, San Salvario
Pets, flowers	12.42	0	83.49	Bertolla, Cavoretto, Falchera, Villaretto	Aurora, Centro, San Donato, San Salvario
Food and drink serving, without kitchen (bar, pubs)	22.87	0.35	131.12	Bertolla, Cavoretto, Sassi, Villaretto	Cenisia, Centro, Crocetta, San Salvario
Restaurants	25.57	0.71	191.05	Bertolla, Falchera, Sassi, Villaretto	Cenisia, Centro, Cit Turin, San Salvario
Total	**181.72**	**1.77**	**1027.49**	**Bertolla, Cavoretto, Sassi, Villaretto**	**Cenisia, Centro, Cit Turin, San Salvario**

After having analyzed the demographic distribution of commercial licenses and their spatial distribution, one can also consider their entry and exit rates, i.e. birth and mortality rates. It is important to estimate the reconfiguration costs, to evaluate the employment effects and, more in general, to understand the intensity of entry and exit barriers.

The entry and exit rates have been measured both for each merchandise category and for each district. They have been measured as the number of incoming (outgoing) licenses in a predefined time span over the stock of licenses at the beginning of that time span.

The analysis of entry and exit rates for categories is the most interesting one.

Figure 11 - Entry and exit rates per category [2010-2019]

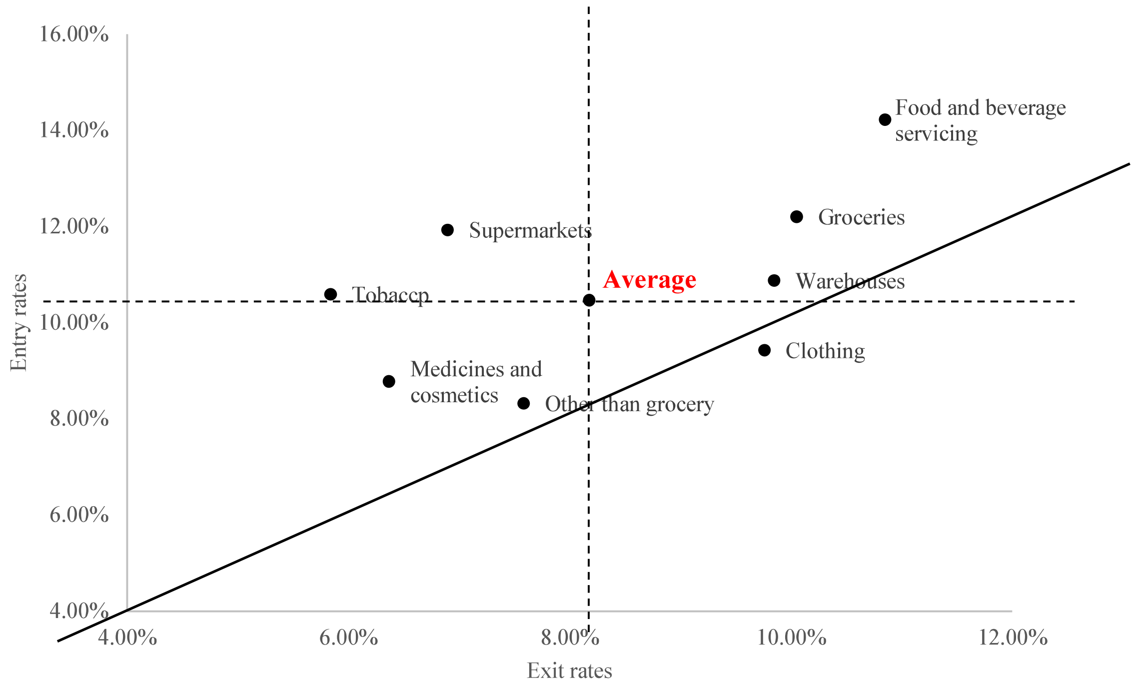


Table 10 - Entry rates per category [2010-2019][13]

	2010	2011	2012	2013	2017	2018	2019
Supermarkets	11.93%	7.30%	5.82%	16.67%	13.77%	9.40%	8.36%
Minimarkets, fronzen products and other stores selling grocery	14.96%	19.40%	14.64%	17.02%	12.16%	12.28%	8.96%
Stores selling other than grocery	12.43%	21.76%	12.62%	13.86%	5.61%	11.00%	5.63%
Grocery and beverage	11.07%	11.62%	12.17%	14.57%	10.98%	10.61%	5.53%
Tobacco	18.69%	16.35%	11.95%	15.04%	3.58%	7.56%	4.04%

[13] Tables 10 and 11 show the same information as Figure 11, the only difference is that it is disaggregated in years and the categories are fifteen instead of nine and they show an evident downward trend both for entry rates and exit rates.

Home accessories (furniture and furniture accessories)	8.37%	9.97%	6.91%	8.63%	6.57%	4.86%	4.47%
Cultural and recreational items (books, newspapers, toys)	10.19%	10.31%	8.72%	7.81%	6.26%	2.98%	3.41%
Clothing, shoes, clothing accessories	12.96%	13.13%	12.18%	9.47%	7.76%	6.37%	4.18%
Medicines and cosmetics	13.44%	9.98%	11.88%	7.83%	8.24%	7.06%	3.72%
Electronics, computer appliances and sporting goods	12.01%	15.80%	15.25%	14.86%	9.52%	8.80%	3.61%
Pets, flowers	12.06%	12.91%	13.83%	12.20%	7.43%	7.89%	4.45%
Food and drink serving, without kitchen (bar, pubs)	12.63%	14.82%	13.65%	12.93%	10.50%	11.20%	5.79%
Restaurants	19.33%	19.41%	17.79%	20.76%	16.83%	15.99%	8.68%

Table 11 - Exit rates per category [2010-2019]

	2010	2011	2012	2013	2017	2018	2019
Supermarkets	10.80%	1.12%	4.23%	9.90%	6.07%	6.02%	4.00%
Minimarkets, fronzen products and other stores selling grocery	14.32%	13.60%	13.21%	13.97%	9.77%	10.25%	8.96%
Stores selling other than grocery	8.11%	15.03%	14.56%	10.89%	7.94%	9.09%	9.86%
Grocery and beverage	10.39%	10.03%	9.49%	9.46%	8.84%	10.20%	8.75%
Tobacco	7.58%	8.73%	6.05%	8.36%	2.88%	5.50%	3.25%
Home accessories (furniture and furniture accessories)	9.10%	8.36%	8.49%	6.96%	5.80%	6.13%	6.99%
Cultural and recreational items (books, newspapers, toys)	9.76%	10.39%	8.81%	7.22%	7.02%	5.46%	8.30%
Clothing, shoes, clothing accessories	11.33%	11.60%	11.73%	11.64%	8.13%	9.09%	8.85%
Medicines and cosmetics	9.48%	7.39%	7.56%	5.98%	4.80%	6.75%	4.22%
Electronics, computer appliances and sporting goods	9.06%	9.87%	9.15%	8.95%	7.51%	6.96%	8.52%
Pets, flowers	10.16%	10.76%	10.18%	10.77%	6.22%	6.50%	8.37%
Food and drink serving, without kitchen (bar, pubs)	11.36%	11.92%	11.90%	11.52%	9.79%	9.05%	7.70%

Restaurants	11.93%	11.46%	10.99%	11.77%	11.37%	11.22%	8.98%

As already known, the *clothing and clothing accessories* category is the only one with a negative net growth rate (it is below the bisettrix in Figure). Moreover, Figure shows that the *prepared food and beverage* stores, the *warehouses*, the *food stores* and the *clothing and clothing accessories* ones are subject to a great ‘turbolence’, i.e. they are characterized by high entry rates and high exit rates. On the opposite, *medicines and cosmetics stores*, *supermarket* have low entry and exit rates.

Chapter 3

Consumers, categories, urban network[14]

Retailers choose their locations based on the objective accessibility measures and the subjective attractiveness of the potential places. It is believed that a place is more accessible for retailers if the generic routes taken by consumers often cross it. However, while previous studies have measured a single betweenness, based on random consumers flows within the city, they failed to consider that there are at least two possible consumer routes: job commutes from residential to work places and shopping trips from store to store. In this chapter, empirical simulations of both trips were used in probit regression models, to analyze commercial patterns in Turin. We first show that stores selling homogeneous products and stores selling comparable goods can differently benefit from being located in population hotspots and in commercial areas. Second, we demonstrate that daily commutes to workplaces do not benefit a retailer along the trip, as much as journeys for shopping purposes do. Indeed, the higher the number of times a potential consumer passes by a location when getting to workplaces, the lower the probability of finding a store

[14] Chapter 4 is an extended version of the paper titled "*Using betweenness metrics to investigate the geographical distribution of retailers*" written by the author and others in 2020.

in that location. By contrast, a typical building located on the route trips from store to store presents a high probability of hosting a store. Finally, we show that the benefits that a store can have when localized on the route, for both kinds of trips, depend on the kind of goods it sells.

3.1 Hypotheses

The empirical literature focusing on the location of retailers within the city introduces several metrics for capturing location attributes that affect the probability of finding a retailer, in a specific position (Guy, 1998; O'Kelly, 1983). Porta et al. (2009) and Sevtsuk (2014), in particular, use the *closeness (gravity)*, *betweenness*, and *straightness* measures. While the closeness to customers and other stores can capture the demographic and agglomeration, straightness and betweenness can capture the impact of the urban morphology and the consumer flows within the city.

Conceptually, closeness indicates how far each location is from the others. When we compute the closeness of a given location in relation to consumers' houses, we can capture whether the location is conveniently located near places where they live (demographics). When we compute the closeness of a given location to retail stores, we can measure the extent to which the location might either enjoy economies of agglomeration—making product and price comparison easier for consumers—or the presence of natural advantages of that location. While closeness can capture whether a given location is near large populations, the betweenness can capture whether people pass by a given location while moving within the city. Hence, the betweenness measure indicates the potential traffic passing by a given location, and traffic is a key driver of the attractiveness of any given location for retailers.

Sevtsuk (2014) investigates the impact of the flow of people within the city on retail locations. The betweenness index used by Sevtsuk estimates the number of times a consumer passes by a location, using the shortest paths from every building (source) to every building (destination) in the city. He demonstrates that such an index is a useful metric to investigate retail location distribution. He states that all retailers appear to value places with higher levels of passing traffic, i.e. high betweenness values.

Finally, the straightness metric assumes that shopping strategies often include unplanned purchases at more visible physical stores; places highly noticeable from the shopping path are consequently more frequented, making them more attractive as retailer locations, because the traffic passing the stores around that location could easily deviate there.

In this chapter, I analyze location attractiveness and accessibility (Krizek, 2003), using all of the above metrics. The primary contribution, however, focuses on the betweenness metric, i.e., the impact of urban morphology on the flow of people within the city. The present research argues that a betweenness index, where sources and destinations are all the buildings in a city, may not fully capture the type and the magnitude of consumer flows. Indeed, we argue that different kinds of trips might have different impacts on retail stores, thus, determining the real importance of "being in the middle" for the retail system.

Akin to Sevtsuk, we measure the flow of people surrounding a location through betweenness. Moreover, similar to Sevtsuk, due to the lack of information on actual people flow within the city, we simulate and estimate its intensity through specific locations. However, we disentangle the contribution of two kinds of people flows that, we argue, might have a different impact on the retail industry.

On the one hand, we estimate the daily commutes of people from their home to their workspaces. These trips are relatively long and might be performed with faster means of transportation (e.g. cars) and, thus, might not generate a large positive impact on consumer demands and attractiveness of the location for retailers.

On the other hand, we can simulate shopping trips from store to store; these trips are relatively shorter and typically on foot, thus making it easier for consumers to stop by and shop, increasing the attractiveness of the location for retailers.

Subsequently, for each building in the city, we will be able to calculate the betweenness for both the first (daily commutes from home to work) and second set of routes (shopping trips from store to store). A second original contribution of the work lies in the attempt to identify the differential appeal of various location attributes for retailers selling different types of goods.

The literature about consumers' shopping behaviors highlights two possible conflicting habits (Carlson and McAfee, 1984; Urbany et al., 1996). Consumers either want to minimize the time cost for searching for the goods they want to purchase, or they want to maximize the opportunity to compare goods. Moreover, they try to combine purchases through 'multipurpose shopping' that tends to increase the variety of complementary shops in a given area, and thus, the dispersion of competing retailers (Burdett and Malueg, 1981; McLafferty and Ghosh, 1986).

