INSTITUTE OF RETAIL ECONOMICS

POPULATION DECLINE AND CHANGES IN FOOD STORE PROXIMITY

HELENA NILSSON

HFI WORKING PAPER No 13

Population decline and changes in food store proximity

Helena Nilsson[§]

This study examines changes in the access to food stores between 2000-2013 using Swedish register data. The analysis shows that there have been increases of, on average, 400 meters in the neighborhood-level average distance to the nearest store. These changes are not systematically biased towards socioeconomically disadvantaged groups. I also analyze the relationship between population density and the distance from the neighborhood centroid to the nearest food store using a spatiotemporal analysis of panel data covering the same period. The analysis finds that a decline in population density is related to an increase in the distance to the nearest food store. This indicates that as long as there is a decline in population density, food store access by proximity will continue to decrease.

Keywords: food, urbanization, retail, panel-data, spatial-analysis

JEL codes: B23; L81; C33; D22

[§] The Institute of Retail Economics, Stockholm, Sweden and Center for Entrepreneurship and Spatial Economics, Jönköping International Business School, Jönköping, Sweden. Corresponding author, e-mail: Helena.nilsson@ju.se.

1 Introduction

By the early 1990s, progress in the technology of production and logistics had stimulated the growth of hypermarkets (which is a supermarket combined with a department store) and malls in European retail markets. Before this, households' car ownership had become widespread, and large-scale retailers could concentrate in out-of-town areas, where access by car was better and land costs were lower (Dawson, 2006). Parallel to this development, a long-term trend of urbanization caused the population size to shrink in many rural areas (Turok & Mykhnenko, 2007).

When the market size declines below a certain level, firms will close or relocate. This decline in services, such as food stores and gas stations, means that the remaining inhabitants in rural areas have to travel a longer distance to reach alternative facilities (Hodge et al. 2000; Lucas 2004; Woods, 2005). Thus the process of urbanization puts a strain on service delivery in rural areas, and this is a trend that is viewed with growing concern in many OECD countries (Grediaga & Freshwater, 2010). The process of a dispersed spatial configuration of people is a problem also in the urban environment. Urban sprawl, which characterizes many of the European cities today (Arribas-Bel et al., 2011) incurs costs of service provision for the local governments (Varela-Candamio et al, 2019) and it decreases the viability of grocery stores (Hamidi, 2019).

Food access has recently received renewed interest in the context of a general interest in regional inequalities, for instance, in Allcott et al. (2019) and Hamidi (2019). Low food store access is considered a problem from a variety of perspectives. From a health perspective, a poor food environment is claimed to be linked to adverse health outcomes. A poor food environment is often referred to as a "food desert", and this is, according to the US Centers for Disease Control and Prevention (CDC), defined as an area "...that lack access to affordable fruits, vegetables, whole grains, low-fat milk, and other foods that make up the full range of a healthy diet" (CDC, 2017). Thus, an area where there are few or no food stores may be a contributor to unhealthy consumption habits for its inhabitants since there are fewer healthy alternatives to fast food (Glanz et al. 2005).

In addition to its direct function as a provider of foods, the local food store – particularly in rural areas – arguably plays an important role as a local hub (Clarke & Banga, 2010) and often provides additional services. such as postal services and drugstore products (SOU, 2015). Last, food store access is viewed as a basic service that is an integral part of a welfare state and therefore should be available to the population in all parts of a country (Wiberg, 1983). Hence, low access to food stores is considered a problem and is therefore one of the target areas for EU subsidies in the rural development programs of 2007-2014 and 2014-2020 (The Swedish Board of Agriculture, 2019). Swedish food stores in areas with declining market sizes have been eligible to receive governmental subsidies in addition to EU subsidies since 1994 (Swedish Parliament, 1994; 2000). Together with other commercial services (such as gas stations and postal service points) food stores in areas with declining populations received up to a total of 230 M SEK (23 M EUR) between 2002-2007 (The Swedish Consumer Agency, 2009) and 203 M SEK (20,3 M EUR) between 2011-2013 (Swedish Agency for Economic and Regional Growth, 2015).

The topic of food access is thoroughly researched, particularly in the context of food deserts, and there is much variation in how the term is defined. Food access may be defined in a purely geographical sense. Some studies use the distance to a food store (such as the distance from a neighborhood centroid or from inhabitants' place of residence) to define access (Clarke et al. 2002; Babey et al., 2008). Other studies (e.g., Alwitt & Donley, 1997) instead measure food access as a count variable within a predefined area, such as the number of stores within a postal code.

Additional variables are sometimes included in the definition of food store access, such as socioeconomic indicators. These additional factors include, for instance, a high share of low-income earners (e.g., Alwitt & Donley, 1997; Rose & Richards, 2004) or unemployed and elderly inhabitants (e.g., Guy et al., 2004). The argument behind the inclusion of such variables is that these groups face barriers that make poor physical access to groceries particularly problematic. Low income earners are, for instance, more vulnerable to the additional transport costs that low physical access to groceries entails (Alwitt & Donley, 1997).

For an overview of the food desert literature, see Walker et al. (2010) and Lamb et al. (2015). In this paper, however, the main interest is in food access by proximity, and the term "food access" hereinafter refers to the geographical proximity to food stores. Moreover, the type of food retailer that is considered in this study is one that, according to its industry code, provides a broad range of foods. Hence, specialized food stores that often only contain one or a few types of food products, such as bakeries and confectionaries, are not considered.

Central place theory (CPT) (Christaller, 1933; Loesch, 1964) stipulates that one of the main determinants of store location is access to nearby demand. For consumer goods firms, the size of demand may be proxied by the population. According to CPT, given the assumption that population density is uniformly distributed, the distance between two firms selling the same good is determined by the size of the market that is required for each of them to break even costs of production and revenues. The distance from the firm to the perimeter of this market area is termed the threshold distance. An increase in population density should decrease the threshold distance, which then would translate into more firms within the same area and hence better access to the stores for the consumers that reside in the area.

There is a large body of literature that examines the relationship between store location and market size. However, few of these studies examine food store retail (e.g., Salyards & Leitner, 1981; Mulligan et al 1985; Harris et al. 1996; Chakraborty 2012; Larsson & Öner, 2017). Moreover, those that do study food retail examine food store access as a binary variable or as the number of stores in a predefined area. This leaves out the possibility that there is a store right across the border of that area and thereby causes a so-called "edge effect" (Sadler et al., 2011), which may bias the estimates. It also means that physical access (distance) is not measured.

There are some studies that more directly capture physical access in the definition of the dependent variable (e.g. Dutko et al., 2012; Hamidi, 2019) [1]. However, these studies use cross-sectional data, which means that time invariant heterogeneity (e.g. road access or distance from a metropolitan area) is not considered, which may bias the estimates. The present paper contributes to the literature by exploring how the development in neighborhood-level food store access is related to the size of nearby demand while accommodating time-invariant heterogeneity. It also explores the interdependencies between the neighborhoods by examining the presence- and magnitude of spatial spillovers. It thereby informs policy makers on how food store access by proximity can be influenced through residential planning.

The mechanisms at work can be illustrated as in the conceptual model below, Figure 1. When the population density of a neighborhood increases, the local market size increases, which may attract food stores to the area. This results in an increase in the

[1] These two studies define "low food store access" as a binary variable that takes the value 1 if a certain share of the population resides more than 1 mile from a food store.

number of food stores within the area, which in turn will improve food store access by proximity for the inhabitants in the neighborhood. Access to consumption possibilities such as retail is a recognized amenity in the literature (e.g. Öner, 2017), and since amenities may influence individuals' choice if residence (e.g. Biagi et al, 2011), the relationship between food retail and population density is endogenous. Thus, the change in access to food stores in the neighborhood may influence the attractivity to live there, and this will in turn affect the population density. The present study will, however, focus on the effect of population density on food access (as indicated by the grey arrows).

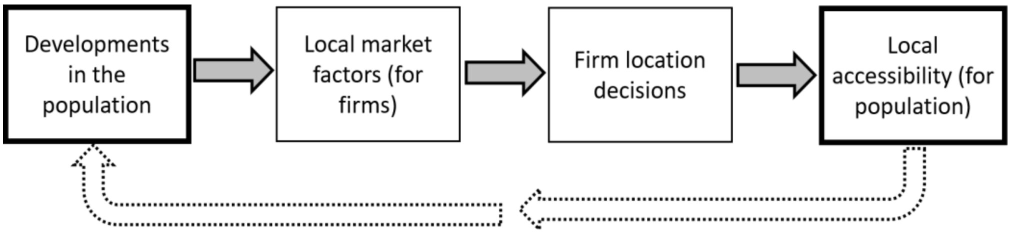

Figure 1. Conceptual model.

Since the size of demand in one neighborhood may influence the number of food stores and, hence, the access in that area, this will also influence the access to food stores in the adjacent neighborhoods. Therefore, there is a spatial dependence between the neighborhoods that must be accounted for to avoid biased estimates.

The first part of the study examines the changes in neighborhood-level food store access by proximity – in terms of the magnitude of change and the characteristics of the neighborhoods that have been affected by a change – that took place between 2000 and 2013. Using geocoded location data on firms and individuals, the distance between the home location of every resident in Sweden and his or her nearest food store is calculated, and for each neighborhood, the average of this distance is calculated annually. The change in the average distance is measured between 2000 and 2013, and the socioeconomic characteristics of the residents in the affected neighborhoods are examined in 2013. By examining variation in food store access on neighborhood level, this part of the study complements a previous analysis by Amcoff (2017) in which food store access was examined for clusters of different socioeconomic groups in Sweden between 1998 and 2008. In Amcoff (2017) it was found that the socioeconomically affluent groups in general had a longer distance to the nearest food store and that the same groups had also experienced an increase in distance during the period.

In the second part of the analysis, the relation between food store access and population density is modeled over time. To assess changes in food store location only, the distance variable is now measured as the distance from the centroid of a neighborhood to the nearest food store. Using an extended Salop (1979) circular market model, the relationship between changes in population density and food store proximity can be derived. The resulting model is estimated using a fixed effects spatial econometric framework that accommodates both time invariant heterogeneity and spatial dependence between neighborhoods.

The descriptive analysis reveals the disappearance of 600 food stores, 10 percent of the total, during the period and indicates that the average neighborhood-level distance of the residents to the nearest food store has increased by approximately 400 meters. When examining the socioeconomic characteristics of the areas where the population has experienced increases in the distance to the nearest food store, these are primarily socioeconomically affluent areas. The results from this part of the analysis are thereby similar to the findings in Amcoff (2017).

The findings from the regression analysis indicate that population density is significantly negatively correlated with the distance from the neighborhood centroid to the nearest food store by 0.18. Thus, an increase of one percent in population density is correlated with a decrease of 0.18 percent in the distance from the centroid of a neighborhood to the nearest food store. When accounting for the spatial dependencies between the neighborhoods, the coefficient of population density remains of a similar magnitude (-0.181), but significant spillover effects (-0.0045) are recorded. Thus, the main part of the effect of the population density is directly correlated with the spatial proximity of food stores, but there are significant spillover effects on adjacent neighborhoods that must be taken into consideration. Since there is a significant negative relationship between the population density and the distance to food stores, areas that have experienced a declining population have also experienced a decline in food store access by proximity. The results thereby indicate that the government- and EU-funded subsidies that are targeted towards maintaining these types of services in areas of population decline may not be efficient in their current form. Other variables that are found to be correlated with the distance to food stores are average income (with a positive elasticity of 0.154 and positive spillover effects of 0.004) and the share of the population with children below the age of 18 living at home (with a negative correlation of 0.38 and negative spillover effects of 0.01).

In the next section, the background and motivation for the paper are elaborated. In section three, the theoretical model adopted is introduced. In section four, the previous literature on the subject is overviewed. In section five, the data and descriptive analysis are presented, and in section six, the empirical method is presented. In section seven, the empirical results are discussed, and section eight summarizes and concludes.

