Estimating Geographic Market Size Nonparametrically: An Application to Grocery Retailing*

Nicolas Eschenbaum†

This draft: November 20, 2018

Abstract

This paper develops a nonparametric approach to empirically determine geographic market size. I exploit highly detailed spatial data and provide estimates of business-stealing effects across distance by studying the impact of store entry on competitors in an increasing range to the entry site. Entropy balancing is employed to control for systematic differences across local markets. I estimate that markets for Swiss grocery retailing stores are highly localized in a tight four kilometer radius. I further document evidence that the impact weakens with increasing distance and that smaller retailers compete in a more narrow market of only two kilometers in size.

Keywords: Market definition, spatial competition, retail trade, supermarkets

JEL-Classification: L81, L11, L4

*I am grateful to Stefan Buehler, Berno Buechel, Maximilian Conze, Philemon Kraehenmann, Christine Zulehner, and seminar participants at the SSES Annual Congress 2018, the University of Mannheim, the University of St. Gallen, and the Swiss Competition Commission for helpful discussions and comments. The usual disclaimer applies.

†Nicolas Eschenbaum: University of St. Gallen, Institute of Economics (FGN), Varnbüelstrasse 19, 9000 St. Gallen, Switzerland (nicolas.eschenbaum@unisg.ch)

1 Introduction

Antitrust law in numerous jurisdictions, including the EU and the US, requires competition authorities to define the relevant market before proceeding to analyse the case. No matter what is the firm conduct that is to be investigated, market definition is generally considered to be "central to reasonableness analysis" (Baker and Bresnahan, 2008, p. 1). The European Commission for example provided its first notice on merger regulation explicitly on the issue of market definition (The European Commission, 1997), while in the US, the 2010 Horizontal Merger Guidelines continued to emphasize the important role of market definition (Federal Trade Commission and U.S. Department of Justice, 2010). Arguably, "the outcome of more cases has surely turned on market definition than on any other substantive issue" (Baker, 2007, p. 1)

In practice however, empirically determining the relevant market for analysis is challenging and a variety of different methods have been used to delineate markets in antitrust investigations. Some of the more ad-hoc approaches have been criticized in the past (Katz and Shapiro, 2003; Danger and Frech, 2001; Capps et al., 2002) and may compare unfavourably to more structural, empirical antitrust market definitions (Gaynor et al., 2013). But regardless of the method of choice, the focus of the analysis generally lies on determining the degree of demand substitution by consumers between a selection of different products or store locations (Baker and Bresnahan, 2008).

This paper develops a nonparametric approach to empirically estimate geographic markets that abstracts from the difficulty of obtaining estimates of demand substitution. I focus on the business-stealing effect of store entry in order to estimate the extent of local geographic markets in the grocery retail industry without pre-specifying the underlying demand function. The approach rests on the simple idea of contrasting existing retailers that experienced entry to others that did not. To isolate the effect, I control for crucial elements of consumer demand and store competition by exploiting the unique spatial detail of my dataset covering the entire universe of Swiss retailers and resident population geocoded to their exact location with a precision of one square meter. This allows me to estimate the causal effect of entry on the competitive outcome of existing stores that is due to demand substitution by consumers. Studying the effect across an increasing range of distance to the site of entry then allows inferring the extent of the geographic market.

Exploiting the competitive impact of store entry to study geographic market size has been done before in antitrust investigations, in the case of grocery retailing for example by the Bundeskartellamt in Germany on the merger of the supermarket chains Edeka and

Tengelmann (Bundeskartellamt, 2015). Similarly, the study of entry effects and their propagation across distance has been of interest to previous researchers, for example on Wal-Mart stores in the US (Ellickson and Grieco, 2013). Compared to previous approaches, the method developed here allows for entirely nonparametric, causal estimates of market size which do not require or pre-suppose assumptions on market structure or consumer demand that in turn may be relevant for subsequent analysis. In addition, I explicitly tackle the challenge of estimating the business-stealing effects of entry nonparametrically in the presence of possible interdependence of observations. The existence of store competition may violate the usually imposed 'stable-unit-treatment-value assumption' which requires the outcome of each store to only depend on its own treatment assignment, but not on others'. Instead, I require that such possible 'spillover' effects be limited in range and propose a simple approach to test for a violation.

The empirical approach is made possible by the richness and degree of spatial detail available to me in the application on Swiss grocery retailing. The majority of work in the literature on grocery retail competition has made use of information only available administrative level, such as municipalities, and has been forced to define local markets based on this aggregate level (Aguirregabiria and Suzuki, 2016). In contrast, my dataset contains all individual retail stores and all individual residents of Switzerland and each of their precise geographic locations, allowing me to consider the unique characteristics of each store location in much greater detail.[1]

I make use of this highly detailed information to flexibly control for location and store differences and compare the change in the outcome of treated stores (grocery retailers that experienced entry within a given distance) to the corresponding change of control stores (grocery retailers that did not experience such entry). In order to ensure that the estimated effect is due to the entrant under consideration, I exclude all retailers that experienced more than one entry. Because it turns out to be difficult to find control observations that are sufficiently similar to treated ones, I employ the recently developed method of 'entropy balancing' (Hainmueller, 2012).[2] Entropy balancing constructs weights for the control group so that the weighted average becomes as similar as possible to the treatment group of stores that experienced entry.

[1] Additionally, I make use of the precise locations of Swiss customs offices and border posts to proxy for competition by stores located in neighboring countries across the border and consider all distances between stores or to the border as routing distances to account for infrastructure and geography.

[2] Entropy balancing has become increasingly popular and already been applied to a variety of settings, such as online hiring markets (Stanton and Thomas, 2016), subsidized employment (Hetschko et al., 2016), syndicated loans (Amiram et al., 2017), or social capital (Satyanath et al., 2017).

I find sizeable business-stealing effects due to entry within a highly localized four kilometers radius (measured in routing distance) around an entrant. The estimates also show a more pronounced impact of entry by a competing store within the first two kilometers, suggesting that the degree of competition varies by distance within a local market. I further decompose the effect of entry and find strong evidence that the extent of a local market is considerably smaller for small grocery retail stores compared to large ones and is limited to the first two kilometers only. In addition, the estimates suggest that small stores are less able to react to increased competition due to entry and in turn are more likely to exit the market.

This paper contributes to two strands of literature. First, I contribute to the empirical determination of geographic markets. The standard conceptual frameworks to define the relevant market in antitrust investigations are the hypothetical monopolist test (HMT) or the test for a small but significant non-transitory increase in price (SSNIP).[3] In either case, the test iteratively examines the hypothesis that a firm in a successively expanding product or geographic market could profitably impose a price increase. If sufficiently many consumers in response decide to switch to an alternative product or store, the price rise is unprofitable for the firm and it lacks the necessary market power. The relevant market is then expanded iteratively until the price rise becomes profitable. Similarly, the method employed here considers the competitive outcome of retail stores that are (potentially) impacted by entry in successively greater distances from the entrant. I contribute a reduced-form empirical approach that refrains from any structural assumptions and is, to the best of my knowledge, novel to the literature.

Second, I add to the empirical study of market entry and spatial competition in the grocery retail industry. Previous research has highlighted the overall impact of retail chains and their stores on competitors, consumer prices and welfare, or the labour market, among other issues (e.g. Basker, 2007; Ellickson, 2007; Jia, 2008; Ellickson et al., 2013; Holmes, 2011; Neumark et al., 2008; Matsa, 2011; Nishida, 2014; Grieco, 2014). The focus of this paper is on identifying the local effects of store entry and so the econometric approach chosen intentionally avoids any structural modelling framework of retail chain interaction. My approach can be viewed as an early step in the analysis of retail chains. The results of such a nonparametric estimation can be used to directly inform the correct specification of local markets and the choice set of firms in a subsequent structural framework. My analysis contributes to the literature by providing evidence on the extent

[3]For a comprehensive treatment of the tests and applications in EU antitrust investigations see Motta (2004). Coate and Fischer (2008) in turn provide a detailed study of the application and prevalence of the tests in 116 market definition decisions by the US Federal Trade Commission.

of the geographic local market in this industry and the heterogeneity of business-stealing effects within a local market.

The remainder of this paper is structured as follows. Section 2 describes the database and discusses some descriptive statistics. Section 3 sets out the econometric approach, detailing the estimator and identification, as well as describing the final dataset in more detail. Section 4 presents the estimation results, and section 5 provides some robustness checks. Section 6 concludes.

2 Data

My analysis is based on two main data sources. The first is the Swiss Business Census (STATENT), which covers the entire universe of plants in the manufacturing and services sectors registered in Switzerland. The census is collected on a yearly basis since 2011 by the Swiss Federal Statistical Office and participation is mandatory for firms. It provides detailed information on each plant including firm ownership, industry classification, and employment numbers, as well as each plants exact geographic location precise to one square meter. However, as is common with census-type data, the STATENT does not provide information on prices, quantities, or costs of stores. In line with the existing literature on the grocery retail industry, I focus on the impact of entry on store employment instead. The data is available to me for the years 2011-2013. The second source is the Swiss Population and Household Census (STATPOP) for the years 2011-2014, which depicts all persons with permanent and temporary residence in Switzerland annually since 2010. For each individual the census includes information such as age, sex, nationality or marital status, and also the precise geocoordinates for a persons residence with a precision of one square meter. The STATPOP is also collected by the Swiss Federal Statistical Office.

I complement this wealth of information with an additional data source: the locations of all Swiss border posts and customs offices provided to me by the Swiss Customs Department. Switzerland is a small, landlocked country with a particularly high price level relative to its neighbours Germany, France, Italy, and Austria. Consequently, Swiss citizens regularly cross the border in order to shop at significantly lower prices abroad. For example, the Swiss Federal Customs Administration reported revenues of 48.32 billion CHF in 2015 from Swiss citizens shopping tourism (Eidgenössische Zollverwaltung, 2016).[4] The location of customs offices serves as a proxy for international competition by foreign stores that I am unable to observe directly in the Swiss data.

The resulting dataset is unique in its spatial detail, because it covers both plant placement and the spatial distribution of consumers at a very fine level. For the remainder of the paper, I focus on plants and firms in the grocery retail industry exclusively.