The prevailing effect could depend on the type of goods that are purchased. Copeland, in 1923, outlined a basic distinction between *convenience goods* and *shopping goods*. The first ones are essential consumer goods, characterized by lower product differentiations and lower price dispersions. The latter ones are differentiable according to some characteristics of the goods that the consumer typically compares before purchasing. This distinction between different types of goods is essential, since consumers tend to behave differently when shopping for convenience versus shopping goods. Thus, business retail strategies, among which the choice of location, must reflect the specific nature of the products sold (Holton, 1958).

The results of this literature suggest the following:

- Convenience goods[15] (which are standardized, have a high frequency of purchase, are low involvement goods) tend to be purchased from easily accessible stores. Indeed, consumers are expected to be relatively indifferent towards which store to visit, given their slight product differentiation. Consumers are typically familiar with these products and, as soon as they decide what convenience goods they want to buy, they will reduce the time cost spent searching. This is also because these products are, generally, relatively cheaper and purchased quite frequently. Hence, retailers who choose easily accessible locations for the consumers gain a sizable competitive advantage. These consumers will in fact tend to purchase such products from near their houses, their workplaces or along the routes they frequent most.
- Differentiated shopping goods[16] tend to be compared by consumers, who are pleased to visit many heterogeneous stores to compare products and prices before making their purchase. Consumers often compare products/stores

[15] Convenience goods includes: Grocery, Biological grocery, Funeral articles, Pet supplies, Sanitary wear, Audiovisual products, Soft drinks, Fuel, Stationery, Personal- and home-care products, Colorificio, Electrical households appliances, Electonic items, Wine, Herbalist, Drugstore, Hardware store, Flowers, Photography, Fruit and vegetables, Ice-creams, Toys, Newspapers, Informatics, Hypermarket, Bookstore, Butchery, Minimarket, Vending machine store, Bakery, Pastry Fish market, Phone center, Pizza, Perfumery, Car parts and accessories, Sexy shop, Supermarket, Tobacco shop, Telephony.

[16] Shopping goods include: Clothing, Households goods, Sportswear, Jewellery, Footwear, Furnishing accessories, Lingerie, Numismatics and philately, Precious objects, Artworks, Glasses, Fabrics, Second-hand goods.

> because they have not been able to establish exactly what they desire before the shopping trip and want to compare products and prices during the shopping trip, contrary to what happens with convenience goods. Consequently, the consumer appreciates a spatially concentrated distribution of shopping goods stores, and retailers selling shopping goods can benefit from being near other such retailers, despite the increase in competition.

Therefore,

Hypothesis 1: The spatial distribution of residence has a positive impact on the spatial distribution of stores; the higher the number of people living close to a given location, the higher the probability of noticing a store in that given location.

Hypothesis 2**:** The density of retailers around a location is positively correlated to the attractiveness of that location for all retailers. This effect can be attributed either to the opportunity that the location offers to the consumers to reduce search costs, and/or to the presence of natural advantages of the location (e.g. the proximity to source of consumers' traffic such as an underground station).

Hypothesis 3: The flow of people has a positive impact on the spatial distribution of stores; the higher the betweenness of a given location, the higher the probability of noticing a store in that given location. However, the traffic generated by shopping trips benefits the retail activities more than traffic generated by daily commutes, both because daily commutes are performed on tighter schedules and with quicker means of transportation that make stopovers less convenient. Hence, we expect the impact of the betweenness of daily commutes, with respect to the betweenness of shopping trips on retail activity, to be lower or even negative.

Hypothesis 4**:** The visibility/accessibility of a specific location tends to make that location attractive for all retailers. Notice that keeping other factors constant, almost all retailers prefer a location more accessible, with more shopping traffic passing the front door, and belonging to a shopping district. However, retailers who serve consumers who compare products tend to be planned destinations of shopping trips, while convenience goods stores benefit more from unplanned buying. Thus, we put forward the following hypothesis:

Hypothesis 5**:** The density of retailers, the traffic generated by shopping trips, and accessibility have an impact that is higher for stores selling convenience goods than for stores selling shopping goods.

3.2 Data

The analysis was carried out in the municipality of Turin, excluding the hilly area of the city.[17] All variables are measured, if not differently stated, on 12/31/2016. The area studied corresponds to 109.4 km^2, with a population of 891,916 inhabitants. Within this area, I obtained information about every active retail license (13,447);[18] among the others, their geographic coordinates and their merchandise category, which we categorized as convenience or shopping goods.[19] Second, I obtained information on every one of the 37,394 buildings in the area.[20] We learned the geographic coordinates and the surface area of each building and their prevailing usage (residential, commercial, industrial, public administration, etc.). In particular, 28,026 out of the 37,394 buildings present were exclusively or prevailing in terms of residential use.

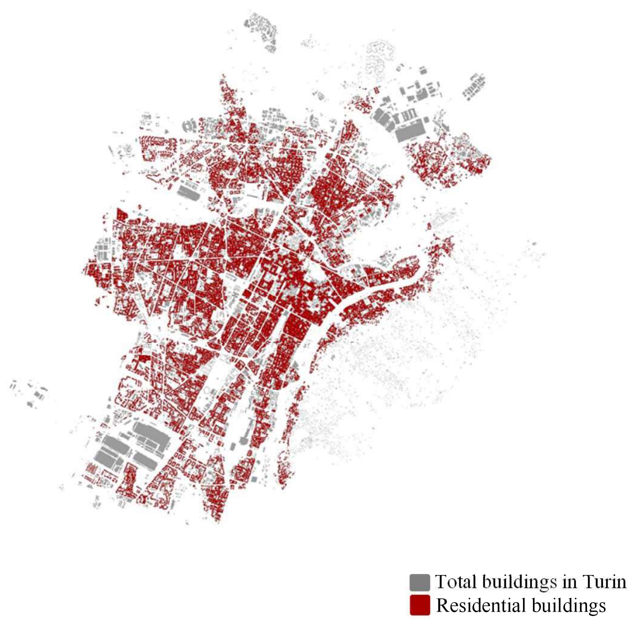
[17] The administrative borders of Turin include two areas delimited by the Po river. The hilly area is a residential area characterized by low density of inhabitants and low density of stores. This area represents 16% of the surface of the municipality of Turin, 4.8% of the population and only 0.03% of Turin's retailers.

[18] A licence is an open trade permit to sell a specific merchandise category related to a geographical address and a store size. It is provided by the municipality as soon as a fee is paid and few requirements are met, except for specific types of merchandise category (tobacco stores and pharmacies) or for very large selling surfaces that need specific administrative approval; the urban retail industry can be considered as a free entry market. The licence data set can be downloaded from the Turin Geoportal (http://geoportale.comune.torino.it/web/).

[19] All the stores have been classified either as shopping goods stores or convenience goods stores. The first are goods subject to comparison shopping (e.g.: clothing stores, clothing accessories, underwear, etc.). The latter are goods that are homogeneous in price and other product features (e.g.; tobacco, pharmacies, etc.). The complete list of the classification is available upon request.

[20] Source: Regional Technical Maps of the Laboratory of Analysis and Urban Territorial Representations (http://www.lartu.polito.it/).

Figure 12 - Map of buildings



The dependent variables we aimed to investigate in all the analyses was the probability that a residential building hosts: (i) a retailer, (ii) a convenience retailer, or (iii) a shopping retailer; these three variables capture the distribution of retailers, given the distribution of buildings in the city of Turin.[21]

[21] The aim is to understand the attractiveness of buildings as locations of retail activities. In order to exclude trivial evidence we have decided to focus the analysis on buildings with exclusive or prevailing residential use. However, this choice regards the dependent variable only: all the retail licenses (including those located in non-residential buildings) will be considered in the other variables. Notice that limiting the analysis to residential buildings also rules out from the sample all of the large selling surfaces whose location is not freely chosen by retailers (see footnote 6).

Figure 13 - Map of buildings hosting at least one retailer selling convenience goods

Figure 14 - Map of buildings hosting at least one retailer selling shopping goods

Third, the geographic layer containing demographic information for each of the 3,850 census zones of Turin[22] provided the number of residents. Within each census zone, the population was attributed to each residential building, in proportion to the surface of the building. Hence, I obtained an estimate of the population living in each building.

Finally, I obtained information on the spatial distribution of travels between home and work in Turin, to capture the daily commutes of workers.[23]

Summing up, the database presents—for every building—information about retail presence, population, morphological attributes and centrality along the flow of citizens. Therefore, it offers a very detailed representation of the city's retail ecosystem. Some descriptive statistics are presented in Tables 12.

Table 12 - Licenses, building and census zones (12/31/2016)

(a) Licenses distribution per building use

	Shopping goods licenses	**Convenience goods licenses**	**Total licenses**
Residential buildings	3,566	8,224	11,790
Commercial buildings	272	418	690
Industrial buildings	110	121	231
Other buildings	238	498	736
Total buildings	**4,186**	**9,261**	**13,447**

[22] Source: Turin Geoportal (http://geoportale.comune.torino.it/web/).

[23] 5T, the Turin Transport Agency, occasionally produces an origin-destination matrix for traffic flows of workers in Turin. Based on interviews to residents and data from traffic detectors, the matrix measures the flows of vehicles of residents who are getting to work between 166 areas of the city. For the origin-destination matrix produced by 5T in 2016, we estimate the flows of residents in area i to their work location in area j (i, j = 1 … 166) by assuming the flows of people as proportionsal to the flows of vehicles.

(b) Licenses and residents: descriptive statistics

	Min	Max	Mean	Median	Standard deviation
Residential buildings per census zone	1	142	8.71	7	8.10
Non-residential buildings per census zone	1	84	3.92	2	5.93
Residents per census zone	4	2644	231.67	184	223.02
Residents per residential building	0.13	1003	30.94	18.77	40.22
Licenses per census zone	0	76	3.74	2	5.16
Licenses per residential building	0	27	0.42	0	1.14

Note: Buildings are classified on the basis of their prevailing use. Official statistics attribute residents to census zones; we have distributed residents among buildings in proportion to the building surface.

3.3 Methodology and results

This research tries to capture the impact of demographics, mobility behaviors and spatial form on the spatial distribution of retail stores in the city of Turin. We want to estimate the probability of a building to host a retailer, a retailer selling convenience goods, or a retailer selling shopping goods, given the set of building characteristics (demographics of consumers, agglomeration of retailers, and urban form).

3.3.1 Demographic, behavioural and morphologic metrics

In this section, we illustrate the variables that we associate to each location (building) for measuring its attractiveness as a retailer location. These variables combine all the morphological, demographic and behavioral information we have, as suggested by the theoretical hypotheses advanced in the previous section; we suppose that consumers trade off search costs and good/price comparisons when shopping, by purchasing either near their home, their workplace or along their typical routes. Hence, below we present density indexes (density of the population and the retailers), betweenness indexes (betweenness of daily commutes and of shopping trips), and a straightness index. These variables will be measured for every location where we estimate the probability of finding a retailer, i.e. for every (prevailing or exclusive) residential building in Turin.

3.3.1 Demographic, behavioural and morphologic metrics

In this section, we illustrate the variables that we associate to each location (building) for measuring its attractiveness as a retailer location. These variables combine all the morphological, demographic and behavioral information we have, as suggested by the theoretical hypotheses advanced in the previous section; we suppose that consumers trade off search costs and good/price comparisons when shopping, by purchasing either near their home, their workplace or along their typical routes. Hence, below we present density indexes (density of the population and the retailers), betweenness indexes (betweenness of daily commutes and of shopping trips), and a straightness index. These variables will be measured for every location where we estimate the probability of finding a retailer, i.e. for every (prevailing or exclusive) residential building in Turin.

Population density index

The first variable that captures demography is the density of resident population DP_i (density of population around building i):

$$DP_i(r) = \sum_{d_{ij}<r} P_j$$

Where, the sum is extended to (P_j)—the population of every building j in the urban landscape—provided that the distance (d_{ji}) between the reference building i and the building j is lower than the radius r. We will use this variable to test whether it is more likely for a retailer to locate where more people live.

Retailer density index

The second density variable adds agglomeration of stores to the story: it is the density of retailers DR_i around building i:

$$DR_i(r) = \sum_{d_{ij}<r} R_j$$

Where, the sum is extended to (R_j)—the number of stores located in every building j in the urban landscape[24] – provided that the distance (d_{ji}) between the reference building i and the building j is lower than the radius *r.* We will use this variable to test whether retailers tend to flourish when located in clusters.