2. Background and motivation

By the beginning of the 1990s, shopping centers had grown drastically in the suburbs of larger Swedish cities, and with that expansion, there had been a parallel decline in the market share of firms in sparsely populated areas (Widerstedt et al., 2006). This trend continued, and at the end of the first decade of the 21st century, the Swedish retail market had come to be characterized by fewer and larger stores (Amcoff et al., 2015). While retail has grown substantially in the past decade and a half (+75 percent from 2000-2015[2]), the growth within the sector has been uneven, with increasing market concentration (increasing by +50 percent for the 10 largest retailers[‡]) and an increase in the share of chain stores (from 2000-2015, this increased by 10 percent[‡]). Moreover, there has been increased competitive pressure on the physical store, in the form of ecommerce, which increased nearly tenfold between 2003-2015, making up 8 percent of retail trade in 2015 (Postnord, 2014; 2018).

The concentration has also occurred in space. In 1992, 75 percent of the sales in retail were concentrated in 88 of the 290 municipalities in Sweden. Approximately 20 years later, in 2015, the same share of sales was concentrated in 66 of the municipalities. These municipalities are primarily urban municipalities in the larger city regions. Parallel to this development, the population has also tended towards a concentration in space. In 1992, 107 of the municipalities had 75 percent of the population, and in 2015, the same share of the population lived in 93[‡] municipalities. The municipalities that hosted this share of the population in 2015 were primarily urban and suburban municipalities. The concentration of both people and services indicates that for the great majority, and for many of the retail categories, the spatial access has been maintained and even improved, on average, for parts of the population. Nonetheless, this development implicates a decline in access for the individuals who remain in rural areas. Regarding food retail, the process of the concentration of stores

[2] Calculations based on statistics from Statistics Sweden.

[‡] see previous footnote.

developed faster than the spatial concentration of people between 1998 and 2008 (Amcoff, 2015). In the same period, the number of stores dropped by 16 percent (Amcoff, 2017). This implies that there is an ongoing decline in food store access for individuals who remain in rural areas.

The population's access to food stores in Sweden began to receive attention in the mid-60s, when a fast decline in the number of stores was noted. For instance, between 1960 and 1970, there was a decrease from 23,000 to 13,000 food stores (Forsberg, 1998). Following a government-induced investigation of consumers' shopping patterns, it was shown that there had been a continuous decline in the number of smaller grocery shops for some time. This development was most clearly evident in the sparsely populated areas of Sweden (SOU, 1972). In 1973, the government began subsidizing food stores in areas where there were few alternative services (The Swedish Consumer Agency, 2008; SOU, 2015), and there were a number of reports and studies mapping access to food stores in rural and later also in urban areas (for instance Forsberg 1998; The Swedish Consumer Agency, 2008, Growth Analysis, 2013). Over the years, there have been additional funds targeting the service in rural areas, with the more recent being subsidies originating in an EU-financed rural development program (SOU, 2015; The Swedish Board of Agriculture, 2019).

For the parts of the population that do not have access to a car or are constrained for other reasons, an increased distance may be particularly problematic (Clarke & Banga, 2010). This is especially true in rural areas where there is a less-developed public transportation system available (Shaw, 2006). Poor access to food stores also lowers the standard of living for people who have cars because it means that more time and money need to be spent to do grocery shopping.

The presence of few or no food stores in an area is problematic since such stores may fulfil a number of complementary functions, in addition to selling groceries. For instance, Clarke and Banga (2010) argue that the local grocery store is important as a meeting place. By providing a social hub, stores in isolated areas become a local community center and a place for the dissemination of information on local events. Moreover, since the stores require staff and they also create job opportunities (Smith & Sparks, 2000). Additionally, they may stimulate business connections, for instance, by creating demand and stimulating the growth of local suppliers (Ilbery & Maye, 2006). From a consumer point of view, there is also the argument that better access to stores increases the local variety of consumption goods (Allcott et al., 2019). It may also be argued that, with no store nearby, traveling to the alternative more distant stores entails environmental costs due to increased car usage.

While food store access is important in many countries, the reasons for this tend to differ, and therefore, the phenomenon has been studied in different contexts. In Anglo-Saxon countries (the U.K. and the U.S.) – where the food desert literature and public policy efforts are more common – the focus of public policies has been to improve access for health reasons as well as socioeconomic equity reasons. This is, for instance, stated in the governmentally issued programs “Healthy Food Financing Initiative” in the US and the “Food Poverty Eradication Bill” of the UK.

As noted by Amcoff et al. (2015), there has been little equivalent focus on health aspects in the Swedish context. Swedish public policy has largely emphasized the additional functions that food stores fulfil, particularly in the rural context. For instance, in the governmental investigation regarding the development of rural areas from 2015 (SOU, 2015), it is argued that basic commercial services, including the local food store, function as a necessary node for both local businesses and a meeting point for people. Moreover, it is argued that the food store may provide additional services, such as postal services, pharmacies and money transfers (SOU, 2015).

Sweden is a country characterized by vast differences in population density, and there are large areas that are sparsely populated and remote. As shown in Figure 2a below, the majority of the country is considered rural.

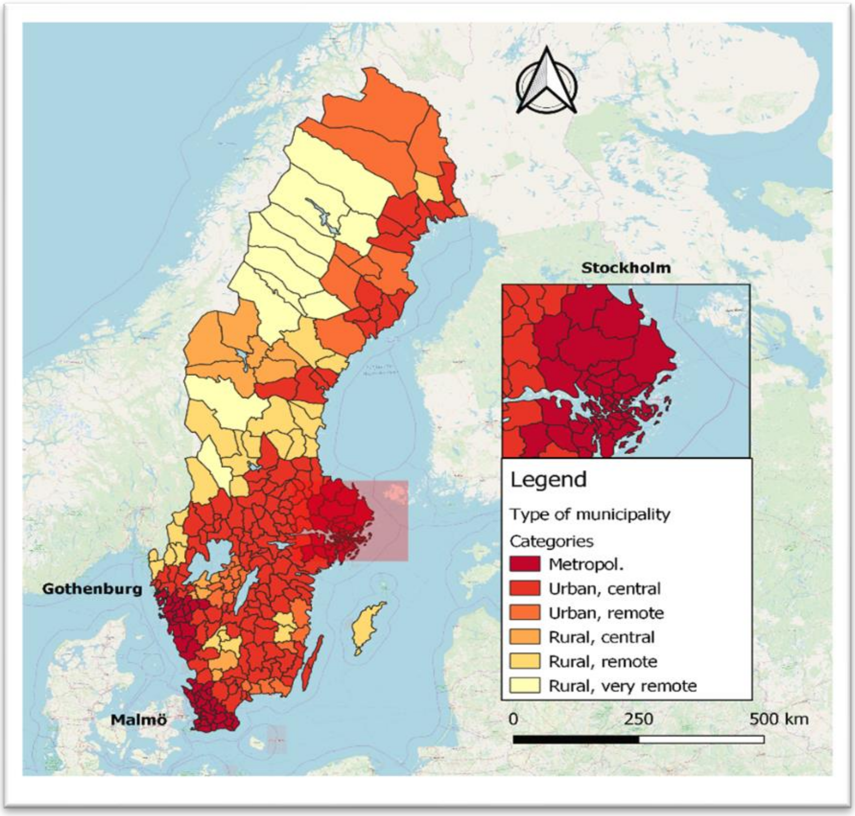

Figure 2a. Swedish municipalities in categories (based on definitions from Swedish Agency for Economic and Regional Growth, 2014)[3]. Source: Statistics Sweden. Program: QGIS. Base map: OpenStreetMap.

During the study period, 2000 to 2013, it is primarily the central municipalities that have grown in terms of population density, as shown in Figure 2b.

[3] The definitions are: **Metropol.** – Municipalities in metropolitan areas. **Urban, central** – close to larger cities – **Urban, remote-** urban municipalities in remote regions. **Rural, central -**rural municipalities close to larger cities

Rural, remote, municipalities in remote areas. **Rural, very remote-** Municipalities in very remote regions. The definitions are from 2014, but they are also expected to be valid for the start of the study period, 2000. Full definitions at: https://www.tillvaxtanalys.se/in-english/publications/reports/reports/2014-04-04-better-statistics-for-better-regional-and-rural-policy.html.

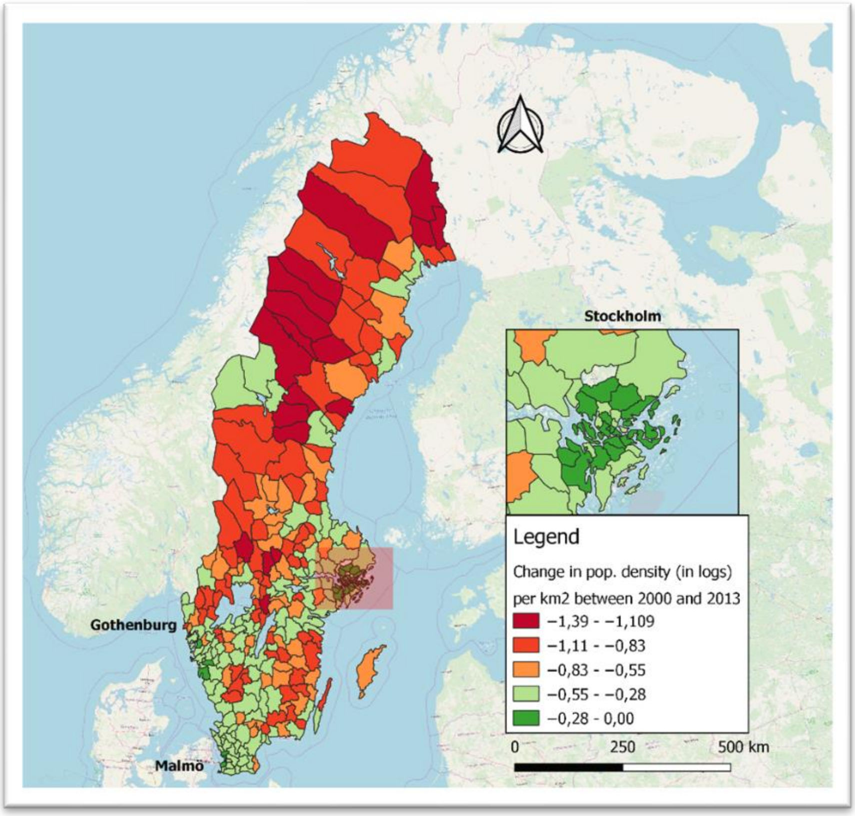

Figure 2b. Change in population density per square kilometer, in logs, 2000-2013. Source: Statistics Sweden. Program: QGIS. Base map: OpenStreetMap[4].

The density of food stores also changed during the period, as shown below in Figure 2c.

[4] The municipality Knivsta was formed in 2003 from Uppsala and is therefore blank in the map.