Table 1 shows some first, basic summary statistics for the retail store dataset, while Table 8 (in the appendix) shows some comprehensive summary statistics for the population dataset. The Swiss grocery retail landscape consists of around 5,000 stores, or approximately one store per 1,600 inhabitants. There are over 2,000 firms operating the stores and both the overall number of stores and firms decrease during the time period studied. The data also suggests that the market is becoming more concentrated with fewer firms relative to stores over time and the mean number of stores per firm increasing each year. In addition, Table 1 shows that entering stores have a significantly smaller average size than established retail stores as measured by both full-time equivalent (FTE) and total employment. This indicates that store size may be important to understand the success of a retailer, or its reaction to experiencing entry. I make use of the industry code classification provided by the Swiss Federal Office of Statistics, which classifies retailing stores by sales area measured in square meters (see Table 9 in the appendix). As it turns out, the variation in size of retailing stores measured both by sales area or employment numbers is significant.

Table 1: Retail data summary table

	2011	Entry 11-12	2012	Entry 12-13	2013
# Firms	2,242	286	2,147	228	2,035
# Stores	5,105	373	5,036	298	4,894
Mean # Stores of Firms	196.3	396	205.6	383.6	214.7
Mean Employment					
Full-time Equivalent	10.65	3.85	10.27	4.00	10.22
Total	13.93	5.48	13.7	5.44	13.67

Notes: The table documents the total number of firms, total number of stores, mean number of stores per firm, and mean employment number at a store (in full-time equivalent and total, respectively) yearly from 2011-2013 for all grocery retailers, as well as separately for entrants in 2011-2012 and 2012-2013.

In particular, both the absolute employment numbers in 2011 of stores of varying sizes, as well as their yearly growth from 2011-2012, indicate two distinct groups of

[4] Buying abroad also allows Swiss shoppers to obtain a refund on the foreign VAT and pay the (generally lower) Swiss VAT rate instead. To do so, they must show their wares at the local customs office, where the tax income is collected.

retailing stores: i) smaller supermarkets and mom-and-pop stores with NOGA codes of 471105, 471104 and 471103, and ii) large supermarkets and supercenters with NOGA codes 471102 and 471101 (see Table 2). The first group is characterized by significantly lower employment numbers and sales area size, relative to the second, as well as consistently negative yearly growth rates. The bigger stores instead appear to be expanding and growing. I will henceforth refer to these two groups as small and large stores, respectively.

Table 2: Retailers by store classification

		FTE Employment		Total Employment		Fluctuation	
Store Classification	N	Mean	Growth	Mean	Growth	Exit	Entry
Supercenters	76	95.00	0.02	116.50	0.02	0%	0%
Large Supermarkets	356	43.08	0.02	55.00	0.01	3%	1%
Small Supermarkets	1302	12.46	-0.05	17.00	-0.06	7%	6%
Large Shops	2315	4.12	-0.06	6.00	-0.05	5%	6%
Small Shops	1056	0.97	-0.10	1.00	-0.11	20%	14%

Notes: The table documents per category of store (classified according to the BfS industry classification index, see Table 9 in the appendix) in 2011 the total number of stores, mean full-time equivalent (FTE) and total employment, growth in FTE and total employment from 2011 to 2012, as well as the percentage of stores in 2011 exiting and percentage (in proportion of the 2011 total number of stores) of stores entering within the next year.

In addition, the number of small grocery retail stores is significantly higher than that of the large supermarkets. It is important to note, that because of the relatively low number of large stores, I observe very little entry and exit for this group. Instead, it is predominantly the small stores that experience a high degree of fluctuation with a sizeable number of stores entering and exiting each year. This finding is in line with a large body of classic research on firm entry and industry dynamics which demonstrates that the size of an entrant correlates negatively with the likelihood of its survival (Sutton, 1997; Geroski, 1995).

Indeed, the smallest category of stores, comprised of tiny retailers with on average only one full-time employee, experience by far the greatest fluctuation: around 20% of existing stores in 2011 exit over the course of the year and are replaced by new entrants in numbers that constitute 14% of the total number of stores in the previous year. This also shows the main source of the overall decrease in the number of stores documented earlier. However, this does not appear to be a clear pattern across store size, but rather a specific attribute of these smallest stores. For example, the second largest store types experience more exit than entry, while the second smallest show the reverse. In addition, the net fluctuation experienced within the four larger categories of stores is small, with the entry and exit rates showing a clear positive correlation. This pattern is well documented in the literature (Geroski, 1995). For small shops though, the difference between the entry and exit rate is more pronounced.

Table 3: Store growth rate and entry by population of area surrounding store locations.

	<25th percentile			>75th percentile		
	N	Growth	Entry	*N*	Growth	Entry
Small	1227	-0.076 (0.47)	25%	1138	-0.062 (0.41)	33%
Large	37	0.01 (0.05)	0%	126	0.010 (0.082)	60%

Notes: Standard deviations in parentheses. *N* is the number of stores open in 2011. Growth is measured in full-time equivalent employment 2011-2012, while entry is the percentage of entrants in one of the groups. Percentiles shown are of the total population in a 5km radius around the store location.

Finally, I document whether the population in the local area of a store affects the employment growth rates of small and large retailers differently. Table 3 shows the average growth rate of full-time equivalent employment from 2011-2012 for large and small stores respectively by the first and third quartile of the population density at their location. Small stores which are located in less densely populated areas fare worse relative to small stores at more populated locations. In either case, on average small stores reduce their employment hours, but the degree to which stores shrink is more pronounced in less populated areas. On the other hand, large stores tend to neither expand nor shrink when located in areas that are thinly populated, but expand and grow in densely populated areas. In addition, small stores are spread relatively evenly across the population distribution, while large stores are mostly located in the densely populated areas. Table 3 also shows the proportion of entrants of the small or large group respectively that choose to locate where many or few people live. It is evident that firms place large stores predominantly in densely populated areas, while small stores enter more evenly in both highly and very little populated locations.

Taken together, this indicates that there is sizable heterogeneity across store types, and that small stores appear to 'play a different game' than large stores. It appears likely that the demand that stores face differs systematically by store type. The results of the estimations bear this out by clearly indicating that the local market for a small store is smaller than for a large store and that small stores are more strongly adversely affected by competition with other small stores.

3 Empirical Approach

In order to estimate the local effect from entry nonparametrically while allowing it to vary by distance, I make use of standard matching and propensity score techniques. I compare the outcome of retail stores experiencing entry within a particular 'bandwidth' of distance from the store (the treatment group), to retailers that do not experience such entry within the given bandwidth (the control group). I consider bandwidths of a range of two kilometers, so that a retail store may experience entry within 0-2 kilometers, 2-4 kilometers and so on. The maximum range I consider is ten kilometers.[5] The estimated effect is simply the average mean difference of the outcome between the two groups.

To account for the fact that assignment into the treatment group may be non-random (as is clearly the case here), I employ entropy balancing (EB) to preprocess the data and obtain a causal estimate of the effect. EB is part of the recent development of synthetic control groups, in which treated units are not compared to single control units or simple averages of controls, but to a weighted average instead (Athey and Imbens, 2017). EB generates individual weights for all observations of the control group, such that the statistical moments of the given sets of observable characteristics, and hence ideally the propensity to be treated, equalize between the treatment and the control group (Hainmueller, 2012). This avoids the difficulty of the usual propensity score modeling approach of correctly specifying the propensity score model in order to obtain satisfactory balance between the two groups. Instead, EB ensures that the covariate distributions are balanced by construction. Since balancing the covariate distributions using a propensity score model turns out to be quite difficult in the setting I consider, EB becomes particularly useful.

Specifically, EB employs a loss function that minimizes the entropy distance of control group individuals' base weights, where each observation is given the same base weight, and EB weights upon the condition that the set of control group covariate moments are as similar as possible to the treatment group moments. Zhao and Percival (2017) show that this approach can be viewed as a propensity score weighting method, where the solution to the EB maximization problem is the logistic regression model with a different loss function than is used in maximum likelihood estimations. In fact, EB implicitly also fits a linear outcome regression model and is 'doubly robust' (Zhao and Percival, 2017): it is sufficient that only one of the two models (logistic propensity score model and outcome regression model) is correctly specified for EB to provide a consistent estimate of the average treatment effect for the treated (ATT).

[5] I increase this range in a robustness check in section 5 and find that the results continue to hold.

It is important to note that in order for the results above to hold and the estimate to measure the causal effect of entry, the following standard assumptions A1 and A2 must be imposed (Zhao and Percival, 2017). In addition, to ensure that the estimated effect is an individual level treatment effect devoid of any interference (such as general equilibrium or spillover effects), assumption A3 is needed.[6] These are well-established assumptions in the literature since the seminal work of Rosenbaum and Rubin (1983).

A1. (unconfoundedness). $\{Y(0), Y(1)\} \perp X | \mathbf{Z}$

A2. (overlap). $0 < \hat{\pi}(\mathbf{Z}_i) < 1$

A3. (stable unit treatment value). $Y_i = X_i \cdot Y_i(1) + (1 - X_i) \cdot Y_i(0)$

where $X_i \in \{0, 1\}$ is the treatment of unit i, Y_i is the outcome of unit i, $\hat{\pi}(\mathbf{Z}_i) = Pr(X = 1|\mathbf{Z})$ is the estimated propensity score and $\mathbf{Z}$ is the vector of confounding variables.

The assumption of unconfoundedness (A1) states that whether a retail store experiences entry in the bandwidth considered is independent of the outcomes, after conditioning on the confounding variables. Put differently, all variables that affect both the treatment of observed entry and the outcome of employment adjustments are measured. A1 ensures that treatment assignment is 'ignorable', that is, analyses on the correspondingly matched data or weighted original data will yield an unbiased estimate of the treatment effect. This is a fairly strong assumption, however the breadth and fine detail of my dataset allows me to control much more precisely for confounding factors than in previous studies in the literature. It should also be noted that A1 is equivalent to the assumption of exogeneity in the error term used in regressions or structural models and is simply made explicit in propensity score or matching analysis. I construct my dataset in order to satisfy A1 as follows.

First, I consider the distances to rival stores from the location of an existing retail store as possible confounding variables. It seems straightforward to assume that retailers take into account competing stores located in their vicinity. However, without making any ad-hoc assumptions about the size and extent of local markets, it is unclear across what distance the rival stores need to be considered in the analysis, so I attempt to be as generous as possible. I group rival stores in increasing bandwidths of distance from the retail store under consideration. To limit the number of variables in the estimation without losing too much information, I group competing stores in bandwidths of one kilometer.