We must say that the *DR* variable incorporates at least two phenomena, other than agglomeration economies. On one hand, competition forces tend to push store owners to minimize or maximize the distance (Economides, 1986) from other competitors, as a function of structural characters of the market where they operate;[25] On the other hand, density also measures the natural advantages of a single location (all the advantages not captured by the other independent variables of the analyses). In this sense, DR_i can be considered as a control variable incorporating other various advantages of location *i*.

Betweenness indexes

The third piece of the study deals with morphology and flow of people. People live not only at home but also in other places, such as workplaces. The betweenness variables introduce measures of the concentration of consumers along the routes of their daily travels within the city.

In general, the betweenness indexes (B_i) attribute all the routes from a generic building *j* and generic building *k*, passing next to specific building *i* (see, Porta et al., 2009; Sevtsuk, 2014):

$$B_i = \sum_j \sum_k \alpha_{jk}^i W_{jk}$$

So, α_{jk}^i is equal to 1, if the shortest path from *j* to *k* passes through *i*, and 0 otherwise, and W_{jk} is the (possibly estimated) number of individuals taking the *j*–*k* route.

[24] As already underlined in footnote 9, the sum here is extended to every building in the city.

[25] Competition forces affect location decisions of retailers belonging to the same sector (e.g., pharmacy locations are affected by the distribution of other pharmacies). As we observe retailers' locations of large sets of retailers, in each set, competition forces are expected to impact the distribution we observe mildly. Indeed the vast majority of retailers in each set are not competing directly.

The original claim is that travelers behave differently, depending on the scope of their travel, so that different betweenness measures are expected to variously affect the probability of finding different types of retailers; in this sense, we are able to estimate the concentration of individuals along their travel for reaching their working location (BW_i), and along their shopping travels (BS_i):

$$BW_i = \sum_j \sum_k \alpha_{jk}^i W_{jk}$$

$$BS_i(r) = \sum_j \sum_{\substack{k \\ d_{jk} < r}} \alpha_{jk}^i$$

In the BW_i variable, W_{jk} is the estimate of the number of workers who live in building j and work in building k.[26] In $BS_i(r)$ we have included, with equal unitary weight, the shortest routes connecting each of j–k stores, located in different buildings—provided that the distance between j and k is smaller than r. In other words, we are attributing the same probability to each shopping trip between j and k, within a circle radius r, centered on the building i. We will use these variables to test whether retailers tend to locate along routes covered by consumers when they go to work and shop.

Straightness index

Finally, we also introduce a variable that measures the visibility of building i, a straightness measure (STR_i). Indeed, a generic consumer shopping in the city is expected to be more attracted by stores that are visible along the shopping path, than by stores that are around the corner in parallel or perpendicular streets to the one that they are walking through. This idea can be captured by the following index proposed by Porta et al. (2009):

$$STR_i(r) = \frac{1}{N(r)} \sum_{\substack{j \\ d_{ij} < r}} \frac{d_{ij}}{D_{ij}}$$

[26] Thanks to the 5T origin-destination matrix we have an estimate of the flows of workers from/to aggregate areas (166 areas, see again footnote 11). In line with the definition of the variable DR, the amount of outcoming (incoming) workers belonging to an area have been distributed among the buildings of the origin (destination) proportionally to their surface.

Where, the sum is extended to every one of the N(r) stores, whose (Euclidean) distance dij from building i is smaller than r.[27] Dij is the actual distance on the road network between building i and store j. STRi, consequently, is higher when the stores (many or few) close to the building under examination are located on a more straight connection with building i; so that we can say that the building is more visible from those stores. We will use this variable to test whether the store locations can be considered more attractive for retailers when more visible.

The two density indexes (population and retailers), the betweenness from store to store, and the straightness index are computed using a predefined radius; we are aware that changing the radius may change the impact of the index on the phenomenon we are investigating. However, the reasons why we opt for the 200 mts. radius for the density of the population, 600 mts. for the density of retailers and the betweenness from store to store, and 100 mts. for the straightness, are summarized below.

On the one hand, the density of population is introduced in the model to test whether consumers purchase, for some goods, from a store that is near their home, rather than going on a shopping trip. In our hypothesis, the concept of 'near' can be fairly approximated at 200 mts.[28]

On the other hand, the literature suggests that typical consumers' shopping trips lay in areas which we could approximate as a circle within a 600 mt. radius.[29] Therefore, we approximate the consumers' shopping behavior, which measures both for the density of retailers and the betweenness from store to store.

Finally, the straightness index aims at measuring the visibility of other stores from a particular building. 100 mts. is supposed to be a good approximation of the radius of action for the eyes of distracted consumers out shopping.

[27] Notice that for the 〖STR〗_i (r) we have considered the location of all the licenses in the database, not only the ones located in residential buildings as for the dependent variable.

[28] We have also conducted sensitivity analysis using alternative radii around 200 mt and results are not significantly changing. Sensitivity checks are available upon request.

[29] This idea follows the Handy and Niemeier (1997) gravity indexes formulation. They have demonstrated that store attractiveness decreases as travel distance increases: it is about zero when travel time by foot is higher than 10 minutes. As in the paper written by Sevtsuk (2014), 600 metres can be a good approximation of 10 minute walk and the radius can better take into account the willingness to consumers to walk among stores. However, we have also conducted sensitivity analysis using alternative radii around 600 mt. and results are not significantly changing. Sensitivity checks are available upon request.

Though we acknowledge that the variables, as often is the case in the social sciences, are proxies of the ideal measures, the next section investigates their empirical relevance when trying to explain the distribution of stores in a city, using data from Turin. The results will reveal whether these metrics provide insights regarding the spatial distribution of stores in an urban context.

3.3.2 The probit models

In Tables 12, the summary statistics of dependent (a) and independent (b) variables and a correlation matrix (c) are provided, respectively.

To test hypotheses 1–5, this work uses the probit regression models.

Table 13.a shows the outcome of a probit model, where the dependent variable is the probability that a building hosts at least one retailer. In Table 13.b, the dependent variable is the probability that a building hosts at least one convenience goods retailer. Finally, in Table 13.c, the dependent variable is the probability that a building hosts at least one shopping retailer.

For each dependent variable (all stores, convenience stores only, shopping retailers only), we consider three different sets of independent variables: *Models 1* consider only the density of the population (DP) and of the retailers (DR); *Models 2* also include the betweenness of daily commutes (BW) and the betweenness of shopping trips (BS). *Models 2* investigate the impact of different mobility patterns within the city in terms of retail activity; *Models 3*, add to *Models 2* the effect of the visibility of stores through the straightness metric.

Models 1 are testing hypotheses 1 and 2. Additionally, the comparison between *Models 1, 2* and *3* provide us with insights about the consistency of results and the marginal impact of every single set of variables. Moreover, *Models 1* provide a test for the quality of the data, since the outcomes are somehow obvious and known in the literature. Following this, the density of population and retailers are control variables in *Models 2* and *3*.

Table 13 - Summary statistics of the variables

(a) Dependent variables

	Yes	No	Total
Presence of one or more retailers in the building	5,670	22,356	28,026
Presence of one or more convenience goods retailers in the building	4,624	23,402	28,026
Presence of one or more shopping goods retailers in the building	2,429	25,597	28,026

(b) Independent variables

	Label	Min	Max	Mean	5th percentile	Median	95th percentile	Standard deviation
Density of population within a 200 mt. radius	***DP***	2	5,390	2,177.8	499.7	2,204.9	3,953.5	1,057.2
Density of retailers within a 600 mt. radius	***DR***	1	1,257	261.7	41.0	234.0	570.0	188.3
Betweenness of daily commutes[30]	***BW***	0	12.7	0.2	0.0	0.0	0.4	0.7
Betweenness of shopping trips within a 600 mt. radius[31]	***BS***	0	12.9	0.6	0.0	0.2	2.4	0.9
Straightness within a 100 mt. radius	***STR***	0.0	1.00	0.67	0.42	0.70	0.84	0.14

(c) Correlation matrix between independent variables

	DP	***DR***	***BW***	***BS***	***STR***
DP	1.00				
DR	0.45	1.00			
BW	-0.06	0.00	1.00		
BS	0.13	0.26	0.45	1.00	

[30] The BW index (see page 6) is normalized to the total number of people getting to workplaces (so that it is the share of workers of Turin passing next to the building) and, then, it is multiplied by 1,000.

[31] The BS index (see page 6) is normalized to [(N–1)(N–2)] (so that it is the probability that a generic shopping route from each couple of stores of the city will pass next to the building, considering that only the routes within a radius of 600 mt. will be actually covered) and, then, it is multiplied per 100,000.

STR	0.03	0.05	-0.01	0.07	1.00

Table 14 - The probability for a building to host a retailer. Probit models (n=28,026)

(a) All retailers

	Model 1			Model 2			Model 3		
	coeff		dydx	coeff		dydx	coeff		dydx
constant	-1.459	***		-1.539	***		-1.803	***	
	(0.023)			(0.023)			(0.053)		
DP	0.335	***	0.09	0.194	*	0.05	0.194	*	0.05
	(0.092)			(0.094)			(0.094)		
DR	0.192	***	0.05	0.160	***	0.04	0.159	***	0.04
	(0.005)			(0.005)			(0.005)		
BW				-0.111	***	-0.03	-0.106	***	-0.03
				(0.014)			(0.014)		
BS				0.309	***	0.08	0.304	***	0.08
				(0.010)			(0.010)		
STR							0.398	***	0.11
							(0.071)		
Pseudo R^2	0.073			0.109			0.110		
Log-likelihood	-13,086.95			-12,578.65			-12,562.63		
LR (χ^2)	2,053.39***			3,069.99***			3,102.03***		
Wald (χ^2)	1,919.28***			2,794.14***			2,819.17***		
Correct pred. (%)	74.08%			76.68%			76.69%		

Note: * p<0.05, ** p<0.01, *** p<0.001

(b) Convenience goods retailers

	Model 1			Model 2			Model 3		
	coeff		dydx	coeff		dydx	coeff		dydx
constant	-1.584	***		-1.661	***		-1.965	***	
	(0.0241)			(0.025)			(0.056)		
DP	0.643	***	0.15	0.518	***	0.12	0.519	***	0.12
	(0.0947)			(0.098)			(0.097)		
DR	0.160	***	0.04	0.127	***	0.03	0.126	***	0.03
	(0.005)			(0.005)			(0.005)		
BW				-0.110	***	-0.03	-0.105	***	-0.02
				(0.014)			(0.014)		
BS				0.291	***	0.07	0.287	***	0.07
				(0.0102)			(0.0102)		
STR							0.458	***	0.10
							(0.075)		
Pseudo R^2	0.060			0.095			0.097		
Log-likelihood	-11,799.92			-11,358.69			-11,339.79		
LR (χ^2)	1,503.17***			2,385.62***			2,423.42***		
Wald (χ^2)	1,450.01***			2,251.12***			2,279.79***		

Correct pred. (%)	77.15%	79.48%	79.50%

Note: * p<0.05, ** p<0.01, *** p<0.001

(c) Shopping goods retailers

	Model 1			Model 2			Model 3		
	coeff		dydx	coeff		dydx	coeff		dydx
constant	-1.901	***		-1.954	***		-2.073	***	
	(0.030)			(0.030)			(0.067)		
DP	-0.182		-0.03	-0.281	*	-0.04	-0.283	*	-0.04
	(0.112)			(0.115)			(0.115)		
DR	0.189	***	0.03	0.166	***	0.02	0.166	***	0.02
	(0.005)			(0.005)			(0.005)		
BW				-0.069	***	-0.01	-0.067	***	-0.01
				(0.017)			(0.017)		
BS				0.202	***	0.03	0.200	***	0.03
				(0.011)			(0.011)		
STR							0.179	*	0.02
							(0.090)		
Pseudo R^2	0.086			0.106			0.106		
Log-likelihood	-7,553.14			-7,386.35			-7,384.34		
LR (χ^2)	1,415.82***			1,749.41***			1,753.41***		
Wald (χ^2)	1,386.72***			1,695.78***			1,698.90***		
Correct pred. (%)	87.40%			87.90%			87.90%		

Note: * p<0.05, ** p<0.01, *** p<0.001

Before we use Tables 14.a, 14.b and 14.c to test hypotheses 1–5, the robustness and significance of the models deserve a comment. All models are very significant, according to both likelihood and the Wald tests. Furthermore, when we compare the coefficients of *Models 1, 2* and *3*, in every single table, they are (in most cases) very significant. This is due to the large amount of observations in the database (n = 28,026). Moreover, coefficients are stable across Models 1, 2 and 3, but they are different across the probit models in Tables 2.a, 2.b and 2.c. Indeed, the dependent variables change with the tables. Thus, the models are statistically significant and stable. Moreover, the coefficients of some variables significantly differ across tables; the divisions we made in terms of retailers, retailers selling convenience goods and retailers selling shopping goods, generate different results and insights.