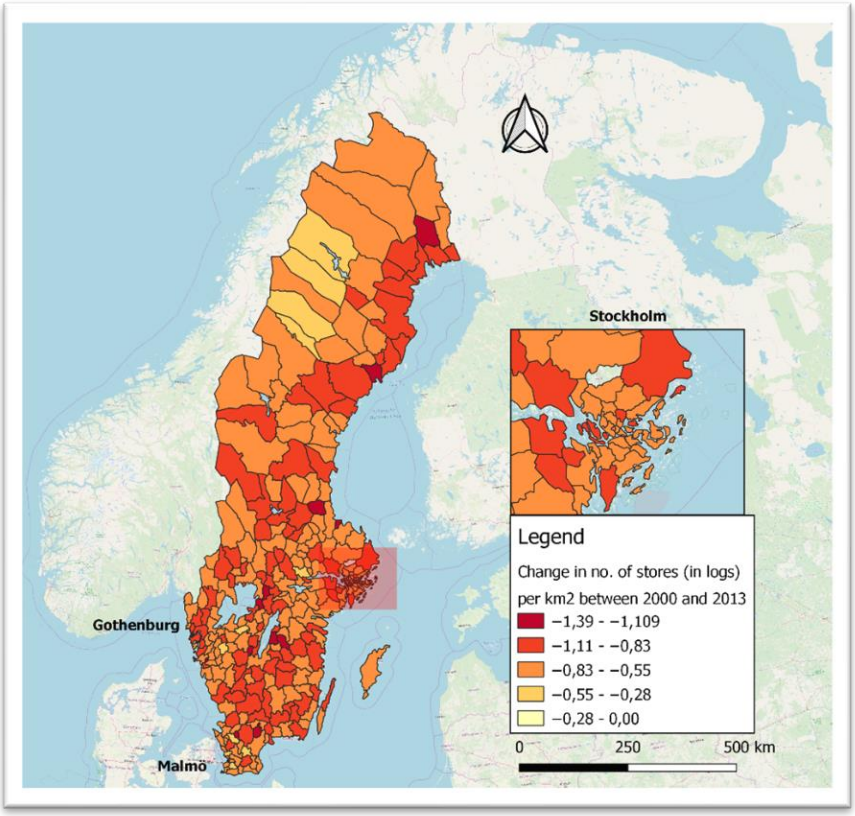

Figure 2c. Change in food stores per square kilometer, in logs, 2000-2013. Source: Statistics Sweden. Program: QGIS. Base map: OpenStreetMap.

As seen in Figure 2c, there has been a decrease in food store density per km2 in all municipalities, which reflects that the number of stores has decreased overall. It appears to be mainly the city regions around Malmö, Gothenburg and Stockholm that have seen the largest relative decrease in the number of stores, while the rural parts have been less affected.

3. Theory

In CPT, Christaller (1933) (and later Loesch, 1964) modeled the uneven pattern of economic activities in space as a function of the consumer's transport costs and the minimum size of demand required to support an economic function (Brakman et al., 2009). The theory is based on the notion that places where market agents interact, central places, are allocated in space in a regular pattern. The assumptions that the theory builds on are that the interactions take place on an isotropic plain, in which a uniformly distributed population of consumers reside. Consumers travel to the nearest central place to purchase a good, and the consumer's utility is thereby a function of the value of the product, its price and the costs of traveling to the location of production. Moreover, the firms on the market are competitive and perfectly informed agents that maximize their profits based on the aggregate demand and costs of production (Parr & Denike, 1970).

As described in Parr and Denike (1970), CPT states that the spacing between the central places can be derived from the demand and production functions of the consumers and producers, respectively. The consumer's demand for a good, at zero distance from the firm, is a downward sloping function of the price. Faced by the aggregate demand curve, the firm finds the optimal price and output where the marginal cost equals the marginal return.

As the distance to the firm increases, the demand for the good, at a given price, decreases, and at one point in space, it drops to zero, and the consumer is supplied by a competitor. This point in space is referred to as the "range" of the good. Given a uniformly distributed population, the radius of the minimum market size that must be covered by a firm to break even is termed the "threshold distance". If the range is larger than the threshold distance, the firm will make profits, and more firms will enter the market. If the opposite holds, the firm will make losses and consequently leave the market.

Since firms that sell different products face different production costs and demand structures, the ranges and the threshold distances will vary for different types of goods. Hence, products that have a low production cost, that are purchased more frequently and for which willingness to travel is low, such as groceries, termed lower order goods, typically have a short range and threshold distance. The opposite is true for higher-order goods, such as furniture. These are the types of goods that have high production costs and that consumers buy rarely and hence are willing to travel a longer distance to obtain (Dicken & Lloyd, 1990). Thus, the spacing between firms that sell lower order goods is smaller than the spacing between firms that sell higher order goods. Consequently, a regular but sparse pattern of higher order firms is layered over a more narrowly spaced pattern of firms supplying lower order goods. Places with higher order firms thus also host lower order firms and are therefore referred to as higher order places.

Since the higher order places host more firms than places that host only lower order firms, they also have a higher rank in the hierarchy of central places (Brakman et al., 2009). In the original theory, Christaller derived a rigid hierarchical structure where higher-order firms in a place guaranteed the presence of lower-order producers. In the Loschian framework, however, this was less rigid and allowed for the sectoral clustering of economic activities and thus specialized centers. The implication of this, brought forward, for instance, by Isard (1956), was that the assumption of a uniformly distributed population no longer held. Thus, the threshold distance would vary with population density. The threshold distance would be smaller and therefore allow for a higher density of producers of the same product in more densely populated areas than in more sparsely populated areas (Dicken & Lloyd, 1990).

Agglomeration externalities were not present in the Loschian framework, but later models, such as those of Berry and Garrison (1958) and Parr and Denike (1970), could show that the relation between the size of nearby demand and the number of producers of a certain good would be positive at a declining rate. Thus, with increasing demand, more producers of a certain good would be required, but at a lower rate. Therefore, with a higher density of demand, the threshold distance for the producers would decrease but at a declining rate.

Against this backdrop, it is possible to derive a theoretical model that describes the relationship between consumer density and store location. I use Salop's (1979) circular market model as a point of departure to express the relationship between consumer density (ρ) and threshold distance (D). The original model assumes a constant density of consumer demand, and in this modified version, I allow consumer density to vary. The steps taken to arrive at the final model presented below are available in appendix A. The final model is as follows:

$$D = \sqrt{\frac{F}{\rho t}} \qquad (1)$$

where F is firm-level fixed costs and t is the costs of transportation per unit of distance for the consumer. It can be shown that the threshold distance is a decreasing function of population density:

$$\frac{dD}{d\rho} = -\frac{F}{2\rho^2\sqrt{\frac{F}{\rho t}}t} < 0 \qquad (2)$$

Equation (1) can be expressed in logarithmic form:

$$lnD = \frac{1}{2}(lnF - lnt - ln\rho) \quad (3)$$

In the regression analysis in this study, the threshold distance D, is proxied by the distance from the centroid of a neighborhood to the nearest food store, and ρ is proxied by population density. The distance to the nearest food store is therefore expected to be negatively affected by an increase in population density, albeit at a decreasing rate. Over time, in areas where population density declines below a certain level, food stores will exit the market, and the distance to the alternative food store will increase.

Analogously, in areas where the population density increases over time, there will be store entries, which would result in a decrease in the distance to the nearest food store. Due to the effects of agglomeration externalities, the relationship between distance D and demand density ρ is assumed to be negative but at a decreasing rate.

4. Previous studies

Disparities in food store access gained attention with the coinage of the term "food desert" in a 1998 report on rising inequalities in the level of health in Britain. The study, conducted by The Social Exclusion Unit (1998) described a pattern of withdrawal of these services in areas where the residents were of poor health and low income (Wrigley, 2002). After the publication of this report, a great number of intervention studies followed. These studies were generally focused on examining the effects of increased physical access to food retail on health outcomes (see, for instance, Wrigley et al., 2003; Cummins et al., 2005). The issue also received attention in the US and gave rise to a large body of literature in a range of research fields, and in the US, government actions were also taken to improve food store availability (Donald, 2013).

In the food desert literature, the focus has been to examine whether there are any systematic socioeconomic inequalities in food store access, and relatively few inferential studies have examined what may explain low food access. There are exceptions, and one is the study by Hamidi (2019), where the relationship between poor food store access and urban sprawl is examined. Poor food store access is defined as an area where at least 500 people and/or 33 percent of inhabitants reside beyond 1 mile (for urban areas) or 10 miles (for rural areas) from a food store. As explanatory variables, an indicator for urban sprawl – defined as an index of eight variables, including population density and employment density – along with socioeconomic variables, such as poverty, education and income, were used. A sample of 30 337 census tracts was studied, and the probability of a census tract being a food desert was estimated using data from 2014. The findings from this study indicated that urban sprawl was one of the main determinants of the presence of food deserts. In other studies, such as that of Dutko et al. (2012), the same definition of a food desert as in Hamidi (2019) was used, and explanatory variables such as population density, income and level of education were included to explain the likelihood of an area being a food desert in 2000. In this study, a low population density was a significant contributor to the probability that an area was a food desert.

Food stores are also included as subcategories in studies of threshold demand analysis. These are studies that are concerned with estimating the minimum size of demand needed to maintain the presence of one or more facilities in a location. Mulligan et al. (1985) studied the determinants of the number of retail stores in 20 Arizona communities in 1981. The variables included were the population, share of men, share of the population below 16 years of age, share of the population above 65 years of age, average income and share of the population of Hispanic origin. It was found that population size and the shares of individuals below 16 and above 65, men and Hispanics had a positive effect on the number of food stores, while income had a negative effect.

Harris et al. (1996) studied the number of retail firms in 2,126 communities in the US in 1986. In this study, they used the population size to proxy demand and the proximity to metropolitan regions to capture spatial interdependencies. They found that the population size and income per capita of a community had a positive effect on the number of grocery retailers, while the proximity to metropolitan areas had a negative effect. Wensley and Stabler (1998) examined the threshold demand for a dozen business activities at approximately 600 centers in rural Saskatchewan, Canada, in the 1990s. In this study, it was also found that the population size of the centers and the proximity to urban areas were important determinants explaining food retail store location.

Mushinski & Weiler (2002) studied a broad variety of retail stores in nonmetropolitan places in eight states in the western US using census data in 1992. Population size had a significant effect on the number of grocery stores, while proximity to metropolitan areas and income were insignificant. Similar results were found in Thilmany et al. (2000).

Chakraborty (2012) studied the threshold demands for a variety of retail stores in 2,201 counties in the US in 2000. Additionally, in this study, the location in relation to urban areas, such as larger towns and cities, was an important explanatory variable, where the proximity to metropolitan areas had a negative effect on the number of food stores. Other variables that were found to have a negative influence are education, poverty and income. Population was found to have a positive but declining effect on the number of food stores, which was captured by a quadratic term. These studies, however, all consider food store presence as a binary or count variable within a predefined administrative area; thus, the edge effect, as referred to previously, is also problematic here. Another caveat with these studies, as with the food desert studies described in the previous section, is that the change in food store access by proximity and the relationship with population density over time are not examined.

5. Data

The data used in this study are obtained from Statistics Sweden's individual and firm-level register database. The analysis is conducted on the neighborhood level and is based on church territories (parishes) from 1999, the year the church was separated from the state, and the neighborhoods remain fixed during the study period. The use of these neighborhoods is an advantage because they have long historical roots and are therefore likely to reflect the boundaries of traditional movement patterns for their inhabitants. Since consumers' willingness to travel is low for lower order goods, such as groceries, it is likely that a parish may represent a natural market area for a food store. The precision of the coordinates of firms and individuals provided by the database, however, are for reasons of confidentiality limited to 1*1 km squares in rural areas and 250*250-meter squares in urban areas. This means that a location may be up to 1.4 km (the hypothenuse of a 1*1 km square) away from the true location. Since this introduces noise in the distance measurements, changes that are of smaller magnitude must be interpreted with caution.

Using geocoded locational data on individuals (over the age of 16)5 and firms, the distance to the nearest food store for each individual inhabitant is measured and the average and median of these are calculated for each of the 2511 neighborhoods in 2000 and 2013. Thus, for each individual i in neighborhood6 n, the distance d to the nearest food store is calculated at time t: d_{int}. Next, the neighborhood average distance ($\bar{d}_{nt}$) (4), and the median distance ($\hat{d}_{nt}$) (5) is calculated for each neighborhood n at time t.

$$\bar{d}_{nt} = \frac{\sum d_{int}}{n_{nt}} \tag{4}$$

where n_{nt} is the size in terms of the population of neighborhood i at time t.

$$\hat{d}_{nt} = \begin{cases} distance\ of\ the\ middle\ number, if\ n\ is\ an\ odd\ number \\ average\ of\ the\ two\ middle\ numbers \end{cases} \tag{5}$$

5 All individuals over the age of 16 are included in the register data.

6 There are 2,511 neighborhoods included in the analysis of a total of 2523 neighborhoods. Due to missing information on population in the remaining 12 neighborhoods, these are excluded from the analysis.