[6]For a comprehensive discussion on interference effects in causal analysis see e.g. Huber and Steinmayr (2017).

Second, due to the nature of Swiss geography, controlling for infrastructure and geography is crucial. Locations in Swiss alpine towns may have little competition from rival stores, but also have difficult access and high transportation cost. I calculate all distances between competing stores as routing distances, rather than straight or geodesic, using the Open Source Routing Machine (OSRM). The OSRM makes use of the OpenStreetMap Project to provide routing distances in the same manner as e.g. Google Maps. It has the advantage of being able to run locally, rather than having to access it via an API on a server, and hence there are only computational limitations to the number of distances that can be considered.[7] In addition, it ensures that every result is directly reproducible.[8]

Third, I attempt to control for competition from stores located across the border from Switzerland. As discussed in section 2, due to the higher price level in Switzerland relative to the nations bordering Switzerland, cross-border shopping of Swiss consumers is commonplace. These foreign stores however are by definition not part of my retail dataset. I was provided with a list of the locations of all Swiss border posts and customs offices by the Swiss Customs Department and consider the nearest routing distance of a grocery retail store to any one of the border posts as a confounding factor. Since the list includes customs offices located inside Switzerland (for example at the Zurich airport) and many border posts are located very close to each other, I manually narrow down the list from 152 locations to only 54. These are border posts at (or in the immediate vicinity of) actual border crossings and should provide a very close approximation. In addition, I also include a dummy variable for the particular neighbouring country of each border post (Germany, France, Austria, Italy and Liechtenstein).

Finally, I consider the resident population at any location as a confounding factor. It seems likely that the number of consumers in close proximity is the most important driver of demand for grocery retail goods. Previous work has been forced to use e.g. the population per municipality as a market-level, control variable. Instead, as before I consider bandwidths around each location and aggregate the number of residents in each bandwidth. I include bandwidths of one kilometer width up to a ten kilometer distance from the location, in addition to bandwidths of 100 meter width for the first 300 meter.[9]

[7]In the case of Google Maps, there are strict limitations on the free use of the Google Maps API for calculations, making it very expensive for use here.

[8]Since both Google Maps and OpenStreetMap are constantly updated, running the same distance calculations at different points in time will lead to (slightly) varying results. I run all calculations locally with the map of Switzerland as it existed on 05.10.2016. Unfortunately, there is no way to consider changes to infrastructure over time, since changes in the map over time may simply represent users uploading details for the first time that have in reality existed for much longer, rather than actual infrastructure changes.

[9]I also have access to detailed municipal-level statistics on population, population density and taxable income. When correlating employment outcomes on the municipal-level with these variables, it turns out

Lastly, I also include a dummy variable for the language area of Switzerland that the store is located in (German, French, Italian, or Romansh).

The assumption of overlap (A2) requires that each store in the dataset has a non-zero probability of experiencing entry. It ensures that the support of the distributions of treated and control observations overlap, so that sufficient data is available to contrast treatment and control groups and hence extrapolation is not required.

The assumption of the stable unit treatment value (SUTVA, A3) in turn states that the outcome of a particular retailer Y_i is assumed to only be affected by the treatment assignment of that retailer, but not of any other store, or $Y_i(d_i)$. In essence, the SUTVA explicitly rules out any general equilibrium, spill-over, or interaction effects that relate to the treatment assignments of individual observations. This may be a problematic assumption when applying it to a setting of competing retail stores. For example, there may be a business-stealing effect due to entry for one particular store that causes an indirect (or spillover) effect transmitted via competition to a second store. Then, the potential outcome of that second store would no longer be independent of the treatment assignment of the first. This implies that there are two issues that need to be resolved: i) if such interference effects exist, entry may have both a direct and indirect effect on a treated store, ii) with interference, the treatment assignment of entry may have a non-zero impact on the potential outcome of the untreated as well.[10] Finally, I also need to deal with the fact that the outcome of a retailer may be affected by multiple entrants, so that $X_i \in \{0,1,2,...\}$, rather than the treatment being binary.

To solve these issues, I take two steps. First, I follow Hong and Raudenbush (2013) and assume that a store has a limited set of influence, i.e. it only competes within its local market. Specifically, each store i has a set of influence that includes all those competing stores that might affect the outcome of store i. They are denoted by the set of $\mathbf{D}_i$, which contains all stores within a distance x from store i. This is formalized in assumption A4. It allows me to limit the (potential) competitors that need to be controlled for in the estimation and avoids having to consider possible entry effects that seem highly unrealistic (e.g. entry in a different part of the country affecting a store's success). Additionally, I introduce assumption A5 which states in its weak form (a) that for a given local market and store there are no interferences between different local markets and that local markets are intact, such that stores do not react by migrating from one local market to another

that the population variables captures almost all variation that can be explained using these factors and other information such as total or average taxable income does not add much explanatory power.

[10]The mean difference in the observed outcome would then estimate the difference between the treatment effect for the treated (composed of the direct and indirect effect) and the treatment effect (i.e. spillover effect) for the untreated.

in response to entry. In its strong form (b) it essentially reintroduces the SUTVA in the context of local markets and imposes in addition to (a) that there are no spillovers between stores within one local market that are due to entry.

A4. (local market). Y_i only depends on competitors in $\mathbf{D}_i$, where $\mathbf{D}_i$ is the set of all stores $-i$, where the distance between i and $-i$ is smaller or equal than x.

A5 (a). (weak spillovers). Second-order spillovers across local markets are negligible and local markets are intact.

A5 (b). (strong spillovers). In addition to (a), second-order spillovers between firms that share a common market (local markets intersect) are negligible, so that Y_i only depends on the treatment assignment of i.

Consider the implications of assumptions A4 and A5 in turn. Figure 1 illustrates the implication of A4. The local market assumption states that each store $1,2,3$ has a limited set of influence (the local market) that is indicated by the circle. It only competes with other stores within this range. The effect of one store on a competing store is indicated by the arrows. For example, store 1 competes with store 2 only, which in turn competes with both store 1 and store 3.

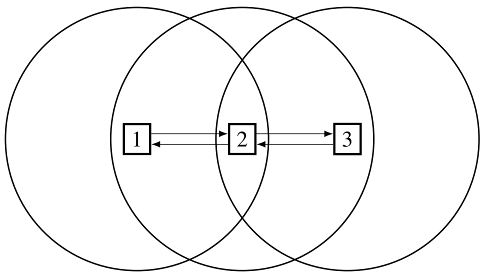
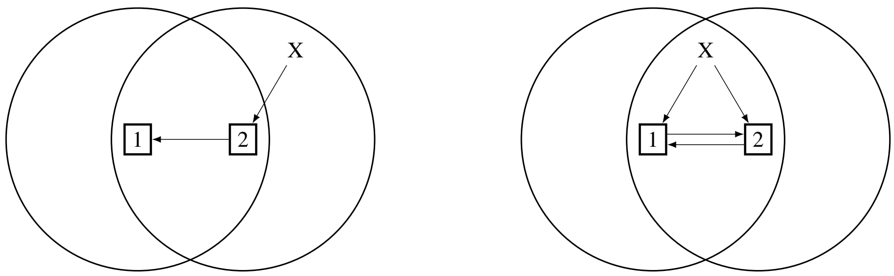


Figure 1: Interactions between three stores

Notes: The figure shows the implication of assumption A4 for the case of three firms. The arrows represent the effect of one competitor on another.

A5 in its weak form (a) in turn states that the effect of the treatment assignment of store 3 on store 1 via the effect on store 2 is negligible. Similarly, it implies that the treatment assignment of store 2, if the entry occurs outside of the local market of store 1, has only a negligible effect on store 1. In its strong form (b), assumption A5 additionally imposes that when both store 1 and 2 are affected by an entrant, the effect of the treatment assignment of store 2 on store 1 is negligible and hence the outcome of store 1 only depends on the treatment assignment of store 1.

The implication of this assumption can be seen in Figure 2. Panel (a) shows the case of entry occurring (X) within the local market of one of the stores only, or outside the common market of the two stores. Here, A5 (a) implies that store 2 is affected by entry, but store 1 is not. Panel (b) in turn shows the case of a retailer entering (X) within the local market of both store 1 and 2. A5 (b) implies that the entrant may have a direct effect on both existing stores, but that there is no indirect effect from entry via a competing store, say from 2 to 1. Note that a violation of the strong form of A5 in the sense of panel (b) would still allow me to consistently estimate the causal effect of entry, however the estimate would reflect both the direct and indirect effect. It is unclear how these two could be disentangled. If instead the weak form of A5 is violated as depicted in panel (a), the resulting estimates would be biased. I examine whether the assumption A5 (a) holds in the robustness checks in section 5 and find no evidence of a violation of the weak form of A5.



(a) Indirect effect outside common market

(b) Indirect effect within common market

Figure 2: Indirect effects

Notes: The figure shows the implication of assumption A5. Panel (a) shows the indirect effect that may arise when entry occurs outside of the local market boundary for store 1. Panel (b) shows the indirect effect that may arise when entry occurs within a common market of stores 1 and 2 (i.e. within both of their respective local markets).

The second step I take to resolve the issue and to isolate the effect of entry, is by only comparing retail stores that experienced *one* entry within one of the bandwidths to retailers that did not experience any entry within the maximum ten kilometers radius. Hence, I implicitly assume that the local market extends over no more than ten kilometers, or $x = 10$ kilometers. Previous work for example on Wal-Mart has found that the impact of a store entry is localized to the first four miles around the entry site (Ellickson and Grieco, 2013), indicating that a restriction to ten kilometers is not overly strict. In the robustness checks I increase this radius to 12 kilometers and find that the results continue to hold. If assumptions A1-A2 and A4-A5(a) hold, the estimates show the causal effect of store entry on existing retail stores. If A5(b) holds as well, the estimates show the causal, direct effect of entry. The estimated effect is then the average difference of the control observations, which receive a weight according to the EB estimation and the treatment observations, which receive a weight of one.