In the rest of the section, the results will be described according to the list of hypotheses from section 2. When not differently specified, we will refer to Models 2. The coefficients of the models are not elasticities, hence, marginal effects at the mean (dydx measures in the tables) will also be discussed. Moreover, along with the discussion, we will also provide (in relevant cases) the partial effects of the discrete variables (50th percentile–95th percentile).

Hypothesis 1: Table 14.a confirms the hypothesis that population density supports retail activity; the population has a positive impact on the probability that a building hosts one or more retailers. Indeed, the impact is positive and significant in all three models. The result holds for convenience retailers as well (Table 14.b), and the impact of the resident population is even larger than for the average retailer. As an additional measure of the effect, an average building (i.e., a building with median values of the independent variables) but characterized by high values of DP (95th percentile), has a 0.84% higher chance of hosting at least one retailer, and a 2% higher chance of hosting one or more retailers selling convenience goods. This is consistent with the fact that for convenience goods, consumers should pay more attention to the search costs in their shopping choices.

Interestingly, the impact of population on the probability of finding a shopping goods retailer in a building is negative, and once we account for the consumers' trips in the city (*Models 2* and *3*—Table 14.c), it is statistically significant. Indeed, shopping goods retailers tend to concentrate in areas with relatively low population density. These (sets of) stores are typically the destination of shopping trips, and consumers might be willing to incur higher search costs and move considerable distances (on an urban scale).

Hypothesis 2: All the models confirm the hypothesis that retail density supports retail activity. Notice that this outcome (positive coefficients of *DR*) is very significant and stable in all three models. This might be due to strong agglomeration effects; they seem to prevail over possible competition forces that lead retailers to locate far from each other. Moreover, consider that the wide set of product categories in the database is such that competition occurs within small subsets of competing retailers (e.g. competition among hardware stores[32]) and thus, shall be rather weak in our framework. On the other hand, stores selling complementary merchandise categories (e.g., benefit of an apparel store, for a shoe store) or

[32] Competition strategies of firms in single product categories when they decide to locate in the geographic space could be actually very complicated (Huang and Levinson, 2011). Baum and Haveman (1997), for example, show that location decisions in the Manhattan hotel industry determine a distribution where hotels of similar price tend to agglomerate in order to avoid the hazards of localized price competition, while competition pushes hotels far apart from hotels of similar size. In our dataset that includes tens of different product categories, we expect single industry competition effects to be rather weak and agglormeration or natural advantage forces to prevail.

merchandise categories subject to shopping comparison (e.g., benefits of a store selling clothes to another selling clothes) might be generating proximity externalities. This makes agglomeration forces rather strong. However, we shall mention that this outcome can also be consistent with a different explanation—retailers might concentrate in locations with some sort of natural advantage (e.g., railway stations generating large traffic of daily commuters). Indeed, stores might tend to be located near sources of competitive advantage, such as traffic generators. The data does not provide evidence to support one explanation over the other.

Hypothesis 3: The evidence provided shows that the betweenness of daily commutes and the betweenness of shopping trips have an opposite effect. In every model, the coefficients of *BW* are negative and significant, while the coefficients of *BS* are positive and significant. This confirms the basic hypothesis behind this work—while previous literature (Sevtsuk, 2014) has shown that the city form leads people flow, and these have an impact on the distribution of retail activity, this empirical test shows that the contribution of such flows shall be disentangled, since their nature produces very different impacts. From Table 14.a, on the one hand, an average building characterized by high values of *BW* has a 3.77% lower chance of hosting a retailer, than a location with all the values of the independent variables at the median. On the other hand, the average building characterized by high values of *BS* is 21.10% more likely to host a retailer, than a location with all the values of the independent variables at the median. Tables 14.b and 14.c show similar results for convenience and shopping goods.

Hypothesis 4: *Models 3*—Table 14.a shows that the straightness, i.e. the visibility of stores, has a positive and statistically significant effect on the likelihood to find at least one retailer in a given residential building. This confirms that the urban form has a relevant impact on the spatial distribution of retail activities. The models confirm this hypothesis, both for convenience stores and for stores selling shopping goods.

Hypothesis 5: All the results we have illustrated so far, show that the density of retailers, the traffic of consumers during their shopping activities, and the visibility of locations support retail activity, both for convenience goods and shopping goods. However, the impact of these variables proves to be much more relevant in magnitude for convenience stores, than for stores selling shopping goods. This is evident by comparing the marginal effects at the mean (dydx measures) in Table 14.b and 14.c.

As for the partial effect of the discrete variables,

- An average building characterized by high values of *DR*, has a 9.83% higher chance of hosting one or more convenience retailers, and a 9.09% higher chance of hosting one or more shopping retailers, as compared to a location with all the values of the independent variables at the median.
- An average building characterized by high values of *BS*, has a 17.75% higher chance of hosting one or more convenience retailers, and a 7.46% higher chance of hosting one or more shopping retailers, as compared to a location with all the values of the independent variables at the median.
- An average building characterized by high values of *STR*, has a 1.68% higher chance of hosting one or more convenience retailers, and a 0.38% higher chance of hosting one or more shopping retailers, as compared to a location with all the values of the independent variables at the median.

Interestingly, many retail models focus on the distribution of the population (i.e., where people live); in fact, the empirical test shows that the distribution of retail stores matter more than the distribution of residents in the attractiveness of locations for retailers. *Model 3*—Table 14.a shows that the marginal effect at the mean for *DP* is 0.05, while it is 0.09 for *DR*.

Moreover, the impact of the population is even lower than the impact of the ease of access, as measured by the betweenness of shopping trips and straightness; a retailer has a 21.10% increase in probability, when the *BS* passes from median value to the 95th percentile, and a 1.68% increase in probability, when the *STR* increases. In other words (at the least at the urban scale), the distance from where people live seem to matter less than the patterns of trips within the city, and the ease with which the location can be accessed. Similar conclusions hold in the specific case of convenience goods and shopping goods; however, the density of population has a positive and large impact in the case of convenience goods, while in the case of shopping goods, the impact is small and negative. In Table 14.b and Table 14.c, a building that has the 95th percentile value of *DP*, is 2% more likely and 0.6% less likely to host a convenience goods store and a shopping goods one, respectively, as compared to a building that has a median value of *DP* (keeping, as usual, all the other variables at the median).

Hence, the shape of the urban form (straightness), the dynamic patterns of consumer movements within the city (betweenness of shopping trips), and the spatial distribution of retailers matter even more than a static perspective on where potential consumers live.

3.4 Discussion

This empirical test contributes to the stream of research on spatial analyses about urban retail. Previous researches showed that stores tend to concentrate where people live (e.g. Seim, 2006; Waldfogel, 2008) while Sevtsuk (2014) proved that, not only where people live but also the flow of people in the city network tends to drive the spatial distribution of retailers. It investigates very detailed data obtained about the city of Turin (28,026 buildings) and shows, at least at the urban scale, the density of other stores, the flow of commuters and the flow of consumers from store to store drive the density of stores, more than the distribution of residents.

Thus, planning often tends to investigate ratios between residents and stores, the chapter suggests that the internal dynamics of the retail industry, and the flow of people, rather than just the static distribution of where people live, shall be considered when planning a retail space. Moreover, these findings might be of interest for retailers. Indeed, the results demonstrate the features that make a location attractive and provide the metrics to measure them.

Moreover, the chapter shows that there are different types of flows of people in the city, and their contribution shall be disentangled. Indeed, while the contribution of the daily commutes in the city is negative, the contribution of shop to shop trips is positive, significant and large. Again, this fundamental understanding of what helps retail activities thrive can contribute not only to the knowledge but also to the planning and location choices of retail investors.

In the chapter, these metrics on the flow of people in the city are proxies of actual trips, estimated from publicly available data. Though the significance of the metrics and estimates are proved by the robustness of the results, the relevance of the results deserve further investigation and the use of actual traffic data—collected from cell-phone GPS records—which might provide additional opportunities for further research. Such data may pave the way for various, more detailed analyses, such as the impacts of hourly people flow, sources of natural advantages (e.g. sources of consumers, such as underground stations), new large store openings on traffic in the area, and more.

This work also investigates the impact of the urban form and proves that straightness matters, in other words, the visibility of one store from other stores increases the probability of a building hosting a store. Finally, it also shows that the above phenomena vary for different kinds of stores. Namely, stores selling convenience goods differ from stores selling shopping goods. Interestingly, the density of population increases the probability of a building hosting a store (as suggested by literature). However, the density of population lowers the probability of a building hosting a store selling shopping goods. Indeed, these latter stores are destinations of ad hoc shopping trips. This result is consistent with another observation: stores selling shopping goods prevail in shopping malls that tend to be located outside of cities, in areas where the population density is relatively low, but the concentration of stores is high. This result has potential implications for urban planners. Indeed, it seems to suggest that large concentrations of stores selling shopping goods can be created, regardless of the concentration of residents-consumers, since consumers are willing to travel both to suburban malls and concentrations of stores in the city.

Chapter 4

Consumers, categories and store size

More than ever before, traditional urban retail systems are endangered. Physical stores throughout the city are facing a slow but steady transformative process, leading to the closure of physical stores throughout the city. If the presence of empty spaces within the urban tissue is an issue by itself, the closure of Mom-and-pop retailers[33] is even worse. Indeed, they play a vital role on the liveability of a city by contributing to the vitality and the attractiveness of the neighbourhood. Furthermore, the closure of retailers negatively interferes with the economic metabolism of the city and, thus, with its viability. The chapter aims at investigating whether one of the causes of the closure of retailers can be the opening of supermarkets. Indeed, by offering a wide range of products and lower prices, they can attract consumers that were used to frequent small stores before the new opening. Empirical tests have been developed around a wide dynamic dataset of commercial licenses in Turin in order to assess the intensity of the negative spill over that opening a supermarket may generate on nearby retailers and to estimate how infectious it is; I finally measure the risk for a small store to close if a

[33] In general, independent small stores. In the thesis, with a reference to chapter 1, Mom-and-pop stores are small (sqm. <= 250) specialty stores and small (sqm. >= 250) convenience stores.

supermarket opens in the neighbourhood and I state that the risk changes if the merchandise category sold by the small retailer changes.

4.1 Hypotheses

Retailing is a local economic activity that should deal with the classical demand-supply mechanisms. On the demand side, population distribution, customer preferences, building accessibility to consumers drive the presence of a retailer as stated in the third chapter of the thesis. On the supply side, spatial competition dynamics, prices, product quality and other product features are considered (Aguirregabiria and Vicentini, 2016).

Spatial competition, first, arises between same-format stores; as Basker, Aguirregabiria and Suzuki wrote in 2016 *"the distance of a store to potential consumers, wholesalers and competitors can have substantial effects on demand and costs, and consequently on prices, quantity, profits and consumer welfare".*

Indeed, retailers should deal with three key decisions beyond price: they should decide the product to sell, the distribution channel (brick and mortar retailer or internet retailer, firstly; large-scale retailer or brick-and-mortar store, secondly) and the location to open (Baum and Haveman, 1997).

Traditional retailers, they desire to be located in the best place selling the best product. The place can either be the best one because of its relative position with respect to consumers or its relative position with respect to competitors (Fox et al., 2007); they can decide to locate as near as possible either to client or to stores of the same category or to store of other categories or to traffic generators (Basker et al., 2016).

Moreover, the selection of the best location cannot prescind from the product the store sells.

The merchandise category in the present work are divided into two groups, as presented in the previous chapter. With reference to the usual classification of merchandise categories made by Holton (1958):

- stores selling convenience goods, slightly differentiated, tend to be located far from each other and as near as possible to consumers. Indeed, consumers are familiar with those products (that are standard in price and other products' characteristics) and they aim at reducing the time cost spent searching. Hence, the competition that arises from being located one near the other is cannibalizing demand instead of producing shopping externalities.