In a similar fashion, the socioeconomic and demographic variables used in the analysis are calculated. Following Amcoff (2017), the definition of a food store used in this study is an establishment that is categorized as Standard Industrial Classification (SIC) 5211, the Swedish equivalence to the European NACE standard. This code is defined as "Retail sale in non-specialized stores with food, beverages or tobacco predominating". The code is valid between 1992 and 2007. For the period between 2008 and 2013, the code SIC 4711 is used. Establishments that provide foods in specialized stores are not included in the analysis. These stores are typically smaller and more specialized, such as butchers and bakeries, and they are excluded because these types of stores, by definition, do not guarantee the complete assortment of food products required for a household and are not relevant for the purpose of this study.

In Table 1 below, descriptive statistics of the stock of food stores on the neighborhood level in Sweden at two points in time are presented. The variables are the median and average distance of the inhabitants in a neighborhood to the nearest food store (Median dist., $\hat{d}_{nt}$; Average dist., $\bar{d}_{nt}$), size of the nearest store (the store to which the median distance is the lowest; Nearest store size) in terms of employees, the number of stores in the neighborhood (No. of stores) and their average size (Average store size) in terms of employees. Moreover, the characteristics are measured in areas that are classified as either rural (Rural area), urban (Urban area) or metropolitan (Metropolitan area). A neighborhood is rural if it is located in a municipality in which more than 50 percent of the population is located in rural areas; it is urban if it is located in a municipality in which less than 50 percent of the population lives in rural areas. Metropolitan areas are neighborhoods located in municipalities with less than 20 percent of their population in rural areas. These definitions are provided by Growth Analysis (2014).

Table 1. Neighborhood-level distance to the nearest food store in 2000 and 2013.

	All: Obs. 2,511			Rural area: Obs.1,000			Urban area: Obs 1,318			Metro. Area: Obs.193		
Year 2000	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Median dist.	3.32	0.00	41.67	3.86	0.00	41.67	3.26	0.25	28.89	0.94	0.00	7.25
Average. dist.	3.60	0.02	40.73	4.25	0.33	40.73	3.49	0.20	26.21	1.02	0.02	7.02
Nearest store size	9.64	1.00	192.00	8.93	1.00	71.00	9.89	1.00	192.00	11.69	1.00	69.00
Average store size	5.44	0.00	100.63	4.51	0.00	208	5.37	0.00	155.00	10.84	12.13	104.00
No. of stores	2.12	0.00	36.00	1.21	0.00	14.00	1.95	0.00	34.00	8.05	0.00	36.00
Year 2013	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Median dist.	3.70	0.00	28.62	4.24	0.00	28.62	3.70	0.25	21.69	0.99	0.00	7.98
Average. dist.	4.00	0.04	26.93	4.63	0.10	25.97	3.95	0.15	26.93	1.05	0.04	7.61
Nearest store size	13.84	1.00	188.00	12.22	1.00	174.00	14.53	1.00	188.00	17.48	1.00	144.00
Average store size	6.81	0.00	100.63	5.48	0.00	89.00	6.72	0.00	202.00	14.20	0.00	126.33
No. of stores	2.04	0.00	35.00	1.08	0.00	13.00	1.84	0.00	26.00	8.42	0.00	35.00

In the year 2000, people living in rural areas had to travel an average of 4.25 km to the nearest food store, people residing in urban areas had to travel 3.5 km, and people in metropolitan areas faced a distance of 1 km. A similar relationship holds for the median distances. Examining the average size of the closest food store, people in rural neighborhoods had a store with, on average, 9 employees, and people in urban and metropolitan neighborhoods had stores that, on average, had 10 and 12 employees, respectively. The average store size was similar in urban and rural areas and had approximately 5 employees, while it was twice that, 11 employees, in the metropolitan areas. People living in rural neighborhoods had, on average, 1 store in their neighborhood, people in urban neighborhoods had 2 stores, and people in metropolitan neighborhoods had 8 stores.

Between 2000 and 2013, the total number of stores in Sweden decreased by 600, from 5838 to 5238. In this period, there was an increase in the population's distance to the nearest food store by 30 to 400 meters in metropolitan areas, urban areas and rural neighborhoods. Thus, the increase was of a relatively marginal size, and the internal relationship between the three categories of neighborhoods remained largely the same. In contrast, the maximum distance to the nearest food store decreased substantially for the inhabitants in rural areas, from 41 km to 26 km. When examining the data more closely, this change is driven by only one neighborhood and can therefore not be considered a general development trend. When examining the number of stores per neighborhood type and the average store size, stores have become somewhat fewer and larger in both rural and urban areas, while the stores in metropolitan areas have become larger and more. This indicates that economies of scale have steered the development towards larger and fewer stores in rural and urban areas. However, as the average distances from the population to the stores have not increased much, this also means that the population has become more concentrated in space. Based on these figures –it seems that the situation has not become substantially worse.

5.1. For whom has the distance changed?

What are the socioeconomic characteristics of the population in the neighborhoods that have experienced a change in the distance to the nearest food store? In the literature on food store access, it is argued that a long distance to the nearest food store is especially problematic for individuals who are already facing additional barriers to obtaining groceries (Ver Ploeg et al., 2009). These may be groups that have a low income or are unemployed and, hence, are more vulnerable to the transportation costs that a distant food store requires. Single households with children < 18 years old may be faced with more severe time constraints, which are exacerbated by a remotely located food store. Last, individuals with disabilities that decrease their mobility are also more vulnerable to the implications of low physical access to groceries. To capture groups that are vulnerable to poor physical access to groceries, I therefore include variables that show low income, low employment, single-parent households and those aged over 80 years old, measured in 2013. The share of the population that is low-income earners is measured as the share of the neighborhood population that has an income that is below 50 percent of the national median worker's income. Due to data constraints it is not possible to measure unemployment on individual level and therefore I use share of employment instead. The share of the population that is employed is measured as the neighborhood population between the ages of 20 and 74 that is registered as employed in the current year. The share of single parents in the population is proxied by the share of the population in the neighborhood that do not have a registered partner and with at least one child that is below the age of 18 living at home. This is a crude measure as individuals that live with a child may live together with an unregistered partner thus one must be cautious with the interpretation. The share of the population over the age of 80 is the share of the neighborhood population that is above the age of 80 years old. Following Amcoff (2017), the neighborhoods are categorized based on the size of the changes in the inhabitants' distance to the nearest store. There is one category of neighborhoods whose population has experienced an increase in the distance to the closest store that is above 400 meters (~ 5 min walking distance), one for a decrease in distance that is below -400 meters, and a category for neighborhoods that are within this interval. The results are found in Table 2 below.

As shown in Table 2, the population in 1459 out of the 2511 neighborhoods experienced neither a positive nor a negative change in the average distance to the nearest food store that is larger than 400 meters. The share of elderly people above 80 years old were 1 percentage point higher in the categories that had experienced an increase and a decrease. The share of the population that is employed is 1 percentage point higher in the neighborhoods that experienced a decrease. Low-income earners were similar in all three categories of neighborhoods, and the same is true for the share of single parents. The share of low-income earners is 1 percentage point lower in the areas that experienced a change of less than 400 meters. However, overall, there are only minor differences in socioeconomic status between the three categories.

Table 2. Characteristics of neighborhoods that experienced change in the average distance to the nearest food store.

	Increased 2000 - 2013					+0.4 km < Ch. in distance > -0.4 km					Decreased 2000 - 2013				
	Obs	Mean	Std. Dev	Min	Max	Obs	Mean	Std. Dev	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Change in average dist. (km)	717	2.61	2.50	0.40	20.68	1,459	0.03	0.16	-0.40	0.40	335	-2.74	3.37	-26.92	-0.40
Median dist. (km)	717	6.10	3.78	0.35	28.62	1,459	2.52	2.66	0.00	16.71	335	3.73	3.15	0.00	18.06
Average dist. (km)	717	6.33	3.49	0.83	26.93	1,459	2.80	2.51	0.04	16.16	335	4.22	2.86	0.26	17.60
Share of pop. that is employed	717	0.74	0.07	0.49	0.99	1,459	0.72	0.07	0.36	0.93	335	0.73	0.07	0.50	0.91
Share of pop. that is low income earners	717	0.29	0.05	0.16	0.64	1,459	0.28	0.05	0.15	0.63	335	0.29	0.06	0.17	0.50
Share of pop. that > 80 years	717	0.06	0.03	0.00	0.17	1,459	0.06	0.03	0.00	0.16	335	0.06	0.03	0.00	0.16
Share of pop. that is single parents	717	0.03	0.01	0.00	0.14	1,459	0.03	0.01	0.00	0.10	335	0.03	0.01	0.00	0.09

Table 3. Characteristics of neighborhoods with changes in average distance in 5 km intervals.

Change:	< - 10 km		-5 to -10 km		-0.4 to -5 km		+ 0.4 to +5 km		+5 to +10 km		> 10 km	
	Obs.	Mean	Obs.	Mean	Obs.	Mean	Obs.	Mean	Obs.	Mean	Obs	Mean
Share of pop. that is employed	14	0.64	35	0.70	286	0.74	613	0.74	89	0.71	15	0.69
Share of pop. that is low income earners	14	0.38	35	0.33	286	0.28	613	0.29	89	0.31	15	0.31
Share of pop. that is > 80 years	14	0.10	35	0.07	286	0.06	613	0.06	89	0.07	15	0.09
Share of pop. that is single parents	14	0.03	35	0.03	286	0.03	613	0.03	89	0.03	15	0.04

Examining the above Table 3, the population in the neighborhoods that experienced the largest increases in average distance to the nearest food store, greater than 10 km, had a higher employment rate, 69 percent, than the areas that experienced the largest decrease in average distance, with an employment rate of 64 percent. The share of low-income earners in the population in the neighborhoods that experienced the largest decrease in average distance was 38 percent, compared to 31 percent in the areas with the largest increase. The same is true for the share of people over the age of 80, which was 10 percent in the neighborhoods with the largest decrease in average distance, compared to 9 percent in the areas with the largest increase in average distance.

The variable capturing the share of single parents was 3 percent in the areas with the largest increases in distance and 4 percent in the areas with the largest decreases in average distance. If anything, the areas that were best off were those that had a generally higher level of disadvantaged socioeconomic characteristics. The neighborhoods whose population experienced decreases in distance, -5 km to -10 km, versus -5 km to -10 km, were relatively similar. However, the neighborhoods where the population had experienced an increase in the distance to the nearest food store were socioeconomically better off than those that experienced a decrease in distance. The neighborhoods experiencing changes of -0.4 to -5 km versus those experiencing changes of +0.4 to +5 km had similar shares of socioeconomically disadvantaged groups.

The above analysis indicates that neighborhoods where the distance increased were similar to those where the distance decreased in terms of socioeconomic characteristics. In the cases where there were differences, however, there was no consistent pattern indicating whether the disadvantaged tended to experience larger increases in distances. If anything, it was the more affluent areas that experienced an increase.

As a last step, the neighborhoods with the lowest and highest percentiles of the socioeconomic variables are analyzed in terms of the sizes of the changes in distance between 2000 and 2013. In Table 4, the results are shown.

Table 4. The top (90%) and bottom (10%) deciles of socioeconomic variables and (mean of) average distance to the nearest food store in 2013 and the change in (the mean of) in this distance 2000 - 2013. Standard deviations within parentheses.