Table 4: Dataset excerpt

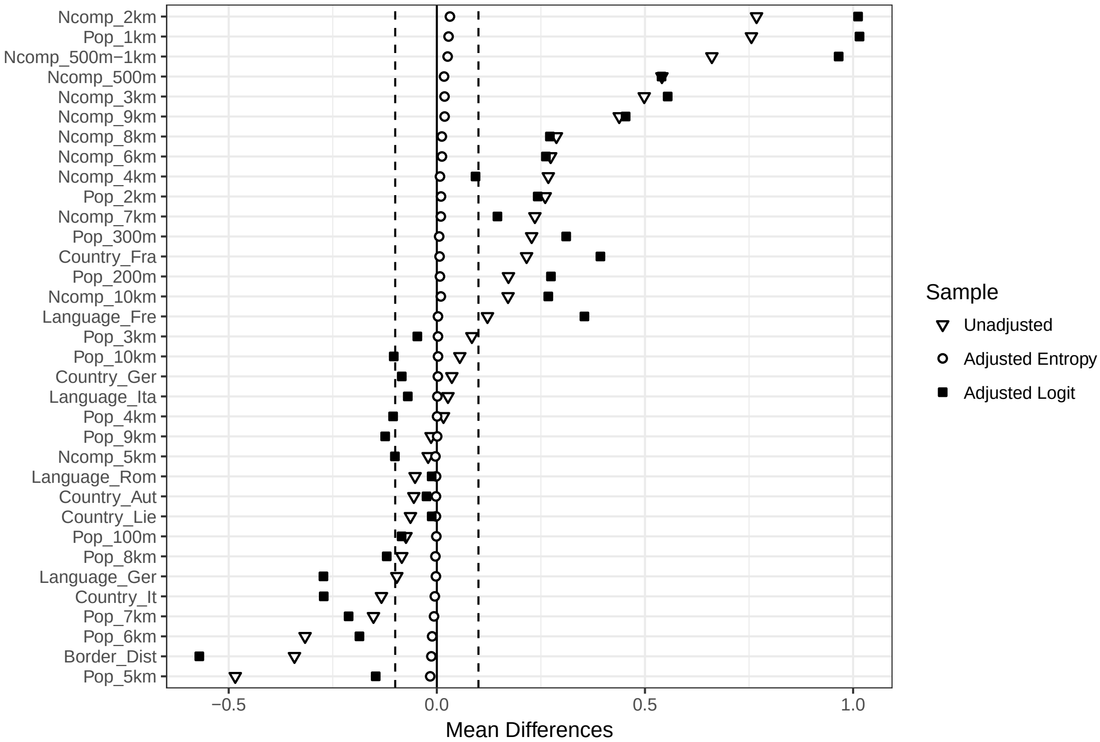
	All	Control	Treatment 1
Total rivals per one kilometer			
0-1 kilometer	3.84	2.23	4.65
1-2 kilometer	5.33	0.90	4.17
2-3 kilometer	7.13	1.07	2.32
3-4 kilometer	7.53	1.22	1.55
4-5 kilometer	7.24	1.36	1.26
Total population per one kilometer			
0-1 kilometer	3,203.07	948.43	2,823.34
1-2 kilometer	1,451.27	236.22	514.84
2-3 kilometer	888.64	193.12	222.32
3-4 kilometer	737.95	149.73	151.59
4-5 kilometer	590.86	160.62	87.25
Mean border distance in kilometers	40.38	48.86	40.15
Mean FTE employment	10.85	7.02	12.72
Mean Total employment	14.43	9.84	16.87

Notes: The table shows for the three groups: all stores, the control group (stores that did not experience any entry within ten kilometers), and treatment group 1 (stores that experienced one entry within two kilometers, but no other entry), the total number of competing grocery retail stores and total resident population in one kilometer bands around a store, the mean border distance in kilometers, as well as the mean full-time equivalent (FTE) and total employment of stores.

Table 4 shows some key variables of the dataset and their means for three different groups: i) the full dataset, ii) the control group (retailers that did not experience any entry within 10 kilometers), and iii) the first treatment group (retailers that experienced one entry within 2 kilometers but no further entry within 10 kilometers). Retailers in the control group appear to be located in areas with a smaller population in their vicinity, relative to retailers in the first treatment group. The difference in the first kilometers is particulary visible. Considering the treatment group concerns entry within the first 2 kilometers, this seems unsurprising.

In addition, retailers in the control group have less competition from other existing stores in their neighbourhood and tend to be further away from the border. A likely explanation is that with a larger share of the population in Switzerland living near the border, rather than in the mountainous geographical center of the country, retailers in less populated areas correlate with retailers more distant from the border. Finally, stores that experienced entry are significantly larger than those that did not, as measured by employment numbers. The full dataset comprised of all 34 confounding variables shown separately for each treatment as well as the control group can be found in Table 10 in the appendix.

Figure 3: Balancing of EB and PS for entry within two kilometers



Notes: The figure shows the mean differences of all confounders for treatment group 1 (one entrant within two kilometers, but no other entry) relative to the control group (no entry within ten kilometers) for three samples: unadjusted, adjusted entropy (where the control group is weighted using the entropy balancing weights), and adjusted logit (where the control group is weighted using the propensity score (logit) weights).

Lastly, to demonstrate the usefulness of EB in this setting, Figure 3 shows how well the covariates are balanced for the first estimation (treatment group 1). It depicts the standardized mean difference between the treatment and control group of each covariate in three samples: (i) the raw, unadjusted data, (ii) the weighted data using entropy balancing (entropy), and (iii) the weighted data using a standard logistic regression (without interaction terms) for the propensity score (logit). It is immediately visible that weighting using the propensity score model does not work well. Only around half of the covariates can be considered balanced (i.e. within the usually considered threshold of a standardized mean difference below 0.1 indicated by the dashed lines) and for some of them, the balancing is in fact better in the unadjusted sample. EB instead by design balances all covariates almost perfectly. The balancing as measured by the standardized mean difference is documented for each estimation using EB separately in the appendix in figures 7 to 11, and in detail in tables 11 and 12. The next section reports and discusses the results of the main estimations.

4 Results

Table 5 reports the main estimates of the effect of retail store entry in 2011-2012 on the growth rate of employment of existing retail stores in 2012-2013. The second and third columns show the parameters of the impact of entry on all rival retail stores regardless of store size, measured using full-time equivalent (FTE) and total employment, respectively. I find strong and consistent evidence for a negative effect on the growth rate of both full-time equivalent and total employment in the first two bandwidths of 0-2 and 2-4 kilometers. These results suggest that grocery retail stores in Switzerland compete in a highly localized radius of up to four kilometers, but not beyond. As one might expect, the effect in the first bandwidth is significantly stronger than in the second for FTE employment, indicating that business-stealing effects are heterogeneous within a local market and decrease with increasing distance. The growth rate of total employment in turn does not show this pattern, however the statistical significance decreases across distance. In addition, the effect on full-time equivalent employment is much stronger. In the first bandwidth, stores react by reducing their FTE employment hours by 12%, while it is much subdued in the second and around 7%. This seems to indicate that retail stores tend to focus on adjusting their employment numbers short-term in response to entry by lowering the hours their employees work, rather than letting them go.

The growth rate takes into account the absolute size of a retail establishment as measured in employment, however as discussed earlier in section 2, the growth rates between small and large retail stores differ significantly. In addition, the location choices of large and small retail stores appear to not be identical. In order to examine the differential effects across store heterogeneity, I estimate the effects separately by the store types. Since small stores consistently make up over 90% of retail stores across the years and the number of entries observed by large stores is particularly small, I focus on examining small stores.

The fourth and fifth column of Table 5 report the estimates of entry by small stores in 2011-2012 on the employment growth rate of small retailers only in 2012-2013. I find a business-stealing effect of a reduced magnitude for the growth rate in full-time equivalent and of greater magnitude for total employment within the first bandwidth, relative to the main results. I do not find a statistically significant effect in the second bandwidth for FTE employment, but a reduced and less statistically significant effect for

Table 5: Impact of store entry on employment growth rate

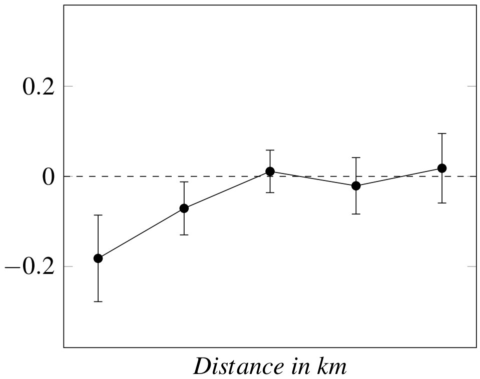
	All stores		Small stores	
Distance to entry	FTE	Total	FTE	Total
0-2 kilometers	−0.128**	−0.065**	−0.092***	−0.092**
	(0.058)	(0.032)	(0.036)	(0.042)
2-4 kilometers	−0.071**	−0.065*	−0.048	−0.069*
	(0.036)	(0.038)	(0.037)	(0.040)
4-6 kilometers	0.011	0.003	−0.012	0.005
	(0.029)	(0.032)	(0.039)	(0.038)
6-8 kilometers	−0.021	−0.026	−0.027	−0.031
	(0.038)	(0.42)	(0.039)	(0.046)
8-10 kilometers	0.018	0.011	0.019	0.021
	(0.047)	(0.039)	(0.040)	(0.041)

Notes: *** indicates significance at the 1% significance level, ** at the 5% level, * at the 10% level. Heteroscedasticity-robust standard errors in brackets.

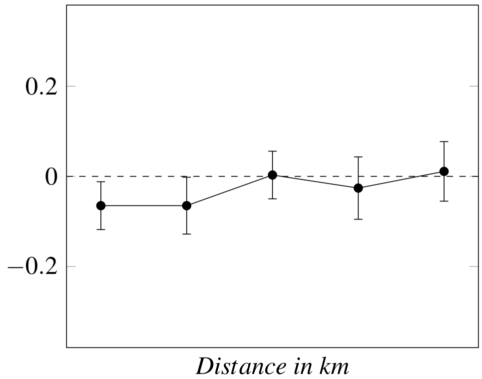
total employment. It appears that smaller stores react more on the extensive margin of employment adjustments to entry by similar competitors than on the intensive. In addition, since the results only show an effect in the 2-4 kilometer range on total employment growth that is marginally statistically significant at the 10% significance level and none for FTE employment, this strongly indicates that the local market size varies by store size: small retailers compete within a shorter distance with other small retailers, compared to large ones. A likely explanation would be that customers are more willing to travel a longer distance for sizable grocery stores, where a great amount of shopping can be done quickly, compared to small grocery retailers.