- stores selling shopping goods, highly differentiated tend to be located near stores selling the same merchandise category given that the good they sell is never equal to the one sold by competitor; it is rather similar. Hence, being located in clusters produces shopping externalities (outweighing the negative effects of competition), mainly generated by consumers' trip-chaining behaviour (Koster et al., 2019) and it increases the potential catchment area of the store itself. Indeed, a shopping cluster increases the probability for a consumer to buy the good that meet his needs (Sevtsuk, 2010) and, hence, its willingness to travel to that cluster.

Second, spatial competition should also arise between distribution channels; traditional retailers, already competing in the physical market with stores belonging to the same distribution channel, should also compete with retailers who decide to play with another one distribution channel.

Among the others, local traditional stores, namely brick-and-mortar retailers, after having chosen the product to sell and the place where to locate, of course compete with other brick and mortar stores (being either small stores or large retailers as supermarkets) as explained before but also with internet retailers.

The empirical literature focusing on the competition between traditional retailers and e-tailers is wide. The difference between the two distribution channel is evident. Traditional retailers have physical contact points (Enders and Jelassi, 2000), who conclude their purchase usually having touched and tried the product. Online retailers have a larger catchment area, larger product assortment and customers can comfortably buy products from their home, at any hour. Moreover, online consumers can access to prices that are typically 9-16% (Brynjolfsson and Smith, 2000) lower than prices in traditional retailers.

This kind of competition, finally, has two main impacts: the first is on urban commercial pattern (Miller et al., 1998), the second is on prices and product assortments. On one side, retailers better survive in to more dense areas in terms of population, traffic generators and other stores (Guo and Lai, 2017). Prices in traditional retailing, on the other side, typically decrease taking the price from the online market leader (Chevalier and Goolsbee, 2003; Goldmanis et al., 2010) if the product sold by retailers and e-tailers is the same. When possible, instead, some supply-side shifts in the brick-and-mortar store can reduce the competition based on prices (Clay et al., 2002; Prince, 2007).

Finally, the academic research focusing on the impact of mass merchandisers on small stores comes mainly from the United States. There, researchers analyse the impact of the growth of Wal-Mart's superstores on local retailers' sales, on wages and on employment (Artz and Stone, 2006; Basker, 2005).

Even if retail employment typically increases in the year after the entry of a new supercenter, the years after local retailers exit the market and the gain disappears (Basker, 2005). However, the negative spill over decreases both with distance from the superstore and with the differences in the assortment of the local stores, compared to the large-scale one (Haltiwanger et al., 2010).

The negative spill over based on retail employment analysis, is confirmed when analysing small retailers' sales and survival; indeed, there are some evidences that prices in supermarkets are lower (Hausman and Leibtag, 2007). Moreover, consumers, by "cherry-picking"[34] (Talukdar et al., 2010) the products can save switching costs. Hence, retailers selling the same products as supermarket can suffer its presence in the neighbourhood, given that they have, on average, higher prices and a narrower assortment of products (Irwin and Clark, 2006).

The chapter aims at investigating the cross-competition between small independent stores and large stores (supermarkets, in this case). Of course, small stores and supermarkets offer a different shopping experience. In principles, the first typically are full-services, salesmen are specialized, products are refined and top-quality, the assortment is limited and prices are higher; the second are almost self-services, salesmen are not specialized, products are mass produced, the assortment is wider and prices are lower. Indeed, supermarkets face scale economies: for example, setup costs decreases, the economy of mass reserves exists, workers are specialized in a single task and fixed costs are shared among higher volumes.

Hence, they serve consumers with different needs, at least in theory, even if some of the differences between the two store format are flattering over time. For

[34] The reference is to the cherry-picking behaviour by buyers who, within a single store visit, pick the best products at the best prices (Fox and Hoch, 2005). The higher the assortment of the store, the higher is the possibility for a consumer to fill the basket of desired products with the best ones, without incurring in the cost of switching from a store selling one merchandise category to another one selling another merchandise category.

example, the product quality in the supermarkets is now comparable to the one of a small retailer.

However, it is possible for small retailer to suffer the presence of a supermarket in the neighbourhood for the reasons explained in the three general hypotheses below. These hypotheses will be tested using data about Turin, Italy; the analysis of the state of the art suggests that could be interesting to investigate the competition dynamics between local retailers and the large ones, indeed.

Therefore:

Hypothesis 1: small stores can either be negatively or positively impacted by the new opening of a supermarket in its neighbourhood[35] (Arnold et al., 1998). In the first case, supermarkets' consumers can find better deals and can cherry-picking the desired products (Fox et al., 2004). In the second case, small store can benefit from the increase in consumers' traffic given the improved attractiveness of the neighbourhood after supermarket new opening (Dennis et al., 2002).

Hypothesis 2: Hence, if small stores sell products that are complementary with respect to the ones sold in the supermarkets, consumers cannot compare prices and cannot fill the basket with the desired product that only the local store sells. So, stores selling complementary goods can better survive in the surrounding of a supermarket than stores selling goods that are substitutes for the ones sold in the large-scale retailers.

Hypothesis 3: Moreover, the impact on a new opening of a supermarket on small retailers decreases over time. Indeed, if a small store can overcome the first year of competition, then it can perform some supply-side shifts moving to a store-specific assortment. So, the competition based on prices is reduced (Clay et al., 2002; Prince, 2007).

[35] The neighborhood of store is defined as an area of 600 mts. radius. This idea follows the Handy and Niemeier (1997) gravity indexes formulation. They have demonstrated that store attractiveness decreases as travel distance increases: it is about zero when travel time by foot is higher than 10 minutes. As in the paper written by Sevtsuk (2014), 600 meters can be a good approximation of 10 minute walk and the radius can better take into account the willingness to consumers to walk among stores. However, we have also conducted sensitivity analysis using alternative radii around 600 mt. and results are not significantly changing. Sensitivity checks are available upon request.

5.2 Data

The analysis is based on an elaboration of two databases about the Municipality of Turin.

First, data about commercial licenses issued in the Municipality of Turin from January 2005 to December 2019 have been collected by the same database used in the previous analysis. There, two kind of data have been grossed up:

- Data about the large retailers selling mainly groceries (hereinafter, supermarkets): 129 supermarkets have been opened from 2005 to 2019;
- Data about the small retailers in the surrounding of the supermarkets selling groceries; furnishing and furnishing accessories; books, newspapers, toys; clothing and clothing accessories; drugs and perfumes; pets and gardening; sanitary and orthopaedics articles.

The key information regarding the large retailers and the small stores were three: (1) the geographic coordinates of the licenses; (2) the opening date of the licenses; (3) the closing date of the licenses. The table below shows how the key information have been used for the two objects of the analysis: supermarkets and small stores.

Table 15 - The key information for the key objects

	Supermarket	**Small stores**
Geographic coordinates	Location of the supermarket, in order to measure the distance between the newly opened supermarket and all the brick and mortar stores that were working	Location of the brick and mortar store to measure whether it has been potentially influenced by the newly opened supermarket (i.e. the distance to the newly opened supermarket is smaller than 600 mt.[36]) or not (i.e. the distance to the newly opened supermarket is bigger than 600 mt.)

[36] See footnote 28.

Opening date	Moment when the potential impact of the supermarket begins; in this date, one can measure the stock of retailers potentially influenced by the new opening (i.e. brick and mortar retailers working in that date, located at a distance that is smaller than 600 mt.)	Day when the brick and mortar store has opened. If it is before the opening date of the supermarket and their distance is smaller than 600 mt., it is a potential victim of the supermarket.
Closure date	Moment when the supermarket has closed and its potential influence on brick and mortar stores ceases to exist.	Moment when the brick and mortar store closed. If it is close to the opening date of the supermarket and the distance between the two is smaller than 600 mt., it could be that the new opening has impacted on its existence.

Second, the same geographic layer of the previous analysis containing the number of residents within each census zone, has been useful to determine the number of people living in the surrounding of each supermarket.

The two databases have been put together in the following four step process, with the aim to describe the variables that may drive the closure of a small store.

1. For each supermarket new opening between 2005 and 2019, I know the squared commercial meters and the list of small stores (independent commercial licenses) that were working in that date in a radius that is smaller than 600 mts from the supermarket;
2. For each license working in the surrounding of the new opening, I count the number of months from the opening date of the license to the closure date of the license (in the econometric model, $duration$)
3. For each license working in the surrounding of the new opening, I know if the license closed (in the model specification, $event$), closed in a year from the new opening (in the econometric model, $event1$), closed in two years from the new opening ($event2$), closed in three years from the new opening ($event3$), closed in four years from the new opening ($event4$) or closed in 5 years from the new opening ($event5$).

4. For each license working in the surrounding of the new opening, I count the number of residents in the neighborhood (in the model specification, DP) and the number of stores in the neighborhood (in the model specification, DR) in the year when it closes.

5.3 Methodology and results

There are some evidences that the hypotheses are fair. On average, the 129 supermarkets newly opened between 2005 and 2019 have 73.54 small stores in the surrounding, on average. In five years from the new opening, instead, the average number is reduced by 30%.

Table 16 - Average number of brick and mortar stores around a supermarket in 5 years from its opening

when they open	1 year after the new opening	2 year after the new opening	3 year after the new opening	4 year after the new opening	5 year after the new opening
73.54	**69.52**	**65.75**	**62.68**	**59.62**	**56.61**

Of course, there can be many reasons for that. It can be attributable to the general economic cycles, to the advent of the internet and to some shifts in customers' tastes; the econometric models presented below are an attempt to disentangle the only supermarket negative spillover from the other conjunctural causes for the death of small stores.

However, it is true that small stores die more easily in the surrounding of a new supermarket opening in the surrounding (see chapter…); they annually decrease by 2.3% in the five years after the new opening.

Moreover, the 18% of the closures happens in the first year from the new opening; the 50% happens from 2 to 4 years after the new opening and the remaining 32% happens from 4 to 5 years after the new opening. Hence, time is an interesting variable to be analyzed.

Table 17 - Percentage of closures in 5 years from the supermarket opening

when they open	1 year after the new opening	2 year after the new opening	3 year after the new opening	4 year after the new opening	5 year after the new opening
18.02%	**16.76%**	**15.04%**	**17.27%**	**32.91%**	

Indeed, a Cox proportional hazard model has been used to explain if the $duration$ of a small store is influenced by the opening of a new supermarket in the surrounding; we are interested in explaining the time it takes for a store to close, given a new supermarket opening.

The duration variable used in the model is the number of months that passes by the opening of a supermarket and the closure of the surrounding stores. The closure of the store in the set of the ones that are working when the supermarket opens, in a radius smaller than 600 mts. between the two, is the event dummy variable. Hence, the basic model presents the baseline hazard function (i.e. he risk for a store to close after the new opening for brick and mortar stores) with $x_i = 0$ and $e^{x'_i\beta}$ adjusts for the set of stores characteristics.

$$\lambda\,(t, x_i) = \lambda_0(t) \cdot e^{x'_i\beta}$$

4.3.1 The set of stores characteristics (x_i)

In order to describe the risk for a store to close given a new supermarket opening, we describe the store using a set of characteristics that we suppose may drive its duration, in order to get rid from those external effects.

The first variables is the key one; it counts for the sq. meters of supermarkets that opens in the surrounding of the store. The second are two variables that combine some demographic and behavioral information we have, as suggested by the theoretical hypotheses advanced in the previous chapter. Third, two dummy variables in order to get rid of any fixed effect.

The key dependent variable

First, small stores already working in a radius that is smaller than 600 mts. from a new supermarket opening in time t are selected. Then, the key independent variable, $sq\ mts\ super_i(t)$, counts for the sq. meters of the supermarket just opened. It has a simple meaning; it aims at describing whether the duration of the selected small stores is influenced by the newly opened supermarket. Moreover, it implicitly considers that the hazard for a store to close given a new supermarket opening in an area with a radius that is smaller than 600 mts, increases if the commercial meters of the supermarkets increases too.

Other dependent variables

The first variable that captures demography is the density of resident population DP_i (density of population around store *i*):

$$DP_i(r \leq 600) = \sum_{d_{ij}<r} P_j$$

Where, the sum is extended to (P_j)—the population of every building *j* in the urban landscape—provided that the distance (d_{ji}) between the store i and the building j is lower than the radius *r*. We will use this variable to test whether it is more likely for a retailer to survive where more people live.

The second density variable adds agglomeration of stores to the story: it is the density of retailers DR_i around the store *i*:

$$DR_i(r < 600) = \sum_{d_{ij}<r} R_j$$

Where, the sum is extended to (R_j)—the number of stores *j* – provided that the distance (d_{ji}) between the reference store i and the retailer j is lower than 600 mts. We will use this variable to test whether retailers tend to flourish when located in clusters.