Dis/advantaged neighborhoods (in 2013)	Percentiles	Obs.	Average dist. (km) 2013	Average change in dist. (km), 2000-2013
Share of pop. > 80 years	p90	251	3.49 (3.68)	+0.26 (2.84)
	P10	252	4.20 (2.54)	+0.14 (1.33)
Share of pop.	p90	251	5.79 (2.73)	+0.43 (1.89)
that is employed	P10	252	3.25 (4.54)	-0.01 (3.97)
Share of pop.	p90	251	4.18 (3.19)	+0.47 (2.95)
that is single parents	P10	252	6.31 (3.30)	+0.56 (3.10)
Share of pop.	P90	251	5.12 (3.67)	+0.60 (3.13)
that is low income earners	P10	251	3.58 (2.62)	+0.28 (1.44)

When examining the results in Table 4, the picture changes somewhat. For the variable share of low income earners, the most socioeconomically disadvantaged areas do face a longer distance (1.5 km) to the nearest food store than the less disadvantaged areas. The opposite holds for the areas in the top decile of the shares of individuals over the age of 80, single parents and employment (the relevant variable is here the bottom decile), which have a shorter distance to the nearest food store. In general, the areas with the top deciles of socioeconomically disadvantaged groups over 80 years old and low income earners have seen a larger increase in the distance to the nearest food store than their counterparts. For the variables share of population over 80 years of age and share of low income earners, the ratio between the top and bottom deciles is close to 2. However, the magnitude of the differences in the changes in average distance between the top and bottom deciles is between 100-300 meters, thus, the magnitude of the differences is marginal.

The definitions of the variables and descriptive statistics for all the variables that are included in the models, along with a correlation matrix, can be found in the appendix, Tables A1, A2 and A3.

6. Empirical design

As observed in Figure 2a to 2c, Sweden has large variations in population density, and over the study period, there has been a continued migration towards the urban and metropolitan areas of the country. How has this affected food store access by proximity? To examine how changes in population density may explain the development of food store access over time, the spatial and temporal dynamics of the change in food store access are modeled.

A regression analysis is conducted where the distance to the nearest food store is used to capture food store access. In the theory section, it was argued that a change in the threshold distance is the effect of closures and new entries. However, since the distance measure that was used in the descriptive analysis in section 5, average distance $\bar{d}_{nt}$, is measured as the average distance for individuals in neighborhood n at time t, a change in $\bar{d}_{nt}$ may also indicate that inhabitants have moved between one year and the next. Therefore, to limit the changes in the dependent variable to those due to changes only in the store location, this part of the analysis is conducted using the distance from the neighborhood midpoint to the nearest food store as the dependent variable. Using a GIS program, the centroid coordinates for each neighborhood[7] are calculated, and the distance to the nearest food store from this point is measured annually and represented by d_{nt}^{G}, where G is the centroid notation. The model for the threshold distance, which was derived in the theory section, is presented below in a general formulation, with n indicating the neighborhood and t indicating the year:

$$D_{nt} = f(F_{nt}, t_{nt}, \rho_{nt}) + g(\boldsymbol{X}) \qquad (7)$$

In order to linearize the model, the dependent variable and the independent variables are expressed in logarithms. D_{nt} is proxied by $ln.\,d_{nt}^{G}$, the log of the distance from the centroid of each neighborhood to the nearest food store, in neighborhood n, at year t. The fixed costs, F_{nt}, can be proxied by the neighborhood-level log of average income, $ln.\,ave.\,inc_{nt}$, as a higher level of income in an area can be argued to be correlated with a higher rent. This variable may also capture the purchasing power of the population living in the neighborhood; therefore, the expected effect of an increase in average income may be of two kinds. Thus, since income captures land rents, it may

[7] The centroid is the center of a two-dimensional figure, here, the neighborhood polygons.

have a positive effect on lnd_{nt}^G because higher rents deter retailers with lower order goods, since such retailers have a lower ability to pay a higher rent, as argued by Garner (1966) and shown by Des Rosier et al. (2009). It may, however, also have a negative effect on lnd_{nt}^G since a higher income indicates higher purchasing power, and because groceries can be considered a normal good in Sweden (as shown in Lundberg & Lundberg, 2012) a higher income translates into a higher demand for food which may attract food retail stores to an area, thus making the distance to the nearest food store smaller.

Consumer density, ρ_{nt}, is proxied by the log of population density per square km in each neighborhood, $ln.\rho op.dens_{nt}$. According to the theoretical model this variable is expected to have a negative correlation with distance to the nearest food store. This is also indicated in previous research (e.g. Lundberg & Lundberg, 2010) which has shown that the proximity of a food store to the household address is an important determinant of Swedish households' choice of grocery store and hence residential population density of the neighborhood is expected to have a negative effect on the distance to the nearest food store. Based on theory of economies of scale and previous research (Parr & Denike, 1970), this effect is expected to be negative and nonlinear, thus that food store distance decreases but at a decreasing rate as population density increases. This is captured by its logarithmized functional form. $\boldsymbol{X}$ is a vector with two additional variables. The first of these is demographic composition. Demographic composition is captured by the share of families with children in the area, $sh.children_{nt}$. This variable is included to account for the fact that neighborhoods may have a similar income profile but still differ in terms of demand due to demographical composition, hence making the area more or less attractive for a certain type of retail. Lundberg and Lundberg (2012) finds that Swedish household expenditures on groceries is positively correlated with the number of individuals in a household that are below 19 years of age. Thus, share of households with children at home should capture a higher demand for food retail, which should attract food stores to the area. Therefore, a higher share of individuals that have at least one child living at home is expected to have a negative effect on the dependent variable.

One of the longstanding critiques of the assumptions of CPT is that individuals do not shop at the store that is closest to their homes but in concurrence with other activities, such as traveling from work (e.g., Fingleton, 1975). To account for the possibility that this may influence store location, the second variable included in X is a variable that captures daytime population density. This is included as the log of the density per square km of the number of people who work in the neighborhood.

Many studies of the determinants of retail location also include the level of unemployment and level of education of the population (e.g., Chakraborty, 2012). However, these are not included in the present analysis since much of their effect is likely to go through the income variable. The transport costs, t, may vary substantially over the Swedish population, which makes it an important determinant of store location. This variable is difficult to capture with the data available. However, this variable is likely to change slowly over time, and thus, the use of neighborhood and time fixed effects can be expected to alleviate the problem. The two-way fixed effect estimation also accommodates other time-invariant aspects of the units of analysis, such as access to road networks, habits, topography, country borders and public transport access.

According to CPT, the size of a market is correlated with the number and order of retail functions of the market. Thus, the size of a market is indirectly correlated with its relative place in the central place hierarchy. This means that there is an interdependence between central places and thus, a change in market size may be correlated with a change in the number and order of the functions in other central places. Thus, a change in in population density in one neighborhood may potentially set off a chain of changes in all neighborhoods, triggering a process of adjustments towards a new equilibrium (Lesage, 2014). Thereby a model that incorporates this dependence may be motivated. I use a Pesearan (2004) cross sectional dependence (CD) test to test for spatial dependence between the neighborhoods. This test is specifically designed for panel data with large N and small T (De Hoyos & Sarafidis, 2006), and as such it is appropriate for this dataset.8 The test shows that there is significant cross sectional dependence between the units. Based on the test and on theory, a model that accounts for global spillovers is justified. I follow the "classical" stepwise approach as defined by Burridge (1980) and Anselin (1988) and recommended in Florax et. al. (2003). This involves the estimation of an OLS model and calculation of a Lagrange Multiplier (LM) statistic for a spatially lagged error term (LME) and spatial lag of the dependent variable (LML). The stepwise approach states that the test that produces the largest LM statistic is the favored model. I use the panel

[8] In this study N=2511 neighborhoods, and T=14 years, thus N is large relative to T.

version of the classical and robust (Anselin et al, 1996) LM-test (RLM). I also include within-effects, and all the tests point towards the spatial autoregressive model (SAR), where a spatial lag of the dependent variable (y) is included. The SAR is a specification where the endogeneity that is due to the spatial dependence between the dependent variable in neighborhood n and neighborhood j, the global spillovers, are accommodated. The specification can be expressed as follows:

$$y = \lambda \boldsymbol{W} y + \boldsymbol{X\beta} + \epsilon, \epsilon \sim N(0, \sigma^2 \boldsymbol{I}) \qquad (8)$$

where λ is the coefficient of the spatially lagged dependent variable, $\boldsymbol{W}$ is a spatial weights matrix. Due to the nature of the neighborhoods, being irregularly shaped and sized spatial units, a row-standardized "Queen" contiguity spatial weights matrix is used. This matrix allows for spatial dependencies between neighborhoods that share borders[9]. $\boldsymbol{\beta}$ is a vector of coefficients for the explanatory variables, $\boldsymbol{X}$. The model in equation (7) is transformed into:

$$ln.d_{nt}^{G} = \lambda \boldsymbol{W}\, ln.d_{jt}^{G} + \beta_1 ln.ave.inc_{nt} + \beta_2 sh.children_{nt} + \beta_3 ln.pop.dens_{nt} + \beta_4 ln.day.pop.dens_{nt} + a_n + year_t + \varepsilon_{nt}, \qquad (9)$$

where

$$\boldsymbol{W} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & \cdots & \cdots & w_{nn} \end{bmatrix}, \qquad (10)$$

where

$$w_{nj} = \begin{cases} 1, & if\ n\ and\ j\ are\ contigous \\ 0, & otherwise \end{cases} \qquad (11)$$

[9] The SAR specification is robust to alternative spatial weight matrix specifications, k=3, k=4 and k=6 nearest neighbors.

where $w_{nj} = 1$ if neighborhood n is adjacent to neighborhood j, otherwise 0. λ is the coefficient of the spatially lagged variable, centroid distance to the nearest food store, β_1 to β_4 are the coefficients of the explanatory variables and ε_{nt} is an error term with an expected mean of zero and constant variance. Last, the model has an individual- and a time-specific fixed effect, denoted a_n and $year_t$, respectively. The spatial model is fitted using the R package splm (Millo & Piras, 2012), and the model is estimated by the two-step maximum likelihood method. The estimation is run in its reduced form, a nonspatial fixed effects model, and in its augmented form, a spatial fixed effects model with spatially lagged dependent variables10.

7. Results

In Table 5, the results from running the regression model are presented in the nonspatial form (FE) and spatial form (SAR). When examining the first column, the fixed effects estimation containing only the main variable of interest, population density, the coefficient 0.21 of $ln. pop. dens_{nt}$ is significant and negatively correlated with the distance from the centroid to the nearest food store. As the dependent variable is also logarithmically transformed, this coefficient can be interpreted as an elasticity. Thus, a one perce increase in population density decreases the distance to the nearest food store by 0.21 percent. The relationship is negative and significant and indicates that, ceteris paribus, an increased demand density decreases the threshold distance, i.e. the required market size for a firm to break even. This is in line with CPT and the expectations from the model. It is also in line with previous studies on threshold demand analysis (e.g. Mulligan et al, 1985; Chakraborty, 2012).

When adding the other variables, the magnitude of the coefficient decreases somewhat to 0.18, but it remains significant and negative. Based on previous studies the expected effect of average income was inconclusive; thus, it could influence the dependent variable both positively and negatively. The correlation between $ln. ave. inc_{nt}$ and the distance from the centroid to the nearest food store is in this model positive and significant with an elasticity of 0.154 percent, which indicates that food stores may be deterred from areas with higher average income, as these may have

[10] Hausman test for the non-spatial and the spatial models are conducted and both tests indicates that the appropriate models contain neighborhood specific fixed effects.

higher land rents. This supports previous findings by, for instance, Chakraborty (2012), where the number of food stores was lower in high-income areas.

The demographic composition, $sh.children_{nt}$, in terms of the share of households that have children living at home, is significant and negatively correlated with the distance to the nearest food store with a coefficient of 0.382. A one unit increase in the share of residents that have at least one child under 18 years of age living at home is correlated with a decrease in distance of 0.382 percent. This result is in line with Lundberg and Lundberg's (2012) findings of a positive correlation between grocery expenditures and number of individuals below 19 years of age in the household. Thus, an increased share of households with children is positively correlated with presence of grocery stores. Lastly, the variable that captures the density of the daytime population is added, and there is no significant effect of this variable.