The main results are illustrated in Figure 4. The four panels show the value of the point estimate and the 90% confidence intervals across the increasing distance from left to right. The top panels show the results for all retailers, documented in columns two and three in Table 5, while the bottom panels show the results for small retailers only, as documented in columns four and five in Table 5. The panels all illustrate the same pattern: the impact of store entry is concentrated on competitors in the immediate vicinity and no statistically significant effect can be observed beyond the four kilometer range. The business-stealing effect from entry for all stores is more sizable within the first bandwidth, relative to the second, while for small stores the effect is only evident within the first bandwidth (for total employment it is marginally statistically significant at the 10% level).

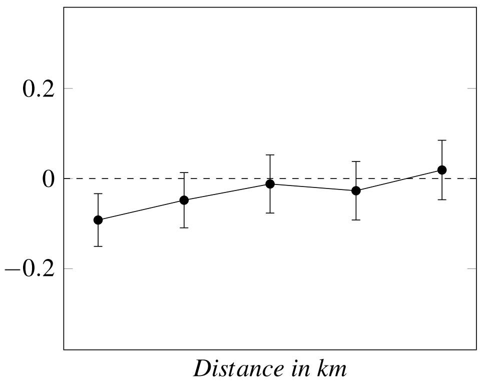
The specifications so far have focused on the intensive margin of changes in the number of employees or their work hours. Yet, the data shows that there is a sizable fluctuation



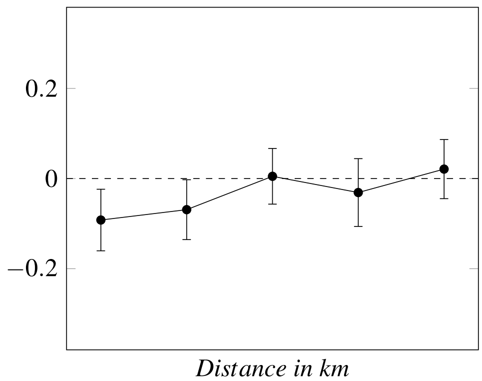
(a) All Stores (FTE)



(b) All Stores (Total)



(c) Small Stores (FTE)



(d) Small Stores (Total)

Figure 4: Estimated effects of entry

Notes: The figure shows the point estimates and 90% confidence intervals for the estimation results documented in Table 5. Distance is measured from left to right in two kilometer bandwidths. The top panels show the effects for all stores, the bottom panels show the effects for small stores only. The left panels show the impact measured in full-time-equivalent (FTE) employment, the right panels show the impact in total employment.

in retail store entry and exit, especially for small retailers (see the discussion in section 2). I now turn to examining the impact that store entry has on employment changes that are due to exit. I incorporate employment destruction that is due to stores exiting the market in the growth rate measure. Columns two and three in Table 6 show the estimated impact of retail store entry. I find that the effect of store entry is significantly more pronounced and that the business-stealing effect is of up to twice the magnitude compared to the previous estimates. It appears that the bigger component of the observed decreases is due to contraction or reduced growth of continuing firms, rather than employment destruction due to exit. This is particularly visible in the FTE employment measure.

Table 6: Impact of store entry on employment growth, including employment destruction due to exit.

	All stores		Small stores	
Distance to entry	FTE	Total	FTE	Total
0-2 kilometers	−0.147***	−0.131***	−0.171**	−0.105***
	(0.057)	(0.049)	(0.056)	(0.065)
2-4 kilometers	−0.121**	−0.120**	−0.097	−0.082
	(0.056)	(0.061)	(0.066)	(0.069)
4-6 kilometers	−0.097	−0.085	−0.044	−0.079
	(0.063)	(0.066)	(0.060)	(0.069)
6-8 kilometers	−0.042	−0.046	−0.049	−0.052
	(0.061)	(0.064)	(0.061)	(0.064)
8-10 kilometers	0.008	−0.036	−0.054	−0.079
	(0.055)	(0.070)	(0.060)	(0.070)

Notes: *** indicates significance at the 1% significance level, ** at the 5% level, * at the 10% level. Heteroscedasticity-robust standard errors in brackets.

In addition, I find more consistent evidence of the impact being centered in the first two bandwidths, with the impact being greater in the first bandwidth. As before, the effect appears to be greater on work hour reduction, rather than on the number of employees and similarly the results again suggest very clearly that the local market in Switzerland is limited to a radius of around four kilometers for a given retailer and that the degree of competition decreases with distance. It appears that the greater the overlap of the catchment area of two stores, the greater the extent to which they compete and negatively impact each other. Finally, I also provide estimates that include employment destruction from exit for small retailers only in columns four and five of Table 6. As before, the effect is greater than for all stores in the first bandwidth, but I find no evidence of an impact of entry by a small retailer on another small retail store beyond the range of 2 kilometers.

The increased size of the effects also closely mirrors the relative increase observed when estimating the impact for all retailers.

The estimates overall suggest that the short-term business-stealing effect from entry is significant and contained in a tight local market. A competing store may be forced to reduce the work hours of their employees by up to 12% and the total number of employees by up to 10% for a small retailer. However, there exists considerable heterogeneity of the impact across both space and retail store type. The impact on both the intensive margin of employment adjustments and on the extensive margin of market exit falls more strongly on small retailers than large ones. In addition, small stores appear to be less able to react by lowering work hours for employees.

It should also be noted that the estimates do not account for job creation due to the entry under consideration. A rough calculation suggests that the change in aggregate, local employment due to job creation by the average entrant relative to the job losses experienced at competing stores depends on the measurement used. Given the average number of competing stores within 0-2 and within 2-4 kilometers distance of entry sites and the main estimates provided in Table 5, the average full-time equivalent employment loss due to business-stealing effects from entry is around 15% larger than the job creation. However, the data also shows that the average total employment of an entrant is much larger than the full-time equivalent. Indeed, the same calculation in absolute terms (that is, for total employment) suggests more jobs are created by entry than are lost at competing stores. The difference is sizable: entrants create approximately 10% more jobs than they destroy in the short-term at competing stores. One possible explanation for this striking difference in the approximate, net impact of entry on employment might be that there are setup costs in the size of the workforce required for opening or running a new store.

In the next section, I examine the robustness of the assumptions introduced in section 3.

5 Robustness Checks

The estimator employed to study the impact of grocery retail store entry provides causal estimates of the effect of entry on the outcome of a store given assumptions A1-A2 and A4-A5 (a). Since it is generally impossible to test for the strong ignorability condition stated in A1 and unclear how assumption A5 (b) could be tested,[11] I focus on examining whether assumptions A4 and A5 (a) hold.

Assumption A4 may be violated if the local market area is larger than the ten kilometers considered in the estimations. In columns two and three in Table 7, I provide the same estimates as in the main results reported in Table 5 for all stores, except with an increased maximum radius of twelve kilometers. I find that the business-stealing effect is clearly visible as before and the estimates confirm that the extent of the local market is limited to the first four kilometers. In particular, the size and statistical significance of the effect of entry on total employment are very similar when increasing the distance by two kilometers, while the size of the impact on full-time equivalent employment changes slightly and becomes smaller in the first bandwidth, but larger in the second.

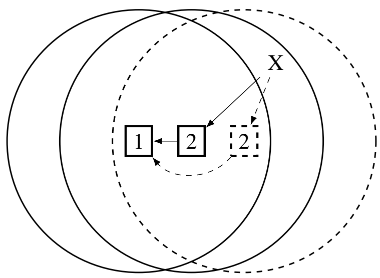


Figure 5: Robustness check for indirect effects

Notes: The figure illustrates the robustness check for spillovers across market boundaries for two separate bandwidths of distance between store 1 and store 2. The farther bandwidth is shown in dashed lines. The indirect (spillover) effect by store entry on store 1 via competition with store 2 is indicated by arrows.

Next, I examine the robustness of assumption A5 (a). It is violated if there is evidence of spillovers across the local market boundary of ten kilometers. In order to determine if the assumption holds, I provide evidence of the size of entry effects that can only occur indirectly, given the local market size, in increasing distance. Specifically, I estimate the effect of entry that occurs just outside of the ten kilometer radius of a store, but that is within the ten kilometer radius of a second store, which shares a common market with the first. This scenario is illustrated in panel (a) of Figure 2.

I proceed to continually move the two retail stores away from each other, while keeping the distance of the entrant to the first store constant, so that as the distance between the competitors increases, the distance between the entrant and the second, 'spillover' store decreases. This approach is illustrated in Figure 5. The figure shows the possible spillover effect from entry on store 1 transmitted through competition with store 2. The distance between store 1 and store 2 is successively increased (illustrated in dashed lines),

[11]The indirect effect from a competitor that is caused by an entrant within the common market of the two firms would be indistinguishable from the direct effect of that same entry.

while entry continues to occur within 10-12 kilometers of distance to store 1. By moving the two stores away from each other, I am able to check for the existence of spillover effects across the whole spectrum of possible interference by distance. The distances between stores are grouped as before in 2 kilometer bandwidths.

Table 7: Estimates of the robustness checks

	Extended Local Market		Spillovers	
Distance to entry	FTE	Total	FTE	Total
0-2 kilometers	−0.092***	−0.077***	−0.012	−0.016
	(0.040)	(0.026)	(0.244)	(0.029)
2-4 kilometers	−0.113**	−0.059*	0.025	−0.040
	(0.056)	(0.031)	(0.249)	(0.053)
4-6 kilometers	0.006	0.009	0.002	0.013
	(0.051)	(0.063)	(0.226)	(0.218)
6-8 kilometers	−0.027	−0.027	−0.034	−0.047
	(0.075)	(0.064)	(0.31)	(0.205)
8-10 kilometers	−0.032	−0.035	0.024	0.001
	(0.071)	(0.073)	(0.261)	(0.217)

Notes: *** indicates significance at the 1% significance level, ** at the 5% level, * at the 10% level. Heteroscedasticity-robust standard errors in brackets. (1) reports the estimates of the impact of entry on all stores when extending the maximum range from ten kilometers to twelve. (2) reports the estimates of the indirect effect of entry.