Finally, we also include two dummy variables. The first ($year\_clos$), aims at capturing time fixed effects. It associates to each selected brick and mortar store, the closing year. The second ($district$), aims at capturing some location fixed effects. It associates to each selected store, the district.

4.3.2 The Cox proportional hazard models

Hence, the general proportional hazard model (*model1*) is:

$$\lambda\,(t, x_i) = \lambda_0(t) \cdot e^{sq\ mts\ super_i \beta_1 +\ DP_{i,r\leq 600}\beta_2 +\ DR_{i,r\leq 600}\beta_3 + year\_clos_i \beta_4 + district_i \beta_5}$$

Then, five other variables have been used and five Cox proportional hazard models have been defined:

- the $event1$ dummy variable assumes value 0 if the small store in the surrounding of the new opening closed in one year from the new opening (*model2*);

$$\lambda_{1\,year}\,(t, x_i) = \lambda_0(t) \cdot e^{sq\ mts\ super_i\beta_1 + DP_{i,r\leq 600}\beta_2 + DR_{i,r\leq 60}\ \beta_3 + year\_clos_i\beta_4 + distric\ _i\beta_5}$$

- the $event2$ dummy variable assumes value 0 if the small store in the surrounding of the new opening closed in two years from the new opening (*model3*)

$$\lambda_{2\,years}\,(t, x_i) = \lambda_0(t) \cdot e^{sq\ mts\ super_i\beta_1 + DP_{i,r\leq 600}\beta_2 + DR_{i,r\leq 600}\beta_3 + yea\ \_clos_i\beta_4 + distric\ _i\beta_5}$$

- the $event3$ dummy variable assumes value 0 if the small store in the surrounding of the new opening closed in three years from the new opening (*model4*)

$$\lambda_{3\,years}\,(t, x_i) = \lambda_0(t) \cdot e^{sq\ mts\ super_i\beta_1 + DP_{i,r\leq 600}\beta_2 + DR_{i,r\leq 600}\beta_3 + year_{clos_i}\beta_4 + district_i\beta_5}$$

- the $event4$ dummy variable assumes value 0 if the small store in the surrounding of the new opening closed in four years from the new opening (*model5*)

$$\lambda_{4\,years}\,(t, x_i) = \lambda_0(t) \cdot e^{sq\ mts\ super_i\beta_1 + DP_{i,r\leq 60}\ \beta_2 + DR_{i,r\leq 600}\beta_3 + year_{clos_i}\beta_4 + district_i\beta_5}$$

- the $event5$ dummy variable assumes value 0 if the small store in the surrounding of the new opening closed in five years from the new opening (*model6*)

$$\lambda_{5\,years}\,(t, x_i) = \lambda_0(t) \cdot e^{sq\ mts\ super_i\beta_1 + DP_{i,r\leq 600}\beta_2 + DR_{i,r\leq 600}\beta_3 + year_{clos_i}\beta_4 + district_i\beta_5}$$

However, small stores can differently be influenced by the opening of a new supermarket.

Then, we specify the five models for each j category[37] and we estimate $\lambda_{1\,year,j}\,(t, x_{ij})$, $\lambda_{2\,years,j}\,(t, x_{ij})$, $\lambda_{3\,years,j}\,(t, x_{ij})$, $\lambda_{4\,years,j}\,(t, x_{ij})$, $\lambda_{5\,years,j}\,(t, x_{ij})$.

[37] Groceries; Furnishing and furnishing accessories; Newspapers, books and toys; Clothing and clothing accessories; Medicines and cosmetics; Flowers and pets; Sanitary and orthopaedic articles

The models that have been specified, finally, are 48. Each set of models refers to a special selection of small stores, as defined above. Table 18, for example, presents the Cox regression considering all the stores and the six models have six different duration variables. Tables 19 present the same model specifications.

Before proceeding with the analysis of the results in order to test the hypotheses, one can notice that the hazard ratios and their significance strongly differ according to the selected merchandise category j.

Table 18 - The Cox proportional hazard model for all the stores

	model 1		model 2		model 3		model 4		model 5		model 6	
	hr	sign	hr	sign	hr	sign	hr	sign	hr	sign	hr	sign
sq mts super	1.00023	**	1.00021	**	1.00021	**	1.00021	**	1.0002	**	1.00029	***
	0.0000		0.0000		0.0002		0.0000		0.0000		0.0000	
DP	0.99991	**	0.99998	**	0.99998	**	0.99994	**	0.99992	**	0.99988	***
	0.0000		0.0000		0.0000		0.0000		0.0000		0.0000	
DR	0.99999	**	0.99999		0.99999		1.00001		1.00001	**	0.99999	***
	0.0006		0.0015		0.0011		0.0009		0.0008		0.0008	
Year_clos	Y		Y		Y		Y		Y		Y	
District	Y		Y		Y		Y		Y		Y	
Log-likelihood	-118,281.25		-19,953.45		-36,460.66		-50,239.25		-62,304.76		-71,781.27	
LR (χ^2)	1,113.09***		1,112.44***		1,693.89***		1,907.98***		1,753.89***		1,560.34***	
N. obs	13902		13902		13902		13902		13902		13902	

Note: * p<0.05, ** p<0.01, *** p<0.001

Tables 19 - The Cox proportional hazard model for the stores divided according to the merchandise category they sell

Groceries

	model 1		model 2		model 3		model 4		model 5		model 6	
	hr	sign	hr	sign	hr	sign	hr	sign	hr	sign	hr	sign
sq mts super	1.00028	***	1.00029	***	1.0003	***	1.0003	***	1.00029	***	1.00029	***
	0.0000		0.0000		0.0000		0.0000		0.0000		0.0000	
DP	0.99999	***	0.99981	***	0.99986	***	0.99977	***	0.99982	***	0.99989	***
	0.0000		0.0001		0.0001		0.0000		0.0000		0.0000	
DR	0.99747		0.99835		0.99722		0.99604		0.99417		0.99446	
	0.0011		0.0026		0.0020		0.0017		0.0015		0.0014	
Year_clos	Y		Y		Y		Y		Y		Y	
District	Y		Y		Y		Y		Y		Y	

Note: * p<0.05, ** p<0.01, *** p<0.001

Furnishing and furnishing accessories

	model 1		model 2		model 3		model 4		model 5		model 6	
	hr	sign	hr	sign	hr	sign	hr	sign	hr	sign	hr	sign
sq mts super	1.0002	**	1.00033		1.00032	**	1.00026	*	1.00022	*	1.00017	
	0.0000		0.0000		0.0000		0.0000		0.0000		0.0000	
DP	1.00019	**	0.9998		0.9999		1.00025		1.00031	*	1.00013	
	0.0000		0.0000		0.0000		0.0000		0.0000		0.0000	
DR	0.99878		0.98825		0.99044		0.9969		0.99962		0.99828	
	0.0000		0.0100		0.0100		0.0100		0.0000		0.0000	
Year_clos	Y		Y		Y		Y		Y		Y	
District	Y		Y		Y		Y		Y		Y	

Note: * p<0.05, ** p<0.01, *** p<0.001

Newspapers, books, toys

	model 1		model 2		model 3		model 4		model 5		model 6	
	hr	sign	hr	sign	hr	sign	hr	sign	hr	sign	hr	sign
sq mts super	0.99999	***	0.99998	***	0.99988	**	0.99988		0.99985		0.99983	*
	0.0000		0.0000		0.0000		0.0000		0.0000		0.0000	
DP	0.99996		1.00003		0.99994		0.99992		1.00006		1.0001	
	0.0000		0.0000		0.0000		0.0000		0.0000		0.0000	
DR	0.99977		0.99405		0.99441		1.00992		1.00506		1.00273	
	0.0100		0.0100		0.0100		0.0100		0.0100		0.0100	
Year_clos	Y		Y		Y		Y		Y		Y	
District	Y		Y		Y		Y		Y		Y	

Note: * p<0.05, ** p<0.01, *** p<0.001

Clothing and shoes, clothing accessories

	model 1		model 2		model 3		model 4		model 5		model 6	
	hr	sign	hr	sign	hr	sign	hr	sign	hr	sign	hr	sign
sq mts super	1.0002	***	1.00024	***	1.00021	***	1.00019	***	1.00019	***	1.00017	***
	0.0000		0.0000		0.0000		0.0000		0.0000		0.0000	
DP	1.00016	***	1.00036	***	1.00016	**	1.0003	***	1.00024	***	1.00015	***
	0.0000		0.0000		0.0000		0.0000		0.0000		0.0000	
DR	0.99777	*	0.99776	*	0.99762	**	0.99886	**	1.00037		1.00088	
	0.0000		0.0000		0.0000		0.0000		0.0000		0.0000	
Year_clos	Y		Y		Y		Y		Y		Y	
District	Y		Y		Y		Y		Y		Y	

Note: * p<0.05, ** p<0.01, *** p<0.001

Medicines and cosmetics

	model 1		model 2		model 3		model 4		model 5		model 6	
	hr	sign	hr	sign	hr	sign	hr	sign	hr	sign	hr	sign
sq mts super	1.00037	***	1.00018		1.00018		1.00016		1.00031	***	1.00034	***
	0.0000		0.0000		0.0000		0.0000		0.0000		0.0000	
DP	0.99674	**	0.97968	**	0.99208	**	0.99951	*	0.9995	*	1.00001	
	0.0000		0.0000		0.0000		0.0000		0.0000		0.0000	
DR	0.99998	**	0.99988	**	0.99987	**	0.99996	**	0.99997	**	1.00003	
	0.0000		0.0100		0.0100		0.0100		0.0100		0.0100	
Year_clos	Y		Y		Y		Y		Y		Y	
District	Y		Y		Y		Y		Y		Y	

Note: * p<0.05, ** p<0.01, *** p<0.001

Flowers, pets

	model 1		model 2		model 3		model 4		model 5		model 6	
	hr	sign	hr	sign	hr	sign	hr	sign	hr	sign	hr	sign
sq mts super	1.00002	**	1.00001	**	1.00001	**	1.00003	**	1.00005		1.00008	
	0.0000		0.0000		0.0000		0.0000		0.0000		0.0000	
DP	1.00024	***	1.00003		0.99994		1.00021	*	1.00015		1.00007	
	0.0000		0.0000		0.0000		0.0000		0.0000		0.0000	
DR	1.00196		0.99405		0.99441		0.99866		0.99724		1.00133	
	0.0000		0.0100		0.0100		0.0000		0.0000		0.0000	
Year_clos	Y		Y		Y		Y		Y		Y	
District	Y		Y		Y		Y		Y		Y	

Note: * p<0.05, ** p<0.01, *** p<0.001

Sanitary and orthopaedic articles

	model 1		model 2		model 3		model 4		model 5		model 6	
	hr	sign	hr	sign	hr	sign	hr	sign	hr	sign	hr	sign
sq mts super	0.99998	**	0.99995	**	0.99997		0.99994		0.99996	**	1.00008	
	0.0000		0.0000		0.0000		0.0000		0.0000		0.0000	
DP	0.99998		1.00012		0.99975		0.99991		0.99997		1.00002	
	0.0000		0.0000		0.0000		0.0000		0.0000		0.0000	
DR	1.01142		1.01304		1.00325		1.00802		1.00766		1.00563	
	0.0100		0.0200		0.0100		0.0100		0.0100		0.0100	
Year_clos	Y		Y		Y		Y		Y		Y	
District	Y		Y		Y		Y		Y		Y	

Note: * $p<0.05$, ** $p<0.01$, *** $p<0.001$

In order to better test the hypotheses one can use tables 20 and 21. There, results are summed up using the increase (decrease) in the probability for a store to close if a medium size supermarket (1375 sq. mts.) opens in a 600 mts. radius from it, keeping all the other variables constant.

Table 20 - Increase (decrease) in the probability for a store to close after the new opening of a medium size supermarket, given the merchandise category the store sells

All the stores	**+37.19%**	**
Groceries	+47.19%	***
Furnishing and furnishing accessories	+32.19%	**
Newspapers, books, toys	**-1.19%**	***
Clothing and shoes, clothing accessories	+32.19%	***
Medicines and cosmetics	+66.19%	***
Flowers, pets	+3.19%	**
Sanitary and orthopaedic articles	**-3.19%**	***

Table 21 - Increase (decrease) in the probability for a store to close 1, 2, 3, 4 years after the new opening of a medium size supermarket, given the merchandise category the store sells[38]

All the stores

when they open	1 year after the new opening	2 year after the new opening	3 year after the new opening	4 year after the new opening	5 year after the new opening
+33.47%	**+33.47%**	**+33.47%**	**+31.65%**	**+48.99%**	

Groceries

when they open	1 year after the new opening	2 year after the new opening	3 year after the new opening	4 year after the new opening	5 year after the new opening
+48.99%	**+51.05%**	**+51.05%**	**+48.99%**	**+48.99%**	

[38] Beta values are black if $p<0.05$, grey otherwise. For more details about significance, see tables 18.