When examining the SAR model, one must be cautious in directly interpreting the first set of results (in column 3) because they cannot be interpreted in the same manner as the partial derivatives in the nonspatial FE model discussed above. It is interesting to note, however, that the significant effect on the spatially lagged dependent variable, $Wln.d_{jt}^{G}$, is positive, which means that there is a clustering of the values of the dependent variable. Thus, neighborhoods that have a shorter distance to the nearest food store are likely to be located close to other neighborhoods that also have a shorter distance to the nearest food store.

Table 5. Regression results of the nonspatial fixed effects (FE) model and the spatial (SAR) fixed effects model. The dependent variable is the neighborhood-level centroid distance to the nearest food store ($ln.d_{nt}^{G}$).

Variables	FE[a]	FE	SAR[bc]	DE	IE	TE
$ln.pop.dens_{nt}$	-0.210**	-0.180**	-0.176***	-0.176***	-0.00453**	-0.181***
	(0.087)	(0.087)	(0.038)	0.0381	0.00192	0.0391
$ln.ave.inc_{nt}$		0.154**	0.153***	0.153***	0.00403**	0.157***
		(0.076)	0.048	0.0479	0.002	0.0492
$sh.children_{nt}$		-0.382***	-0.379***	-0.379***	-0.00996**	-0.389***
		(0.109)	0.071	0.0709	0.00398	0.0728
$ln.day.pop.dens_{nt}$		-0.011	-0.011*	-0.0109*	-0.000288	-0.0112*
		(0.010)	0.006	0.0062	0.00020	0.00639
$W ln.d_{jt}^{G}$			0.0256***			
LME	47.044***	9.068***				
Robust LME	77.444***	12.28***				
LML	48.585***	10.522***				
Robust LML	78.985***	13.734***				
Global PCD test	913.39***	858.25***				
Local PCD test	43.999***	41.315***				
R-sq within	0.015	0.016	0.014			
BIC	20829.32	20817.9	20694			
AIC	20710.77	20673.95	20643.2			
Obs	35,154	35,154	35,154	35,154	35,154	35,154
No. of neighborhoods	2,511	2,511	2,511	2,511	2,511	2,511

*** = significant at the 1% level; ** = significant at the 5% level; * = significant at the 10% level. [a]Heteroscedasticity and serial correlation robust standard errors in parentheses (White, 1980; Newey & West, 1987). [b]Results are robust to estimations using Driscoll and Kraay (1998) standard errors. These standard errors are robust to general forms of spatial and temporal dependence and to heteroscedasticity. [c]Results are robust to alternative spatial weight matrix defined as k=3, k=4 and k=6 nearest neighbors. Included but not reported: time- and neighborhood fixed effects.

The results of the spatial effects are recalculated into three types of marginal effects. Direct effects (DE), which are the effects that a change in an independent variable in neighborhood n has on the dependent variable in the same neighborhood. Indirect effects (IE) are the effects that a change in an independent variable in neighborhood n has on the dependent variable in the adjacent neighborhood j. These effects may be interpreted as spillover effects. The total effects (TE) is the sum of the indirect and the direct effects and thus represent the average effects on the dependent variable. As the relationship between the dependent and independent variables is endogenous, the use of the word "effects" in the following text does not indicate causality.

When examining the marginal effects of the SAR-estimation (columns 4-6), population density has a direct correlation of -0.176 and an indirect correlation of -0.00453 with the distance to the nearest food store. Thus, a one percent increase in the population density in neighborhood n is correlated with a decrease of 0.176 percent in the distance to the nearest food store in the same neighborhood n and a decrease of 0.00453 percent in adjacent neighborhood j. Since there is a spillover effect originating from the adjacent neighborhood, the total effect on neighborhood n is -0.181. It can be deduced that the majority of the effects of a change in population density goes through the direct effects.

In the spatial model, income has a total (indirect and direct effects) positive effect on the distance to the nearest food store, which is of a similar magnitude to the result in the nonspatial model. However, this effect consists of a direct effect of +0.153 percent and an indirect effect of +0.00403 percent, thus indicating a spillover effect on adjacent neighborhoods. Hence, an increase in the average wage in one neighborhood is correlated with a distance increase in adjacent neighborhoods, thus having a negative effect on food store proximity. The share of families with children also has a negative effect in adjacent neighborhoods. This variable has a total marginal effect that is similar to that in the nonspatial model, but there is a significant spillover effect on neighboring areas of -0.00996 percent.

The daytime population density variable was nonsignificant in the nonspatial model, and it is now only weakly significant and remains small also in the spatial model. There is a high correlation between that variable and population density, which may explain that it is insignificant. As a robustness test, the daytime population density variable was included instead of the residential population density while controlling for income and the share of family households. The results (not reported here but available upon request) show that the daytime population variables do not have a significant effect on the dependent variable. Moreover, the inclusion of daytime population density alters neither the magnitude nor the direction or significance of the other coefficients. Hence, the change in food store proximity appears to be driven by population density measured as inhabitants per square kilometer, demographic composition and average income.

The analysis is conducted on the neighborhood level and thus may not hold if the size or shape of the areal unit changes, which is referred to as the modifiable areal unit problem (MAUP; Openshaw, 1984; Fotheringham & Wong, 1991). To assess how the unit of analysis in the present study performs in comparison to alternative specifications, an intraclass correlation (ICC) analysis is performed. This analysis shows at what hierarchical level a certain share of the variation in the dependent variable is explained. In this step, I test the two levels; neighborhoods and municipalities, where the former is nested in the latter. The results are presented in Table 6.

Table 6. Intraclass correlation on two levels of areal units.

Level	ICC	Standard error	[95% Conf. interval]	
			Upper	Lower
Municipality	0.272	0.025	0.225	0.324
Neighborhood	0.900	0.004	0.888	0.903

In total, 27.2 percent of the variation in the dependent variable is explained at the municipality level and 90.0 percent at the neighborhood level. Hence, the variation in the dependent variable is to a considerable degree explained at the neighborhood level. Thus, the neighborhood, as defined in this study, can be theoretically (as discussed in the Data section) as well as econometrically justified.

8. Discussion and conclusion

In many developed economies, the spatial concentration of both economic activity and population has increased over time. The concentration of people in some areas inevitably causes the population to decline in other areas. In areas where the population decreases, the decline in market size makes it difficult for basic commercial services, such as food stores, to remain on the market. When the market size drops too low, these firms may be forced to close down or relocate. In effect, the residents who still live in the area will have to travel a longer distance to obtain the same services, which incurs costs in terms of time and transportation. To maintain the presence of food store access in areas of population decline, EU funding has since 2007 subsidized services in areas where the market is shrinking, and in Sweden, similar subsidies have been in place for over two decades.

The present study examines how physical access to food stores has changed for the Swedish population between 2000 and 2013. It also examines the role of population density in food store access by proximity. This knowledge is relevant for understanding to what extent food store access can be influenced through, for instance, residential planning, and is therefore of importance for policymakers.

The descriptive analysis shows that the number of food stores decreased by 10 percent, or 600 stores, between 2000 and 2013. Moreover, there was an increase of approximately 400 meters in the neighborhood-level average distance from inhabitants' homes to the nearest food store; thus, physical access declined in this period. The individuals who live in the neighborhoods where the largest increases in average distances are measured are more well off – in terms of socioeconomic variables – than individuals in neighborhoods where the distance has decreased the most. These results are similar to those of Amcoff (2017). When examining the physical access for the top and bottom deciles of variables indicating socioeconomically disadvantaged groups, measured in 2013, the most disadvantaged groups (except single parents) have, on average, seen an increase in the average distance to the nearest grocery store. This result is different from the result of Amcoff (2017), where the disadvantaged groups saw a decrease in this distance. The changes in distances are approximately a few hundred meters, and as the precision in the coordinates is limited, changes of this magnitude must be interpreted with caution. Therefore, there is no indication that the most vulnerable parts of the population live in areas that have seen a drastic decline in terms of food store access.

The results from the second part of the analysis revealed that the estimated correlations between the distance to the nearest food store from the centroid and the population density in the non-spatial and spatial specifications differ marginally. The analysis shows that the majority of the total effect of population density goes through the direct effects. The spatially augmented model shows that there are significant – albeit small relative to the direct effects – spillover effects from the changes in population density between neighborhoods. This interdependence between neighborhoods must be taken into account when planning residential areas. The analysis shows that, in addition to population density, demographic composition and average income are significantly correlated with the changes in food store access by proximity.

The results of the model support the findings of previous studies that have analyzed the relationship between food store location and the density of demand (Harris et al, 1996; Chakraborty, 2012, etc.) and thus confirm the previously established positive relationship between food store proximity and the density of demand. The paper complements the literature by illustrating that this relationship also holds after controlling for time-invariant heterogeneity and accommodating spatial dependence.

The results show that increases in population density in a neighborhood are correlated with a decrease in the distance to the nearest food store. This means that areas in Sweden where the population declined between 2000 and 2013 have also seen a decrease in food store access by proximity. This has occurred in the same period that subsidies from the EU and the Swedish government have been focused on maintaining access to basic commercial services for the Swedish population. It is not possible in this analysis to determine whether the subsidies had an effect. It is, however, possible to conclude that the subsidies have not been enough to completely halt the decline in food store access completely. Thus, as long as there is population decline in an area, food stores will continue to disappear.

As discussed in the introduction, the relationship between access to food stores and population density is endogenous; thus, what is captured in this analysis cannot be claimed to be a causal effect. One possibility for future research would be to study the effects of an exogenous decline in population size on food store access for the inhabitants that remain in the area. One such exogenous decline could be the closure of military bases, which for instance is used in Anderson et al (2007) to analyze the effects of the closures on net income growth. This would give indications on the direction of the causality. Another possibility for future research is to estimate demand thresholds in terms of the population density that is necessary to maintain food store access within a certain distance. Utilizing equally sized grids, it would be possible to precisely estimate the population densities required for a certain level of food store access. This could be a valuable tool in the planning of new residential neighborhoods.

Acknowledgements

The author would like to thank (in alphabetical order) Mikaela Backman, Emek Basker, Sven-Olov Daunfeldt, Johannes Hagen, Johan Klaesson, Sierdjan Koster, Agostino Manduchi, Giovanni Millo, Pia Nilsson, Paul Nystedt, Andrés Rodríguez-Pose, Özge Öner, and participants at the Swedish Graduate Programme in Economics (SWEGPEC) Workshop, 2017, the HUI Workshop on Research in Retailing, Mälargården, Bro, Sweden, 2018 and European Regional Science (ERSA) Conference, 2018 and 2019 and for valuable comments that have significantly improved the quality of this paper. Research funding from the Swedish Retail and Wholesale Council (Handelsrådet) and the Institute of Retail Economics (HFI) is gratefully acknowledged.

References

Allcott, H., Diamond, R., Dubé, J. P., Handbury, J., Rahkovsky, I., & Schnell, M. (2019). Food deserts and the causes of nutritional inequality. *The Quarterly Journal of Economics*, *134*(4), 1793-1844.

Alwitt, L. F., & Donley, T. D. (1997). Retail stores in poor urban neighborhoods. *Journal of Consumer Affairs*, *31*(1), 139-164.

Amcoff, J. (2017) Food deserts in Sweden? Access to food retail in 1998 and 2008. *Geografiska Annaler*: *Series B Human Geography 99*: 94-105

Amcoff, J, Mohall M., Waxell A., & Östh J. (2015) Detaljhandelns förändrade geografi. Working Papers in Geography 2015:1 Uppsala University

Anselin, L. (1988). Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity. *Geographical analysis*, *20*(1), 1-17.

Arribas-Bel, D., Nijkamp, P., & Scholten, H. (2011). Multidimensional urban sprawl in Europe: A self-organizing map approach. *Computers, Environment and Urban Systems*, *35*(4), 263-275.

Babey, S. H., Diamant, A. L., Hastert, T. A., & Harvey, S. (2008). Designed for disease: the link between local food environments and obesity and diabetes.

Berry, B.J., Garrison, W. L. (1958) A note on central place theory and the range of a good. *Economic Geography 34*: 304-311

Biagi, B., Faggian, A., & McCann, P. (2011). Long and short distance migration in Italy: the role of economic, social and environmental characteristics. *Spatial Economic Analysis*, *6*(1), 111-131.