The estimated impact is reported in in columns four and five in Table 7. I find no evidence of an indirect effect that is due to the impact of an entrant on a given retail store for any of the different distances considered. The estimates strongly suggest that there is no concern of a violation of assumption A5 (a).

6 Conclusion

This paper develops a novel, nonparametric approach to empirically determine geographic market size and applies it to the grocery retail environment in Switzerland. The approach exploits highly detailed spatial data of the industry and allows inferring local market size while refraining from assumptions on the demand function underpinning consumer substitution. I study the impact of store entry on competitors using the entropy balancing estimator and propose using a different set of assumptions instead of the classic 'stable-unit-treatment-value assumption' that appear more appropriate for the particular application analysed here. Lastly, I provide a simple approach to check for possible violations of the 'local market' and 'spillover' assumptions proposed.

The results show that entry by a retail store does not negatively impact employment at grocery retailers located more than four kilometers away from the entry site. Within this range, I find consistent evidence of heterogeneity in entry effects and competition across both distance and store type. Overall, it appears that differentiation of grocery retail stores by location or type softens competition in the industry. Inside the narrower range of two kilometers, the business-stealing effect of entry is significantly more pronounced. In the wider range of two to four kilometers the effect is much subdued. Moreover, the negative impact in this wider range is mainly driven by competition between larger supermarkets and supercenters. For smaller supermarkets and mom-and-pop stores in turn the impact falls exclusively on the short range of zero to two kilometers, as they appear to be unable to convince consumers to travel more than two kilometers for their goods.

7 Appendix

Table 8: Population data summary table

Statistic	N	Mean	St. Dev.
Sex	8,174,154	1.506	0.500
Marital Status	8,174,154	1.772	0.876
Nationality (State)	8,174,154	8,135.986	131.229
Reporting Municipality	8,174,154	2,930.933	2,249.550
Type of Residence	8,174,154	1.018	0.140
Origin Country ID (CH)	8,174,154	1,050.482	2,766.712
Origin Country ID	8,174,154	917.419	2,609.824
Population Type	8,174,154	1.044	0.276
Age	8,174,154	40.880	22.635
Nationality Category	8,174,154	1.234	0.423
Residence Permit	8,174,154	-0.808	2.307
Population Group	8,174,154	1.978	1.602
Locality	8,174,154	522,786.900	278,705.900
Locality Size	8,174,154	7.455	3.456

Table 9: NOGA classification for retailers

NOGA Code	Original Name	Translation	Size
471101	Verbrauchermärkte	Supercenter	$>$2,500 m^2
471102	Grosse Supermärkte	Large Supermarkets	1,000-2,499 m^2
471103	Kleine Supermärkte	Small Supermarkets	400-999 m^2
471104	Grosse Geschäfte	Large Stores	100-399 m^2
471105	Kleine Geschäfte	Small Stores	$<$100 m^2

Table 10: Treatment and control groups means

Variable	0-2 km	2-4 km	4-6 km	6-8 km	8-10 km	Control
Pop_100m	7.53	17.81	23.91	13.66	7.69	9.18
Pop_200m	13.76	19.59	11.41	8.22	7.69	5.01
Pop_300m	16.15	14.61	11.05	12	10.84	5.32
Pop_1km	2,823.34	2,357.74	1,476.55	1,542.56	1,346.29	948.43
Pop_2km	514.84	870.77	771.11	355.78	311.14	236.22
Pop_3km	222.32	400.67	500.53	291	248.90	193.12
Pop_4km	151.59	198.59	374.40	261.54	236.92	149.73
Pop_5km	87.24	153.70	315.07	297.44	400.54	160.62
Pop_6km	98.77	198.73	333.14	440.45	574.29	185.98
Pop_7km	142.73	168.38	353.12	355.87	337.59	208.79
Pop_8km	173.30	439.38	560.09	160.33	347.02	225.32
Pop_9km	255.46	317.75	514.40	322.63	461.94	251.40
Pop_10km	202.88	237.41	249.46	280.46	391.12	188.97
Ncomp_500m	2.38	1.98	1.91	1.78	1.71	1.62
Ncomp_500m-1km	2.27	1.26	0.82	0.87	1.02	0.62
Ncomp_2km	4.17	2.78	1.64	1.41	1.08	0.90
Ncomp_3km	2.32	3.09	1.52	1.77	1.66	1.07
Ncomp_4km	1.55	3.59	2.50	2.11	1.55	1.22
Ncomp_5km	1.26	2.11	3.28	2.24	1.95	1.36
Ncomp_6km	1.97	1.93	2.82	2.36	2.36	1.37
Ncomp_7km	2.03	2.16	2.75	3.22	2.32	1.62
Ncomp_8km	2.39	2.15	2.73	3.67	2.68	1.80
Ncomp_9km	3.08	2.39	2.71	3.09	3.78	1.94
Ncomp_10km	2.51	2.54	2.99	3.09	4.15	2.16
Border_Dist	40,148.24	46,713.21	47,636.71	42,684.13	43,983.34	48,864.77
Border_Germany	0.28	0.39	0.52	0.50	0.42	0.27
Border_France	0.46	0.34	0.20	0.29	0.29	0.24
Border_Austria	0.02	0.01	0.02	0.01	0.05	0.07
Border_Italy	0.23	0.26	0.22	0.12	0.19	0.35
Language_German	0.54	0.53	0.76	0.78	0.75	0.63
Language_French	0.38	0.28	0.11	0.15	0.20	0.25
Language_Italian	0.08	0.18	0.13	0.06	0.05	0.06
Language_Romansh	0.01	0	0	0.005	0	0.06

Table 11: Mean differences all retailers

	0-2km		2-4km		4-6km		6-8km		8-10km	
Variable	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Pop_100m	−0.074	−0.001	0.116	0	0.088	0	0.066	0	−0.065	0
Pop_200m	0.172	0.007	0.205	0	0.142	0	0.083	0	0.079	0
Pop_300m	0.228	0.006	0.157	0	0.134	0	0.153	0	0.118	0
Pop_1km	0.756	0.028	0.576	0	0.333	0	0.371	0	0.292	0
Pop_2km	0.260	0.010	0.373	0	0.378	0	0.164	0	0.128	0
Pop_3km	0.084	0.003	0.333	0	0.314	0	0.208	0	0.130	0
Pop_4km	0.016	0.001	0.057	0	0.269	0	0.136	0	0.191	0
Pop_5km	−0.484	−0.016	−0.063	0	0.273	0	0.179	0	0.298	0
Pop_6km	−0.317	−0.011	−0.096	0	0.195	0	0.258	0	0.245	0
Pop_7km	−0.152	−0.007	−0.252	0	0.169	0	0.150	0	0.159	0
Pop_8km	−0.084	−0.003	0.147	0	0.240	0	−0.198	0	0.110	0
Pop_9km	−0.014	0.001	−0.009	0	0.222	0	0.098	0	0.217	0
Pop_10km	0.055	0.003	0.039	0	0.044	0	0.132	0	0.237	0
Ncomp_500m	0.541	0.017	0.302	0	0.216	0	0.188	0	0.068	0
Ncomp_500m-1km	0.661	0.026	0.374	0	0.094	0	0.217	0	0.284	0
Ncomp_2km	0.768	0.031	0.570	0	0.312	0	0.228	0	0.082	0
Ncomp_3km	0.498	0.018	0.783	0	0.295	0	0.263	0	0.243	0
Ncomp_4km	0.268	0.008	0.710	0	0.446	0	0.312	0	0.140	0
Ncomp_5km	−0.021	−0.003	0.313	0	0.617	0	0.381	0	0.254	0
Ncomp_6km	0.273	0.012	0.336	0	0.555	0	0.414	0	0.409	0
Ncomp_7km	0.236	0.010	0.209	0	0.363	0	0.528	0	0.322	0
Ncomp_8km	0.287	0.012	0.188	0	0.359	0	0.630	0	0.314	0
Ncomp_9km	0.438	0.019	0.089	0	0.282	0	0.356	0	0.522	0
Ncomp_10km	0.171	0.010	0.164	0	0.267	0	0.338	0	0.587	0
Border_Dist	−0.342	−0.013	−0.093	−0	−0.042	0	−0.161	−0	−0.223	0
Germany	0.036	0.002	0.123	−0	0.228	0	0.230	0	0.149	0
France	0.216	0.007	0.120	0	−0.053	0	0.043	0	0.037	0
Austria	−0.055	−0.002	−0.061	0	−0.041	−0	−0.059	−0	−0.006	−0
Italy	−0.133	−0.005	−0.120	−0	−0.092	−0	−0.225	−0	−0.164	−0
Liechtenstein	−0.063	−0.002	−0.062	−0	−0.042	−0	0.011	0	−0.016	0
German	−0.096	−0.002	−0.117	−0	0.109	−0	0.158	0	0.109	0
French	0.122	0.003	0.069	0	−0.143	−0	−0.114	0	−0.049	0
Italian	0.027	0.001	0.107	0.009	0.092	0.039	0.011	0	−0.002	0.016
Romansh	−0.052	−0.002	−0.059	−0.009	−0.059	−0.039	−0.054	−0	−0.059	−0.016