Furnishing and furnishing accessories

when they open	1 year after the new opening	2 year after the new opening	3 year after the new opening	4 year after the new opening	5 year after the new opening
+57.41%	**+55.26%**	**+42.97%**	**+35.32%**	+26.33%	

Newspapers, books, toys

when they open	1 year after the new opening	2 year after the new opening	3 year after the new opening	4 year after the new opening	5 year after the new opening
-3.00%	**-15.00%**	-15.21%	-18.64%	**-20.85%**	

Clothing and shoes, clothing accessories

when they open	1 year after the new opening	2 year after the new opening	3 year after the new opening	4 year after the new opening	5 year after the new opening
+39.09%	**+33.47%**	**+29.85%**	**+29.85%**	**+26.33%**	

Medicines and cosmetics

when they open	1 year after the new opening	2 year after the new opening	3 year after the new opening	4 year after the new opening	5 year after the new opening
+28.08%	+28.08%	+24.61%	**+53.14%**	**+59.59%**	

Flowers, pets

when they open	1 year after the new opening	2 year after the new opening	3 year after the new opening	4 year after the new opening	5 year after the new opening
+1.02%	**+1.38%**	+4.21%	+7.12%	+11.63%	

Sanitary and orthopaedic articles

when they open	1 year after the new opening	2 year after the new opening	3 year after the new opening	4 year after the new opening	5 year after the new opening
-7.00%	-4.04%	-7.92%	**-5.35%**	**11.63%**	

Therefore:

Hypothesis 1: all the small stores are negatively impacted by the new opening of a supermarket in its neighborhood. Indeed, their probability to close after the new supermarket opening increases by 37%. The same probability changes when time is segmented. In particular, the higher number of closing brick and mortar stores happens in the fifth year after the new opening (49%).

Hypothesis 2: small stores sell products that are either complements or substitutes to the ones that the supermarket sells. Most of the assortment of the supermarket is dedicated to grocery. Indeed, groceries suffer the most the new supermarket opening; here, consumers can cherry-pick and, hence, reduce the transportation costs that emerge from moving from a store to another when they are interested in filling their basket with foodstuffs. So, groceries when a supermarket opens in the surrounding tend to close easily (47%). Even higher (66%) is the negative spillover of the supermarket's new opening for stores selling medicines and cosmetics. Indeed, at least in Italy, 78%[39] of the medium size supermarket has a unit selling those goods. Hence, considering that consumers buying products that are standard in price and other product characteristics (the best example of which is medicines, indeed) are mostly interested in reducing the search cost, the supermarket is the best place where to purchase; when buying groceries, they can also buy medicines without moving from a store to another. On the other hand, sanitary and orthopaedic articles are not sold in the supermarket at all. Hence, they do not suffer from the new opening. Newspapers, books and toys, not only are sold in a small unit in the supermarket, but they are differentiated, also. Hence, consumers prefer to compare products among a wide assortment. Indeed, books and toys have a wide range of

[39] This data emerges from a survey made by the author to all the 91% of the medium sized supermarket in Turin.

product characteristics that makes specialized stores (news vendors, bookstores, stores selling toys) more attractive to consumers.

Hypothesis 3: the negative impact of a new opening of a supermarket on local retailers decreases over time only for stores selling clothing, shoes and clothing accessories, for stores selling furnishing and furnishing accessories. For those categories, supermarkets typically have a very small and basic assortment of the two. However, the ranges of product characteristics for these two categories is wide; clothing, shoes, clothing accessories, furnishing and furnishing accessories are highly differentiable. It is possible, then, for small store to move their product-line to a more store-specific assortment in order to stop compete with the supermarket over time. In this case, the supermarket generates consumers traffic also.

5.4 Discussion

The aim of the chapter was to determine whether supermarket new opening is a danger for the small stores already playing in the local market since spatial competition arises also between distribution channels. Indeed, small stores, selling a defined merchandise category in the place that is supposed to be the best, face both a same-format competition and a competition with large retailers. The chapter focuses on the second kind of competition; it analyzes the negative impact that a new supermarket opening has on its neighbor retailers. First, the negative impact, on average, exists; stores typically suffer from the presence of a supermarket.

However, the empirical test defines that there is a key distinctive factor for letting the two store formats not to compete; it is the product line selection. Indeed, small stores mostly lose the competition with supermarkets when their assortment is fully overlapped with the large retailer one. Indeed, supermarkets propose the best deals and allow consumers to reduce transportation costs.

In order to avoid competition with the supermarket, hence, a small store should diversify its supply with respect to the one of the large retailer; the competition is reduced if small stores can diversify the assortment from the one of the supermarket. This diversification, however, can only happen for stores that sells differentiated products for which the range of products is such wide that assortments can be different from one to the other. Indeed, if products are not differentiated (i.e. prices and other product characteristics are standard), they can not select two different product lines and, hence, they start competing and such a competition is generally won by supermarkets. Hence, stores close with high probabilities. That’s what happens for groceries and drugstores.

Moreover, the assortment of small stores can change over time; they can switch the supply just in response to the supermarket opening. That's why in the first year from the new opening the probability for brick and mortar selling shopping (differentiated; as clothing, shoes, clothing accessories, books and toys) products is higher; time by time, they can select a store-specific assortment that is different from the one of the supermarket. Hence, their survival probability increases year by year from the new opening.

Summing up, the diversification in the assortment can either be a natural diversification or an artificial one.

It is natural if the merchandise category sold by the small store in not sold in the supermarket. In this case, they never start competing and locating near the supermarket can also be a positive choice, given that the supermarket generates consumer traffic.

Diversification is artificial when the category sold by the small store is sold by the supermarket also (even though the supermarket assortment of that merchandise category is small). In this case, if the assortment of the two is overlapped, the small store should make a supply-shift in order not to compete. However, the supply shift is only possible for stores selling merchandise categories whose products are differentiated.

Conclusion

The empirical findings of this study provided a new understanding of the internal dynamics of the retail industry and a new definition of what accessibility is. One of the more significant findings to emerge from this study was that population, which is a location attribute, can either be seen as the stock of people who live around a place or the flow of people passing through a location, driven by the streets in the urban network. Moreover, depending on the destination of the consumers' journey, the flows of people are of different quality in the eyes of a store owner, and the importance for a store to be located where the flows of people are intense strongly depends on the category of the merchandise that the store sells. Further, the category of merchandise that the store sells is a key feature based on which a store owner can select the best location with regard to the population living in the surrounding, the density of other stores in the surrounding and the flow of people passing through the streets in the urban network. Indeed, in chapter 3, we first show that stores selling homogeneous products and stores selling comparable goods can benefit differently from being located in population hotspots and commercial areas. Second, we demonstrate that daily commutes to workplaces do not benefit a retailer along the trip as much as journeys for shopping purposes do. Indeed, the higher the number of times a potential consumer passes by a location when getting to workplaces, the lower is the probability of finding a store in that location. By contrast, a typical building located on the route while making trips from store to store has a high probability of hosting a store.

Thus far, the thesis has considered the commercial pattern made by stores belonging to the same category: small stores. Next, the thesis proposes an original analysis of the competitive dynamics between stores of different sizes; the aim of chapter 4 was to investigate whether the opening of a supermarket is a threat for small stores in the neighborhood. Indeed, by offering a wide range of products and lower prices, supermarkets can attract consumers who were used to frequenting small stores before the new opening. It was observed and quantified that the probability for a small store to exist in a commercial pattern by competing with a newly opened big store depends on the capability the small store has to differentiate its assortment from the new entrant. Indeed, when a small store finds itself being forced to compete with bigger stores, the selection of the small store's assortment assumes a role of primary importance. Therefore, in order to avoid competition with the supermarket, a small store should diversify its supply with respect to one of the large retailers; the competition is reduced if small stores can diversify the assortment from that of the supermarket. This diversification, however, can only happen for stores that sell differentiated products for which the range of products is so wide that assortments can be different among stores. Indeed, if products are not differentiated (i.e., prices and other product characteristics are standard), stores can not select two different product lines and, hence, start competing, and such a competition is generally won by supermarkets.

Furthermore, the empirical test raises a few more research questions.

First, in all the spatial analyses, the scale of the analyses matter. The work shows that the impact of the distribution of consumers is weaker than other factors (distribution of competitors and accessibility to consumers) at the urban scale, but one wonders whether the same conclusion holds for longer distances. Thus, one might want to test the same hypotheses at the province rather than at the city scale. In that case, the impact of the distribution of the population might be more compelling.

Second, in our analyses, we grouped the stores in two large sets: convenience stores and stores selling shopping goods. Given the breadth of our analyses, the impact of agglomeration (e.g. due to complementarity of stores) and natural advantage outweigh the impact of competition. The density of stores tends to have a positive impact on the probability that a building will host a store. One wonders if the same results hold when we consider smaller (though more consistent) samples of stores selling similar, and possibly, substitute products. In this case, the number of observations drops dramatically, and the statistical significance of results might be an issue.

Third, one wonders whether the results, though very robust empirically, partially depend on the sample. Thus the comparison of data from Turin with data from other cities might provide further interesting insights. In this case, the collection of comparable data might be an issue.

Finally, much of the thesis was written in 2020 when most of us were confined to work in our homes and we could only purchase groceries and essential items in the area near our homes. The COVID-19 pandemic has imposed significant changes on our way of living, particularly with respect to households' spending patterns and movements. Of course, this is a transient state, and one can hardly predict the long-term changes in consumer behavior that the pandemic will generate. However, thanks to the original contribution of the thesis, we know that if consumer behaviors and movements within the city change, the urban retail system will likely change accordingly. We also know that retail stores selling non-essential goods have been forced to close for many months. In this scenario, the thesis is well suited to understand the possible impacts of changes in buying behavior on the metabolism of the urban retail system.

References

Aguirregabiria V and Vicentini G (2016) Dynamic spatial competition between multi-store retailers. *The journal of in* LXIV(4): 710–754.

Alonso W (1964) *Location and Land Use*. Cambridge, MA: Harvard University Press.

American Marketing Association (1948) Report of the Definitions Committee. *The journal of marketing* 13: 202–217.

Arnold SJ, Handelman J and Tigert DJ (1998) The impact of a market spoiler on consumer preference structures (or, what happens when Wal-Mart comes to town). *Journal of Retailing and Consumer Services* 5(1): 1–13. DOI: 10.1016/s0969-6989(97)00011-8.

Artz GM and Stone KE (2006) Analyzing the impact of Wal-Mart Supercenters on Local Food Store Sales. 88(5): 1296–1303.

Basker E (2005) Job creation or destruction? Labor market effects on Wal-Mart exapansion. *The Review of Economics and Statistics* 87(1): 174–183. DOI: 10.1002/mde.4090020209.

Basker E, Aguirregabiria V and Suzuki J (2016) Empirical games of market entry and spatial competition in retail industries. In: *Handbook on the Economics of Retailing and Distribution*.

Batty M (2005) *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals*. Cambridge, MA: The MIT Press.

Baum JAC and Haveman HA (1997) Love thy neighbor? Differentiation and agglomeration in the Manhattan hotel industry, 1898-1990. *Administrative Science Quarterly*.

Bavelas A (1948) A mathematical model for group structures. *Human Organization* 7: 16–30.

Bavelas A (1950) Communication patterns in task oriented groups. *Journal of the Acoustical Society of America* 22: 271–282.

Boulding KE (1966) *Economic Analysis*. New York: Harper and Row.

Brown S (1993) Retail location theory: Evolution and evaluation. *The International Review of Retail, Distribution and Consumer Research* 3(2): 185–229.

Brynjolfsson E and Smith MD (2000) Frictionless commerce? A comparison of Internet and conventional retailers. *Management Science* 46(4): 563–585.

Burdett K and Malueg DA (1981) The theory of search for several goods. *Journal of Economic Theory*.