Brakman, S., Garretsen, H., & Van Marrewijk, C. (2009). The new introduction to geographical economics. Cambridge University Press.

Burridge, P. (1980). On the Cliff-Ord test for spatial correlation. Journal of the Royal Statistical Society: Series B (Methodological), *42*(1), 107-108.

Center for Disease Control (CDC). Retrieved from: https://wwwcdcgov/features/fooddeserts/. Accessed 2019-04-20.

Chakraborty, K. (2012) Estimation of minimum market threshold for retail commercial sectors. *International Advances in Economic Research 18*:271-286

Christaller, W. (1933) *Central places in southern Germany* (Prentice-Hall Englewood Cliffs NJ) Harts C D 1954

Clarke, I., & Banga, S. (2010) The economic and social role of small stores: a review of UK evidence. *The International Review of Retail Distribution and Consumer Research 20*:187-215

Clarke, G., Eyre, H., & Guy, C. (2002). Deriving indicators of access to food retail provision in British cities: studies of Cardiff, Leeds and Bradford. *Urban Studies*, *39*(11), 2041-2060.

Dawson, J. (2006). Retail trends in Europe. In Krafft, M., & Mantrala, M. K. (Eds.) *Retailing in the 21st Century* (pp. 41-58). Springer, Berlin, Heidelberg.

De Hoyos, R. E., & Sarafidis, V. (2006). Testing for cross-sectional dependence in panel-data models. *The stata journal*, *6*(4), 482-496.

Des Rosiers, F., Thériault, M., & Lavoie, C. (2009). Retail concentration and shopping center rents—a comparison of two cities. *Journal of Real Estate Research*, *31*(2), 165-207.

Dicken, P. & Lloyd, E.P. (Eds.) *Location in space: Theoretical perspectives in economic geography*. Prentice Hall, 1990.

Donald, B. (2013). Food retail and access after the crash: Rethinking the food desert problem. *Journal of Economic Geography*, *13*(2), 231-237.

Driscoll, J. C., & Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics*, *80*(4), 549-560.

Dutko, P, Ver Ploeg, M, Farrigan, TL (2012) Characteristics and Influential Factors of Food Deserts. Washington, DC: US Department of Agriculture, Economic Research Service.

Fingleton, B. (1975). A factorial approach to the nearest centre hypothesis. Transactions of the Institute of British Geographers, 131-139.

Forsberg H (1998) Institutions consumer habits and retail change in Sweden. *Journal of Retailing and Consumer Services* 5:185-193

Fotheringham, A. S., & Wong, D. W. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A*, 23(7), 1025-1044.

Garner B.J., (1966) The internal structure of retail nucleations, North Western Studies in Geography, 12

Glanz, K., Sallis, J. F., Saelens, B. E., & Frank, L. D. (2005). Healthy nutrition environments: concepts and measures. *American Journal of Health Promotion*, *19*(5), 330-333.

Grediaga, I. O. & Freshwater, D. (2010). *Strategies to Improve Rural Service Delivery* (OECD Rural Policy Reviews). Retrieved from: https://www.oecd-ilibrary.org/urban-rural-and-regional-development/strategies-to-improve-rural-service-delivery_9789264083967-en, Accessed 2019-06-06

Growth Analysis (2013) ”Tillgänglighet till kommersiell och offentlig service 2012”. [Access to commercial and public services in 2012]. Tillväxtanalys. Östersund Report 2013/083

Growth Analysis (2014) Better statistics for better regional and rural policy, Report 2014:04. Retrieved from: https://www.tillvaxtanalys.se/download/18.201965214d8715afd16a35b/1432803718514/rapport_2014_04_rev1.pdf (Accessed 2019-10-11).

Guy, C., Clarke, G., & Eyre, H. (2004). Food retail change and the growth of food deserts: a case study of Cardiff. *International Journal of Retail & Distribution Management*, *32*(2), 72-88.

Hamidi, S. (2019). Urban sprawl and the emergence of food deserts in the USA. *Urban Studies*, 0042098019841540.

Harris, T. R., Chakraborty, K., Xiao, L., & Narayanan, R. (1996). Application of count data procedures to estimate thresholds for rural commercial sectors. *Review of Regional Studies*, *26*(1), 75-88.

Hodge, I., Dunn, J., Monk, S., & Kiddle, C. (2000). An exploration of ‘bundles' as indicators of rural disadvantage. *Environment and Planning A*, *32*(10), 1869-1887.

HUI Research (2018) *Handeln i Sverige Snabbfakta*. [Retail trade in Sweden, quick facts] Retrieved from: http://wwwhandelnisverigese/rapporter/snabbfakta/. Accessed 2019-01-01

Ilbery, B., & Maye, D. (2006). Retailing local food in the Scottish–English borders: A supply chain perspective. *Geoforum*, *37*(3), 352-367.

Isard, W. (1956). Location and Space-Economy, New York.

Lamb, K. E., Thornton, L. E., Cerin, E., & Ball, K. (2015). Statistical approaches used to assess the equity of access to food outlets: a systematic review. AIMS *Public Health*, *2*(3), 358.

Larsson, J. P., & Öner, Ö. (2014). Location and co-location in retail: a probabilistic approach using geo-coded data for metropolitan retail markets. *The Annals of Regional Science*, *52*(2), 385-408.

LeSage, J. P. (2014). What Regional Scientists Need to Know about Spatial Econometrics. *Review of Regional Studies*, *44*(1), 13-32.

Losch A (1964) *Economics of Location* (translated by William H. Woghan). New Haven, London

Lucas, K. (Ed.). (2004). *Running on empty: Transport, social exclusion and environmental justice*. Policy Press.

Lundberg, J., & Lundberg, S. (2010). Retailer choice and loyalty schemes—evidence from Sweden. *Letters in Spatial and Resource Sciences*, *3*(3), 137-146.

Lundberg, J., & Lundberg, S. (2012). Distributional effects of lower food prices in a rich country. *Journal of consumer policy*, *35*(3), 373-391.

Millo, G., & Piras, G. (2012). splm: Spatial panel data models in R. *Journal of Statistical Software*, *47*(1), 1-38.

Mulligan, G. F., Wallace, M. L., & Plane, D. A. (1985). A General-Model For Estimating The Number Of Tertiary Establishments In Communities-An Arizona Perspective. *Social Science Journal*, *22*(2), 77-93.

Mushinski, D., & Weiler, S. (2002). A note on the geographic interdependencies of retail market areas. *Journal of Regional Science*, *42*(1), 75-86.

Newey, W.K. & West, K.D. (1987). A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*, 703–08

Openshaw S, 1984 Concepts and Techniques in Modern Geographyy Number 38. The Modifiable Areal Unit Problem (Geo Books, Norwich)

Pesaran, M. H. (2004). General diagnostic tests for cross section dependence in panels.

Parr, J. B., & Denike, K. G. (1970). Theoretical problems in central place analysis. *Economic Geography*, *46*(4), 568-586.

Postnord (2014) *e-barometern* Report. Retrieved from: https://www.postnord.se/siteassets/pdf/rapporter/e-barometern-arsrapport-2014.pdf. Accessed 2018-12-20

Postnord (2018) *e-barometern* Report. Retrieved from: https://www.postnord.se/siteassets/pdf/rapporter/e-barometern-arsrapport-2017.pdf. Accessed 2018-12-20

Rose, D., & Richards, R. (2004). Food store access and household fruit and vegetable use among participants in the US Food Stamp Program. *Public Health Nutrition*, *7*(8), 1081-1088.

Sadler, R. C., Gilliland, J. A., & Arku, G. (2011). An application of the edge effect in measuring accessibility to multiple food retailer types in Southwestern Ontario, Canada. *International Journal of Health Geographics*, *10*(1), 34.

Salop, S. C. (1979). Monopolistic Competition with Outside Goods. *Bell Journal of Economics*, *10*(1), 141-156.

Salyards, D. M., & Leitner, K. R. (1981). Market threshold estimates: a tool for business consulting in Minnesota. *American Journal of Small Business*, *6*(2), 26-32.

Shaw, H. J. (2006). Food deserts: Towards the development of a classification. *Geografiska Annaler: Series B, Human Geography*, *88*(2), 231-247.

Smith, A., & Sparks, L. (2000). The role and function of the independent small shop: the situation in Scotland. *The International Review of Retail, Distribution and Consumer Research*, *10*(2), 205-226.

Social Exclusion Unit (1998). *Bringing Britain together: a national strategy for neighbourhood renewal*. Cm 4045. London: HMSO.

SOU (1972). *Kommersiell Service i Glesbygder*. [Commercial services in rural areas]. Report 1972:13. Retrieved from: https://lagen.nu/sou/1972:13#sid3-img. Accessed, 2019-09-01.

SOU (2015). *Service i Glesbygd*. [Service in Rural Areas]. Statens Offentliga Utredningar. Report 2015:35. Retrieved from: https://wwwregeringense/contentassets/ddacd084293a4b11b2fb26fdd7950208/service-i-glesbygd-del-1-av-2-sou-201535. Accessed 2018-10-10

Swedish Agency for Economic and Regional Growth (2015). *Uppdrag att fördela medel för insatser inom området kommersiell service i gles- och landsbygder. Slutrapport*. [Assignment to distribute funds for efforts in the area of commercial service in sparse and rural areas. Final report.] Report no. 0612. Retrieved from: http://tillvaxtverket.eprint.se/System/TemplateView.aspx?p=E-View&id=ae4fb11594444602bbe28695807755f7&l=t&view=11&cat=%2FPublikationer%2FEndast%20f%C3%B6r%20nedladdning&tc=%2FWeb%2FText%2FTopp_Listvy_App3&page=6

Swedish Parliament (1994) *Förordning (1994:577) om landsbygdsstöd och stöd till kommersiell service* [Regulation (1994: 577) on rural and commercial service support]. Retrieved from: (https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/forordning-1994577-om-landsbygdsstod-och-stod_sfs-1994-577) Accessed 2019-10-20

Swedish Parliament (2000), *Förordning (2000:284) om stöd till kommersiell service* [Regulation (2000: 284) on support for commercial service]. Retrieved from: (https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/forordning-2000284-om-stod-till-kommersiell_sfs-2000-284) Accessed 2019-10-20

The Swedish Board of Agriculture (2019). Retrieved from: http://wwwjordbruksverketse/amnesomraden/landsbygdfiske/programochvisioner/landsbygdsprogrammet2014202047c4ce2e813deda4d30780004608html. Accessed 2018-03-10.

The Swedish Consumer Agency (2008), *Kommersiell service i alla delar av landet - Redovisning av insatser och erfarenheter 2002-2007* [Commercial service in all parts of the country - Accounting of efforts and experiences 2002-2007], Konsumentverket, Report no. 2008:06. Retrieved from: (http://tillvaxtverket.eprint.se/System/TemplateView.aspx?p=E-View&id=3b0129621c0949e7bcc151f361771a1f&l=t&cat=%2FPublikationer%2FEndast%20f%C3%B6r%20nedladdning&tc=%2FWeb%2FText%2FTopp_Listvy_App3&page=65)

Thilmany, D., McKenney, N., Mushinski, D., & Weiler, S. (2005). Beggar-thy-neighbor economic development: A note on the effect of geographic interdependencies in rural retail markets. *The Annals of Regional Science*, *39*(3), 593-605.

Turok, I., & Mykhnenko, V. (2007). The trajectories of European cities, 1960–2005. *Cities*, 24(3), 165-182.

Varela-Candamio, L., Morollón, F. R., & Sedrakyan, G. (2019). Urban sprawl and local fiscal burden: analysing the Spanish case. *Empirica*, *46*(1), 177-203.