Table 12: Mean differences small retailers

	0-2km		2-4km		4-6km		6-8km		8-10km	
Variable	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Pop_100m	−0.039	−0.003	0.106	0.00000	0.098	0	0.046	0	−0.032	0
Pop_200m	0.144	0.002	0.227	0.00000	0.162	0	0.072	0	0.057	0
Pop_300m	0.236	0.0004	0.160	0.00000	0.119	0	0.123	0	0.128	0
Pop_1km	0.742	0.006	0.585	0.00001	0.300	0	0.322	0	0.304	0
Pop_2km	0.252	0.002	0.364	0.00000	0.369	0	0.191	0	0.126	0
Pop_3km	0.076	−0.0005	0.323	0.00001	0.299	0	0.217	0	0.095	0
Pop_4km	−0.020	0.002	0.081	0.00000	0.236	0	0.128	0	0.173	0
Pop_5km	−0.444	−0.002	−0.050	0.00000	0.254	0	0.193	0	0.308	0
Pop_6km	−0.291	−0.001	−0.110	0.00000	0.174	0	0.228	0	0.249	0
Pop_7km	−0.178	−0.002	−0.281	0.00000	0.160	0	0.137	0	0.169	0
Pop_8km	−0.134	−0.0002	0.138	0.00000	0.250	0	−0.184	0	0.095	0
Pop_9km	−0.059	−0.001	−0.006	0.00000	0.205	0	0.086	0	0.218	0
Pop_10km	0.050	0.002	0.060	0.00000	0.066	0	0.097	0	0.242	0
Ncomp_500m	0.547	−0.001	0.306	0.00001	0.135	0	0.200	0	0.040	0
Ncomp_500m-1km	0.651	0.006	0.409	0.00000	0.091	0	0.184	0	0.277	0
Ncomp_2km	0.781	0.009	0.566	0.00001	0.229	0	0.231	0	0.115	0
Ncomp_3km	0.504	0.005	0.782	0.00000	0.301	0	0.234	0	0.222	0
Ncomp_4km	0.261	0.002	0.704	−0.00000	0.467	0	0.298	0	0.145	0
Ncomp_5km	0.002	−0.002	0.328	0.00000	0.630	0	0.350	0	0.248	0
Ncomp_6km	0.278	0.004	0.352	0.00000	0.563	0	0.396	0	0.407	0
Ncomp_7km	0.254	0.003	0.173	0.00000	0.375	0	0.518	0	0.340	0
Ncomp_8km	0.302	0.004	0.207	0.00000	0.329	0	0.627	0	0.312	0
Ncomp_9km	0.396	0.003	0.111	0.00001	0.273	0	0.345	0	0.531	0
Ncomp_10km	0.177	0.003	0.176	0.00000	0.296	0	0.334	0	0.601	0
Border_Dist	−0.310	−0.0003	−0.127	−0.00001	0.0001	0	−0.180	−0	−0.220	0
Germany	0.037	0.001	0.120	0.00000	0.220	0	0.221	0	0.162	0
France	0.198	−0.0004	0.114	0.00000	−0.054	0	0.044	0	0.031	0
Austria	−0.051	−0.0005	−0.058	0.00000	−0.036	−0	−0.056	−0	−0.009	−0
Italy	−0.123	0.0003	−0.107	0.00000	−0.086	−0	−0.224	−0	−0.159	−0
Liechtenstein	−0.060	−0.0003	−0.068	−0.00001	−0.045	−0	0.015	0	−0.025	0
German	−0.088	0.001	−0.118	−0.00000	0.124	0	0.154	0	0.113	0
French	0.112	−0.001	0.063	0.00000	−0.154	−0	−0.109	0	−0.054	0
Italian	0.031	−0.007	0.117	0.014	0.092	0.046	0.011	0	0.003	0.016
Romansh	−0.054	0.008	−0.062	−0.014	−0.062	−0.046	−0.056	−0	−0.062	−0.016

Figure 6: Distribution of retail stores by type

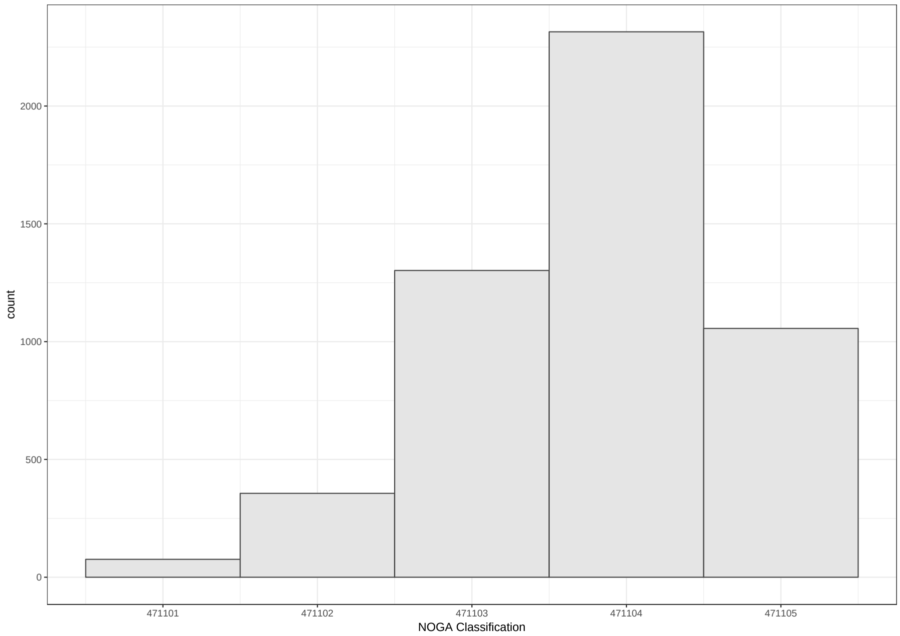


Figure 7: Balancing for entry within 0-2 kilometers.

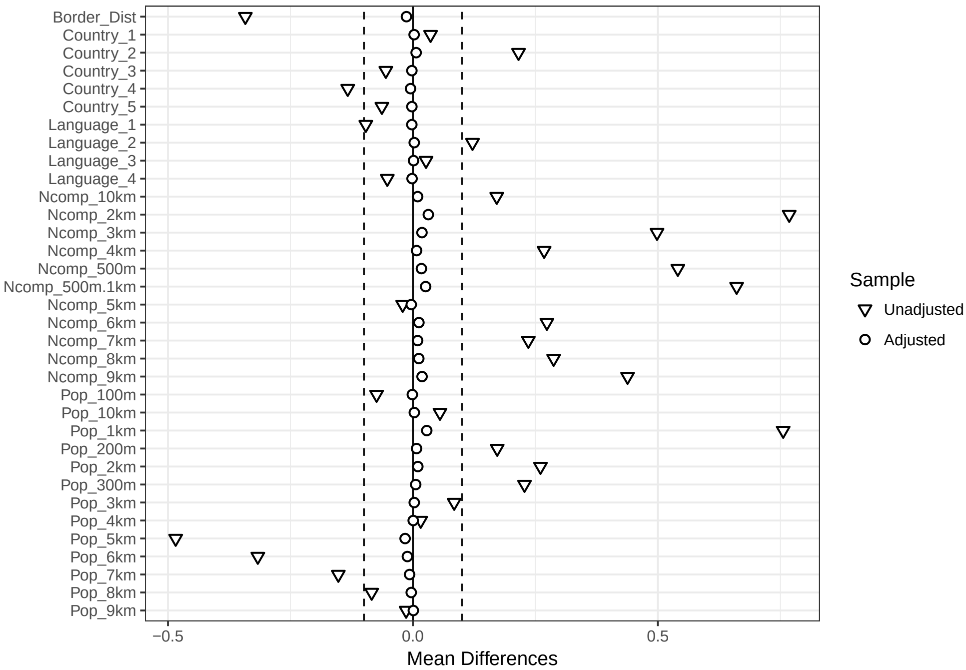


Figure 8: Balancing for entry within 2-4 kilometers.

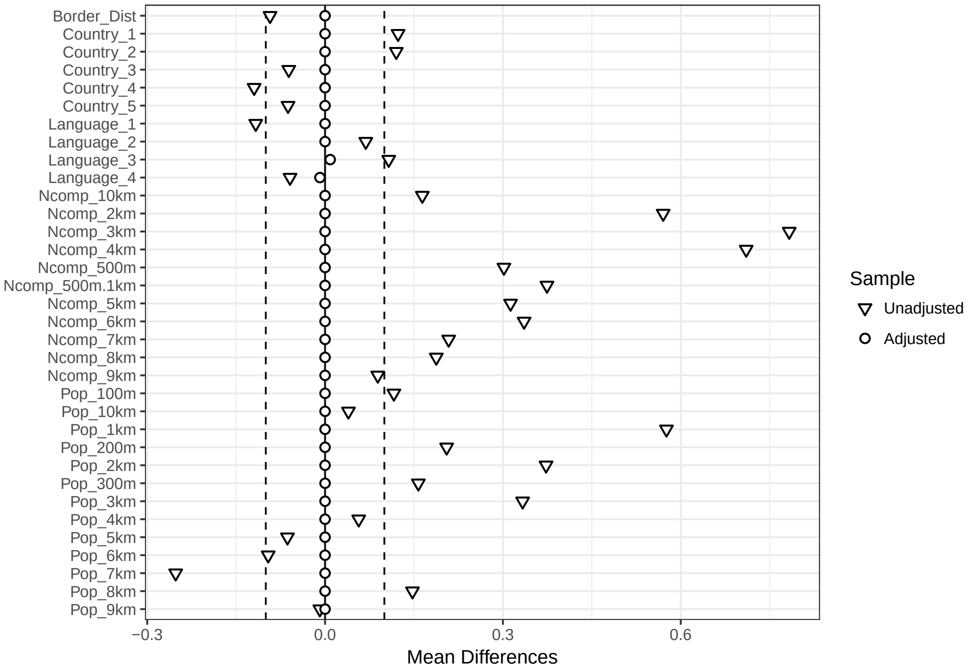


Figure 9: Balancing for entry within 4-6 kilometers.

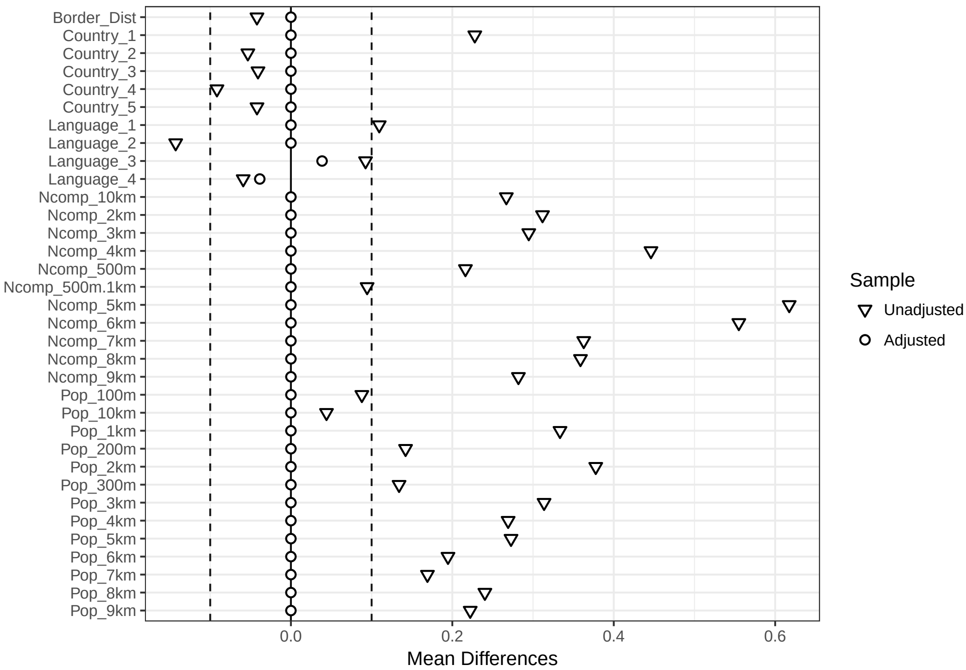


Figure 10: Balancing for entry within 6-8 kilometers.