Buzzacchi L, Leveque P, Taramino R, et al. (2020) Using betweenness metrics to investigate the geographical distribution of retailers. *Environment and Planning B: Urban Analytics and City Science* 0(0): 1–18.

Carlson JA and McAfee RP (1984) Joint search for several goods. *Journal of Economic Theory*.

Carree M and Thurik R (1996) Entry and exit in retailing: Incentives, barriers, displacement and replacement. *Review of Industrial Organization* 11(2): 155–172.

Casalaina V and Horst WJR (1967) *Generating Floor Plans from Adjacency Matrices*.

Chevalier J and Goolsbee A (2003) Measuring Prices and Price Competition Online: Amazon.com and BarnesandNoble.com. *Quantitative Marketing and Economics* 2(1): 203–222.

Christaller G (1933) Christaller Centrality Approach. *A Geography of Urban Places*: 1–9.

Christaller W (1933) *Die Zentralen Orte in Süddeutschland*. Jena: Gustav Fischer.

Clark WA (1968) Consumer travel pattern and concept of range. *Annals of the association of american geographers* 58: 368–396.

Clay K, Krishnan R, Wolff E, et al. (2002) Retail strategies on the web: Price and non-price competition in the online book industry. *Journal of Industrial Economics* 50(3): 351–367.

Copeland MT (1923) Relation of consumers' buying habits to marketing methods. *Harvard Business Review* 1(3): 282–289.

D'aspremont C, Gabszewicz J and Thisse J-F (1979) On Hotelling's stability in competition. *Econometrica* 47(5): 1145–1150.

De Palma A, Ginsburgh V, Papageorgiou YY, et al. (1985) The principle of minimum differentiation holds under sufficient heterogeneity. *Econometrica* 53(4). JSTOR: 767–781.

Dennis C, Marsland D and Cockett T (2002) Central place practice: Shopping centre attractiveness measures, hinterland boundaries and the UK retail hierarchy. *Journal of Retailing and Consumer Services* 9(4): 185–199.

DiPasquale D and Wheaton W (1996) The Urban Land Market: Rents and Prices. *Urban economics and real estate markets*.

Dolega L, Pavlis M and Singleton A (2016) Estimating attractiveness, hierarchy and catchment area extents for a national set of retail centre agglomerations. *Journal of Retailing and Consumer Services* 28. Elsevier: 78–90.

Dudey M (1990) Competition by Choice: The Effect of Consumer Search on Firm Location Decisions. *American Economic Review* 80(5): 1092–1104.

Eaton BC and Lipsey RG (1979) Comparison Shopping and the Clustering of Homogeneous Firms. *Journal of Regional Science* 19(4): 421–435.

Economico M dello sviluppo (2018) Rapporto sul sistema distributivo. Analisi strutturale del commercio italiano, 2018.

Enders A and Jelassi T (2000) The Converging Business Models of Internet and Bricks-and-Mortar Retailers. *European Management Journal* 18(5): 542–550.

Eppli M and Benjamin JD (1994) The Evolution of Shopping Center Research: A Review and Analysis. *Journal of Real Estate Research* 9(1): 5–32.

Eppli MJ and Shilling JD (1995) Large-scale shopping center development opportunities. *Land Economics*.

Fischer JH and Harrington JE (1996) Product Variety and Firm Agglomeration. *The RAND Journal of Economics* 27(2): 281.

Fox E, Postrel S and McLaughlin A (2007) The impact of retail location on retailer revenues: An Empirical investigation. *Southern Methodist University, ...* (January): 1–45.

Fox EJ and Hoch SJ (2005) Cherry-picking. *Journal of Marketing*. DOI: 10.1509/jmkg.69.1.46.55506.

Fox EJ, Montgomery AL, Lodish LM, et al. (2004) Consumer Shopping and Spending across Retail Formats. *Journal of Business* 77(2): 25–60.

Fujita M and Thisse J-F (2002) Economics of Agglomeration: Cities. *Industrial*

Location, and Regional Growth, Cambridge 378: 339–378.

Ghosh A and Craig CS (1983) Formulating Retail Location Strategy in a Changing Environment. *Journal of Marketing* 47(3): 56.

Goldmanis M, Hortaçsu A, Syverson C, et al. (2010) E-Commerce and the Market Structure of Retail Industries. *Economic Journal* 120(545): 651–682.

Golledge RG, Rushton G and Clark WA (1966) Some spatial characteristics of Iowa's dispersed farm population and their implications for the grouping of central place functions. *Economic Geography* 42(3): 261–272.

Guo WC and Lai FC (2017) Prices, Locations and Welfare When an Online Retailer Competes with Heterogeneous Brick-and-Mortar Retailers. *Journal of Industrial Economics* 65(2): 439–468.

Guy CM (1998) Classifications of retail stores and shopping centres: Some methodological issues. *GeoJournal* 45(4): 255–264.

Haltiwanger J, Jarmin R and Krizan CJ (2010) Mom-and-Pop meet Big-Box: Complements or substitutes? *Journal of Urban Economics* 67(1). Elsevier Inc.: 116–134.

Hausman J and Leibtag E (2007) Consumer benefit from increased competition in shopping outlets: measuring the effect of Wal-Mart. *Journal of Applied Econometrics* 22: 1157–1177.

Hidalgo CA, Castañer E and Sevtsuk A (2020) The amenity mix of urban neighborhoods. *Habitat International* 106(September 2019).

Holton RH (1958) The distinction between convenience goods, shopping goods, and specialty goods. *Journal of Marketing* 23(1): 53–56.

Horst WJR (n.d.) Theories of cell configuration. In: Moore GT (ed.) *Emerging Methods in Environmental Design and Planning*. The MIT Press.

Hotelling H (1929) Stability in Competition. *Economic Journal* 39: 41–57.

Irwin EG and Clark J (2006) Wall Street vs. Main Street: what are the benefits and costs of Wal-Mart to local communities. *Choices Magazine* 21(2): 117–122.

Jacobs J (1969) Strategies for Helping Cities. *American Economic Association* 59(4): 652–656.

Jensen P, Boisson J and Larralde H (2005) Aggregation of retail stores. *Physica A: Statistical Mechanics and its Applications* 351(2–4). Elsevier: 551–570.

Kang CD (2016) Spatial access to pedestrians and retail sales in Seoul, Korea. *Habitat International* 57. Elsevier Ltd: 110–120.

Konishi H (2005) Concentration of competing retail stores. *Journal of Urban Economics* 58(3): 488–512.

Koster HRA, Pasidis I and Ommeren J van (2019) Shopping externalities and retail concentration: Evidence from Dutch shopping streets. *Journal of Urban Economics*.

Kotler P and Armstrong G (2010) *Principles of Marketing*. Pearson education.

Krizek KJ (2003) Operationalizing neighborhood accessibility for land use - Travel behavior research and regional modeling. *Journal of Planning Education and Research* 22(3): 270–287.

Leavitt HJ (1951) Some effects of certain communication patterns on group performance. *Journal of Abnormal and Social Psicology* 46: 38–50.

Levin PH (1964) The Use of Graphs to Decide the Optimum Layout of Buildings. *Architects' Journal* 140(15): 809–815.

Losch A and others (1954) The economics of location. Yale University Press.

March L and Steadman P (1971) *The Geometry of Environment: An Introduction to Spatial Organization in Design*. London: Riba Publishing.

McLafferty SL and Ghosh A (1986) Multipurpose Shopping and the Location of Retail Firms. *Geographical Analysis* 18(3): 215–226.

Miller D, Jackson P, Rowlands M, et al. (1998) *Shopping, Place, and Identity*. Psychology Press.

Miller HJ and Finco M V (1995) Spatial search and spatial competition: A probability analysis of basic results from the spatially-restricted theory. *The Annals of Regional Science* 29(1). Springer: 67–89.

Miller HJ and Wu YH (2000) GIS software for measuring space-time accessibility in transportation planning and analysis. *GeoInformatica* 4(2): 141–159.

Mulligan GF (1988) Consumer travel behavior: extensions of a multipurpose shopping model. *Mathematical and Computer Modelling* 10(10): 791.

Netz JS and Taylor BA (2002) Maximum or minimum differentiation? Location patterns of retail outlets. *Review of Economics and Statistics* 84(1). MIT Press

O'Kelly ME (1983) Impacts of multistop, multipurpose trips on retail distributions. *Urban Geography* 4(2): 173–190.

Okabe A and Sugihara K (2012) Spatial analysis along network: statistical and computatonal methods. In: *Statistics in Practice*. Hoboken: Wiley.

Ozuduru BH and Varol C (2011) Spatial statistics methods in retail location research: A case study of Ankara, Turkey. In: *Procedia Environmental Sciences*, 2011.

Peponis J, Bafna S, Bajaj R, et al. (2007) Designing space to support knowledge work. *Environment and Behavior* 39(6): 815–840.

Peponis J, Bafna S and Zhang Z (2008) The connectivity of streets: Reach and directional distance. *Environment and Planning B: Planning and Design* 35(5): 881–901.

Picone GA, Ridley DB and Zandbergen PA (2009) Distance decreases with differentiation: Strategic agglomeration by retailers. *International Journal of Industrial Organization* 27(3). Elsevier B.V.: 463–473.

Porta S, Crucitti P and Latora V (2006) The network analysis of urban streets: A primal approach. *Environment and Planning B: Planning and Design* 33(5): 705–725.

Porta S, Strano E, Iacoviello V, et al. (2009) Street centrality and densities of retail and services in Bologna, Italy. *Environment and Planning B: Planning and Design* 36(3): 450–465.

Prince JT (2007) The beginning of online/retail competition and its origins: An application to personal computers. *International Journal of Industrial Organization* 25(1): 139–156.

Rushton G (1969) Analyzing spatial behavior by revealed preference. *Annals of the association of american geographers* 59: 391–400.

Seim K (2006) An empirical model of firm entry with endogenous product-type choices. *RAND Journal of Economics* 37(3): 619–640.

Sevtsuk A (2010) *Path and Place: A study of Urban Geometry and Retail Activity in Cambridge and Somerville, MA*.

Sevtsuk Andres (2014a) Location and Agglomeration: The Distribution of Retail and Food Businesses in Dense Urban Environments. *Journal of Planning Education and Research* 34(4): 374–393.

Sevtsuk A. (2014) Location and Agglomeration: The Distribution of Retail and Food Businesses in Dense Urban Environments. *Journal of Planning Education and Research* 34(4): 374–393.

Sevtsuk Andres (2014b) Networks of the Built Environment. *Decoding the City*: 144–159.

Sevtsuk A. and Kalvo R (2018) Patronage of urban commercial clusters: A network-based extension of the Huff model for balancing location and size. *Environment and Planning B: Urban Analytics and City Science* 45(3): 508–528.

Sevtsuk Andres and Kalvo R (2018) Patronage of urban commercial clusters: A network-based extension of the Huff model for balancing location and size. *Environment and Planning B: Urban Analytics and City Science* 45(3): 508–528.

Sevtsuk A and Mekonnen M (2012) Urban network analysis: A new toolbox for measuring city form in ArcGIS. In: 2012, pp. 133–145.

Shaw ME (1964) Communication networks. In: Berkovitz L (ed.) *Advances in Experimental Social Psychology*. New York: Academic Press, pp. 111–147.

Shimbel A (1953) Structural parameters of communication networks. *Bulletin of mathematical biophysics* 15: 501–507.

Talukdar D, Gauri DK and Grewal D (2010) An empirical analysis of the extreme cherry picking behavior of consumers in the frequently purchased goods market. *Journal of Retailing* 86(4). New York University: 336–354.

Urbany JE, Dickson PR and Kalapurakal R (1996) Price search in the retail grocery market. *Journal of Marketing* 60(2): 91–104.

Waldfogel J (2008) The median voter and the median consumer: Local private goods and population composition. *Journal of Urban Economics* 63(2): 567–582.

Weltevreden JWJ and Rietbergen T Van (2009) The implications of e-shopping for in-store shopping at various shopping locations in the Netherlands. *Environment and Planning B: Planning and Design* 36(2): 279–299.

Wolinsky A (1983) Retail Trade Concentration Due To Consumers' Imperfect Information. *Bell Journal of Economics* 14(1): 275–282.

Xie F and Levinson D (2007) Measuring the structure of road networks.

Geographical Analysis 39(3): 336–356.

Yoshimura Y, Santi P, Arias JM, et al. (2020) Spatial clustering: Influence of urban street networks on retail sales volumes. *Environment and Planning B: Urban Analytics and City Science* 0(0): 1–17.