Ver Ploeg, M., Breneman, V., Farrigan, T., Hamrick, K., Hopkins, D., Kaufman, P., Biing-Hwan, L., Nord, M., Smith, T. A., Williams, R., Kinnison, K., Olander, C., Singh, A., & Tuckermanty, E., (2009). Access to affordable and nutritious food: measuring and understanding food deserts and their consequences: report to Congress. Administrative Publication No. (AP-036) 160 pp. Retrieved from: https://www.ers.usda.gov/publications/pub-details/?pubid=42729. Accessed 2019-10-10.

Walker, R. E., Keane, C. R., & Burke, J. G. (2010). Disparities and access to healthy food in the United States: A review of food deserts literature. *Health & Place, 16*(5), 876-884.

Wensley, M. R., & Stabler, J. C. (1998). Demand-Threshold Estimation for Business Activities in Rural Saskatchewan. *Journal of Regional Science*, *38*(1), 155-177.
White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: journal of the Econometric Society*, 817-838.
Wiberg, U. (1983). *Service i glesbygd: trender och planeringsmöjligheter*. [Service in rural areas]. (Doctoral dissertation, Umeå universitet).
Widerstedt, B., Bergström, F., Arnberg, J., Blank, M., & Cronholm, M. (2006) *Svensk Digital Handel Detaljhandeln och produktivitetstillväxten. Produktivitetsutvecklingen inom svensk detaljhandel 1993–2004.* [Swedish Digital Trade Retail and productivity growth Productivity development in Swedish retail. Development in productivity, 1993-2004]. Report A2006:002, Swedish Trade Federation. Retrieved from: https://www.tillvaxtanalys.se/download/18.6a3ab2f1525cf0f4f9a70d0/1454413774499/Detaljhandeln+och+produktivitetstillv%C3%A4xten-06.pdf
Woods, M. (2004). *Rural geography: Processes, responses and experiences in rural restructuring*. Thousand Oaks, United States: SAGE Publications.
Wrigley, N. (2002). 'Food deserts' in British cities: policy context and research priorities. *Urban Studies*, *39*(11), 2029-2040.
Öner, Ö. (2017). Retail city: the relationship between place attractiveness and accessibility to shops. *Spatial Economic Analysis*, *12*(1), 72-91.

Appendix A

The main point of this model is to derive a model that shows how the number of sellers increases with population density. In Salop's (1979) circular market model, population density is present, but it is set equal to one. In this modified version, I set the population density $\rho > 0$.

A1. Symmetric equilibrium

I assume there are n independent sellers that operate in a symmetric market. The number of sellers is taken as given and is later endogenized. The sellers are located on a circular market line with a unit length and addresses in the segment [0,1), denoted by $x_i \in (0, D)$. I consider seller S, which is located at 0. Since we assume equal distances between all neighbors, this seller has two neighbors, denoted $\hat{S}$ and $\check{S}$, located at distance D, which I take as given and later endogenize.

A2. Deriving equilibrium prices and profit function

With a given number of firms, consumers are distributed uniformly on the circular market. A single consumer faces a utility function that consists of the value v, minus a constant traveling cost, t, per unit of distance that must be traveled from the consumer's location x to purchase the good for price p_i, at the address of firm i. Seller S is located to the left of the consumer at address 0, and thus, the consumer must travel from x to 0. The consumer therefore incurs a travel cost of $t(x - 0)$. To travel to seller $\hat{S}$, located at the consumer's right, at address D, the consumer incurs the travel cost $-t(D - x)$.

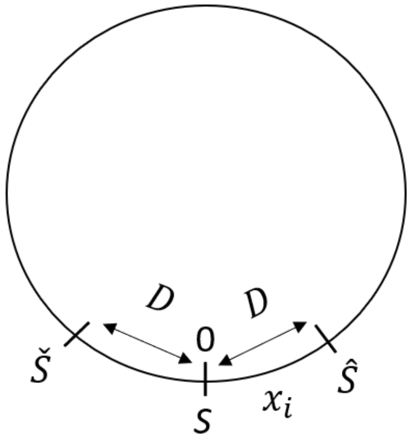

Figure A1. Illustration adapted from Salop (1979).

For seller S, the consumer has the following utility function:

$$U_x(S) = v - t(x - 0) - p_S \quad (9)$$

For sellers $\hat{S}$, the consumer has the utility function:

$$U_x(\hat{S}) = v - t(D - x) - p_{\hat{S}} \quad (10)$$

The address x^*, at which the consumer is indifferent between the two sellers, is found where the utilities of both sellers' goods, (9) and (10), are equal to each other.

$$v - t(x - 0) - p_S = v - t(D - x) - p_{\hat{S}} \quad (11)$$

Solving for x:

$$x^* = \frac{Dt + p_{\hat{S}} - p_S}{2t} \quad (12)$$

The consumers are uniformly distributed over the circle with density ρ, and assuming that the demand is symmetric in both directions from *S*'s location[11], the aggregate market revenues faced by firm *S* can be expressed by:

[11] With a circular market, firm S will face demand both to the right and to the left of its location; thus, the size of the total demand must be multiplied by 2.

$$Q_S(p_S, p_{\hat{S}}) = 2\rho p_S x^* \qquad (13)$$

Substituting in (12) into (13):

$$Q_S(p_S, p_{\hat{S}}) = \frac{\rho p_S(Dt+p_{\hat{S}} - p_S)}{t} \qquad (14)$$

All firms face a fixed cost of production, $F > 0$, and a marginal cost c, which for simplicity is set to equal 0. Thus, firm S has the profit function:

$$\Pi_S(p_S, p_{\hat{S}}) = \frac{\rho p_S(Dt+p_{\hat{S}} - p_S)}{t} - F \qquad (15)$$

As I assume a symmetric market, an analogous profit function holds for all sellers. Differentiating (15) w.r.t. to p_S, the optimal price for seller S can be found:

$$\frac{d\Pi_S(p_S, p_{\hat{S}})}{dp_S} = \frac{\rho Dt+\rho p_{\hat{S}} - 2\rho p_S}{t} \qquad (16)$$

The second-order condition is negative, which shows that this is a maximum:

$$\frac{d\Pi_S(p_S, p_{\hat{S}})^2}{dp_S^2} = -\frac{2\rho}{t} < 0 \qquad (17)$$

As the relationship is symmetric, p_S can be set equal to $p_{\hat{S}}$: $p_S = p_{\hat{S}} = p$.

$$p = \frac{Dt+p}{2} \qquad (18)$$

The equilibrium market price for all firms is thereby:

$$p^* = Dt \qquad (19)$$

If all sellers choose p^*, the profit function of S can be found by substituting (19) into the profit function (15):

$$\Pi_S(p^*) = \frac{\rho Dt(Dt+Dt-Dt)}{t} - F \qquad (20)$$

Of course, the seller *S* could attempt to undercut its neighboring sellers.

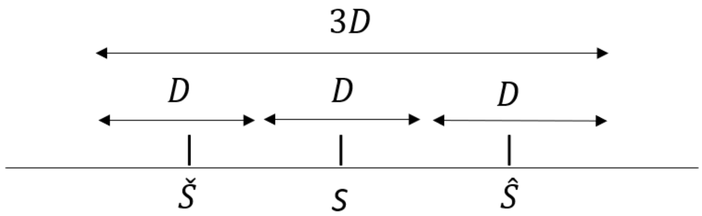

Figure A2. Seller *S* undercutting neighboring sellers and thus capturing a market size that is three times *D*.

To undercut seller $\hat{S}$, *S* must reduce the price by *t*D*. Thus, the new price will be:

$$p = p^* - Dt \qquad (21)$$

Substituting in p^* from (19) into (21) yields:

$$p = Dt - Dt = 0 \qquad (22)$$

A price that is equal to zero yields no profit and is therefore not appealing to seller *S;* therefore, the market size will remain equal to *D*.

Simplifying (20):

$$\Pi_S(p^*) = \rho D^2 t - F \qquad (23)$$

A3. Endogenizing distance

For the firm to be able to stay on the market, the fixed costs must be retrieved because any level of profit that does not cancel out this cost will force the firm to shut down and/or relocate. This free-entry equilibrium occurs at the point in space where equation (21) is equal to zero, which is where profit is equal to fixed costs.

$$\rho D^2 t = F \qquad (24)$$

Solving for the distance at which this occurs gives us:

$$D = \left(\frac{F}{\rho t}\right)^{\frac{1}{2}} \qquad (25)$$

Given that distance is positive, the first-order condition of the positive solution shows that distance decreases in ρ, population density.

$$\frac{dD}{d\rho} = -\frac{F}{2\rho^2\left(\frac{F}{\rho t}\right)^{\frac{1}{2}} t} < 0 \qquad (26)$$

Therefore, it is shown that as population density increases, the distances between the sellers will shrink.

Appendix B

Table B1. Definition of variables, all on the neighborhood level.

Variables	Definition
Median dist.	Median distance from inhabitants' homes to the nearest food store (km).
Average dist.	Average distance from inhabitants' homes to the nearest food store (km).
Distance centroid	Km distance from the centroid of each neighborhood to the nearest food store.
Change in average dist.	Change in average distance between 2000 and 2013
Nearest store employees	The size (in terms of employees) of the nearest food store. Nearest store is identified using the shortest median distance. In the case there are two middle values of the distance, the average size of the two stores connected to those distances are used.
Population	Residents (includes only individuals over the age of 16).
Population density	Residents (> 16) per square km.
No. of stores	Number of food stores.
Average store size	Average number of employees in neighborhood food store.
Change in distance	Change in average distance from inhabitant's address to the nearest food store between 2000 and 2013.
Average income	Average taxable income from owning a business and/or wage employment. In thousands of SEK.
Share of population that is employed	Share of the residents between the ages of 20 to 74 who are registered as employed in the current year
Share of population that is low income earners	Share of the residents (> 16) who have an income that is below 50 percent of the national median worker's income.
Share of population that is over the age of 80	Share of the residents (> 16) who are above the age of 80 years old.
Share of single parents	Share of the residents (> 16) who are living without a registered partner and have at least one child that is below the age of 18.

Table B1 continued.

Variables	Definition
Share of population with children	Share of residents (>1 6) with 1 child or more, who is below 18 years of age and living at home
Rural area	A neighborhood is rural if it is located in a municipality with > 50 % of its residents located in rural areas. Definition according to Growth Analysis[12].
Urban area	A neighborhood is urban if it is located in a municipality with < 50 % of its residents in rural areas (see previous footnote).
Metropolitan area	Metropolitan areas are neighborhoods located in municipalities with < 20% of their residents in rural areas (see previous footnote).
Daytime population density (number of jobs)	Number of people with workplace in the neighborhood, per square km.
$ln.pop.dens_{nt}$	Log of Residents (> 16) per square km
$ln.ave.inc_{nt}$	Log of average income
$sh.children_{nt}$	Share of residents (> 16) with 1 child or more, who is below 18 years of age and living at home
$ln.day.pop.dens_{nt}$	Log of daytime population density

Table B2. Descriptive statistics of the variables in the final model.

Variables	Obs.	Mean	Std. Dev	Min	Max
Ln (Distance (km) centroid)	35,154	1.18	1.07	-3.50	4.46
Ln (Population density)	35,154	1.18	1.07	-3.50	4.46
Ln (Daytime population density)	35,154	1.51	2.27	-13.82	10.96
Ln (Average income)	35,154	7.70	0.19	6.12	8.62
Share of population with children	35,154	0.20	0.06	0	0.57

[12] For more details of these definitions, read more at: https://www.tillvaxtanalys.se/in-english/publications/reports/reports/2014-04-04-better-statistics-for-better-regional-and-rural-policy.html.

Table B3. Correlation matrix of variables included in the full model.

Variables	Ln (Distance (km) centroid)	Ln (Population density)	Ln (Daytime population density)	Ln (Average income)	Share of population with children
Ln (Distance (km) centroid)	1.00				
Ln (Population density)	-0.73	1.00			
Ln (Daytime population density)	-0.72	0.95	1.00		
Ln (Average income)	-0.29	0.42	0.49	1.00	
Share of population with children	-0.004	0.08	0.005	0.03	1.00