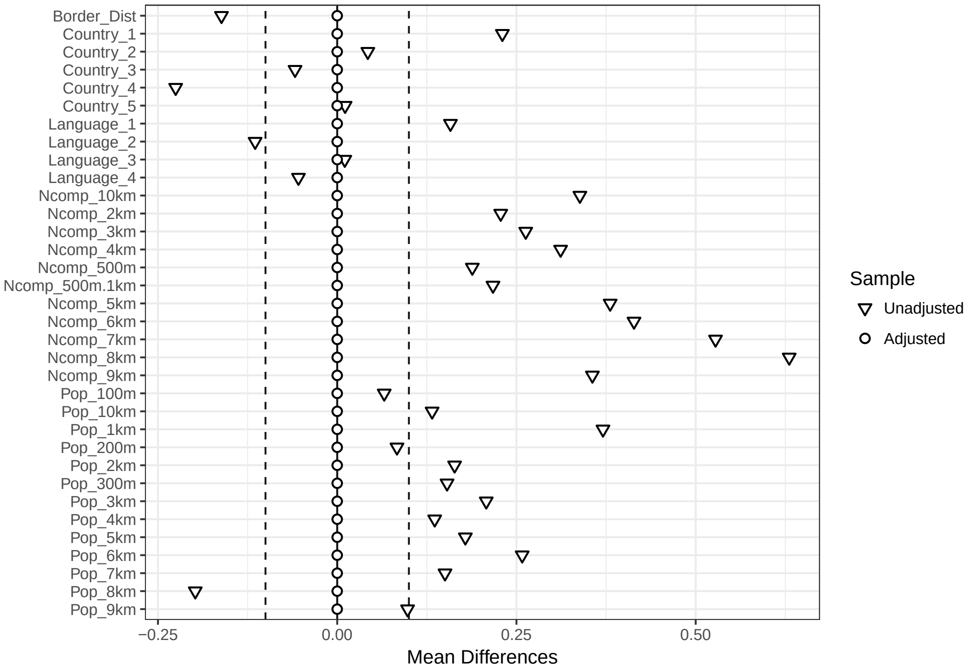
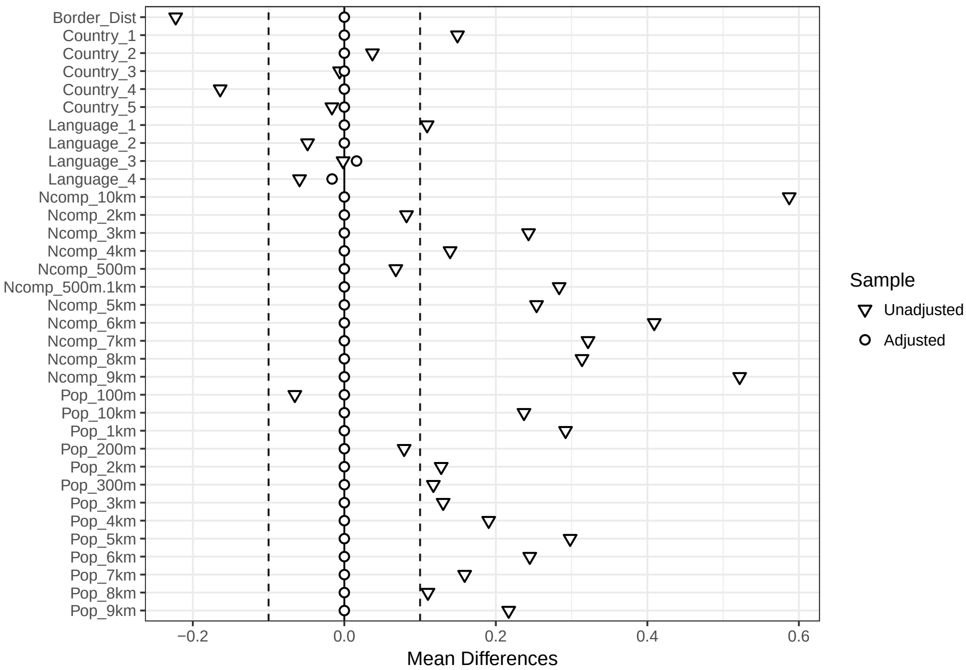


Figure 11: Balancing for entry within 8-10 kilometers.



References

Aguirregabiria, V. and Suzuki, J. (2016). *Handbook on the Economics of Retailing and Distribution*, Edward Elgar Publishing, chap. Empirical games of market entry and spatial competition in retail industries. 201–232.

Amiram, D., Beaver, W. H., Landsman, W. R. and Zhao, J. (2017). The effects of credit default swap trading on information asymmetry in syndicated loans. *Journal of Financial Economics* 126(2), 364 – 382.

Athey, S. and Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives* 31(2), 3–32.

Baker, J. B. (2007). Market definition: An analytical overview. *Antitrust Law Journal* 74(1), 129–173.

Baker, J. B. and Bresnahan, T. F. (2008). *Handbook of Antitrust Economics*, MIT Press, chap. Economic Evidence in Antitrust: Defining Markets and Measuring Market Power. 1–42.

Basker, E. (2007). The causes and consequences of wal-mart's growth. *Journal of Economic Perspectives* 21(3), 177–198.

Bundeskartellamt (2015). Fusionskontrollverfahren Verfügung (B2-96/14). `https://www.bundeskartellamt.de/SharedDocs/Entscheidung/DE/Entscheidungen/Fusionskontrolle/2015/B2-96-14.pdf?__blob=publicationFile&v=3`.

Capps, C. S., Dranove, D., Greenstein, S. and Satterthwaite, M. (2002). Antitrust policy and hospital mergers: Recommendations for a new approach. *Antitrust Bulletin* 47(4), 677–714.

Coate, M. B. and Fischer, J. H. (2008). A practical guide to the hypothetical monopolist test for market definition. *Journal of Competition Law & Economics* 4(4), 1031–1063.

Danger, K. L. and Frech, H. E. I. (2001). Critical thinking about critical loss in antitrust. *Antitrust Bulletin* 46(2), 339–355.

Eidgenössische Zollverwaltung (2016). Einnahmen aus dem Einkaufstourismus im Reiseverkehr der Schweizer von 2010 bis 2015 (in Millionen CHF). Retrieved from https://de.statista.com/statistik/daten/studie/520459/umfrage/einnahmen-aus-dem-einkaufstourismus-im-reiseverkehr-der-schweizer/.

Ellickson, P. B. (2007). Does sutton apply to supermarkets? *The RAND Journal of Economics* 38(1), 43–59.

Ellickson, P. B. and Grieco, P. L. (2013). Wal-mart and the geography of grocery retailing. *Journal of Urban Economics* 75, 1–14.

Ellickson, P. B., Houghton, S. and Timmins, C. (2013). Estimating network economies in retail chains: a revealed preference approach. *The RAND Journal of Economics* 44(2), 169–193.

Federal Trade Commission and U.S. Department of Justice (2010). Horizontal merger guidelines.

Gaynor, M. S., Kleiner, S. A. and Vogt, W. B. (2013). A structural approach to market definition with an application to the hospital industry. *The Journal of Industrial Economics* 61(2), 243–289.

Geroski, P. (1995). What do we know about entry? *International Journal of Industrial Organization* 13(4), 421–440.

Grieco, P. L. (2014). Discrete games with flexible information structures: An application to local grocery markets. *The RAND Journal of Economics* 45(2), 303–340.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis* 20.

Hetschko, C., Schöb, R. and Wolf, T. (2016). Income support, (un-)employment and well-being. *CESifo Working Paper Series No. 6016* .

Holmes, T. J. (2011). The diffusion of wal-mart and economies of density. *Econometrica* 79(1), 253–302.

Hong, G. and Raudenbush, S. W. (2013). *Handbook of Causal Analysis for Social Research*, NY: Springer, chap. Heterogeneous agents, social interactions, and causal inference. 331–352.

Huber, M. and Steinmayr, A. (2017). A framework for separating individual treatment effects from spillover, interaction, and general equilibrium effects. *IZA Discussion Papers 10648*, Institute for the Study of Labor (IZA).

Jia, P. (2008). What happens when wal-mart comes to town: An empirical analysis of the discount retailing industry. *Econometrica* 76(6), 1263–1316.

Katz, M. L. and Shapiro, C. (2003). Critical loss: Let's tell the whole story. *Antitrust* 17(2), 49–56.

Matsa, D. A. (2011). Competition and product quality in the supermarket industry. *Quarterly Journal of Economics* 126, 1539–1591.

Motta, M. (2004). *Competition Policy: Theory and Practice*. Cambridge University Press.

Neumark, D., Zhan, J. and Ciccarella, S. (2008). The effects of wal-mart on local labor markets. *Journal of Urban Economics* 63(2), 405–430.

Nishida, M. (2014). Estimating a model of strategic network choice: The convenience-store industry in okinawa. *Marketing Science* .

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1), 41–55.

Satyanath, S., Voigtländer, N. and Voth, H.-J. (2017). Bowling for fascism: Social capital and the rise of the nazi party. *Journal of Political Economy* 125(2), 478–526.

Stanton, C. T. and Thomas, C. (2016). Landing the first job: The value of intermediaries in online hiring. *The Review of Economic Studies* 83(2), 810–854.

Sutton, J. (1997). Gibrat's legacy. *Journal of Economic Literature* XXXV, 40–59.

The European Commission (1997). Commission notice on the definition of relevant market for the purposes of community competition law. *Official Journal* C(372).

Zhao, Q. and Percival, D. (2017). Entropy balancing is doubly robust. *Journal of Causal Inference* 5(1).