CO-LOCATION, GOOD, BAD OR BOTH: HOW DO NEW ENTRIES OF DISCOUNT VARIETY STORES AFFECT LOCAL GROCERY BUSINESSES?

Charlotte B. Evensen
Oslo Economics

Frode Steen
Norwegian School of Economics and CEPR

Simen A. Ulsaker
Oslo Economics and Norwegian School of Economics

Abstract
We analyse 69 entries and relocations by the largest Norwegian discount variety chain Europris during the period 2016 to 2019, and measure how its location choices affect local grocery stores' performance. We use detailed data from a major Norwegian grocery chain, which enables us to combine local grocery stores' sales and traffic with travelling distance to new or relocated Europris stores, and a two-way fixed effects strategy. Our findings suggest that entries and relocations have significant effects and that the sign of the effect depends on the distance between the stores, creating a non-linear relationship between the effect of entry and the distance between the stores. Sufficiently close entries and relocations increase local demand since more customers are attracted to the market, but, as the distance increases, the competitive effect of a new discount variety store dominates, and local grocery store sales and traffic are reduced. As we move further away, the entry effect is gradually reduced to zero. (JEL: L10, L21, L66, R30)

Keywords: Retail economics, local competition effects, positive agglomeration forces, grocery markets.

The editor in charge of this paper was Pierre Dubois.

Acknowledgments: We are grateful to comments from the editor (Pierre Dubois) and referees that have improved the paper. We are also grateful to the FOOD research project at Norwegian School of Economics, and the grocery chain for data access.

E-mail: cbe@osloeconomics.no (Evensen); frode.steen@nhh.no (Steen); sau@osloeconomics.no (Ulsaker)

1. Introduction

In this paper, we explore the rise of discount variety retail stores and how this has changed the competition situation for grocery stores. Over time, grocery stores have broadened their product range to include everything from books to consumer electronics. Likewise, we see a growing trend where previously specialised retailers like "dollar stores" and general hardware stores add groceries to their product range. In 2019, the American discount variety chain Dollar General expanded its product range to also include fresh grocery products, and since 2003, they have offered food products in a number of stores.[1] Today, Dollar General delivers grocery products to more than 9,000 of its total of 16,500 locations.

Moreover, we see a strong trend in the retail sector towards stores co-locating in shopping centres and business areas. In this new retail landscape, where different chains complement each other but also compete, store location choices have become less obvious. On one hand, differences in the product ranges of stores might lead to an increase in traffic when they are located very close to each other. On the other hand, increased local competition for products offered by both chains reduces incentives for co-location.

To understand how this new mix of product ranges and reduced retail chain specialisation affect store location, we analyse the location choices of Europris, the largest discount variety chain in Norway. In particular, we investigate how its location choices affect one of the largest grocery chains in Norway. Europris has been one of the most successful retail chains in Norway, establishing a number of new stores across the country. The grocery chain in question is one of the leading discount grocery retailers in Norway. It has a share of over 20% of the national market and has a presence in all major regions of Norway. Europris offers more than one-third of the grocery chain's product categories, contributing to as much as one-fourth of the grocery chain's total turnover in terms of sales.

Benefiting from a detailed data set covering all transactions at the grocery stores before and after the arrival of competing discount variety stores, we employ a two-way fixed effects approach to estimate the effect of entries. More specifically, we analyse how sales and customer traffic in local stores within the grocery chain are affected by Europris' establishment of stores and relocations as compared to a large control group of grocery stores that are not affected by changes in Europris locations. In all models, we control for local competition and municipality level demographics. We also have detailed information about the product overlap between Europris and the grocery chain, allowing us to estimate separate effects for products that are offered by both chains and products that are only offered by the grocery chain.

In the cases we consider, an incumbent grocery store may be affected in two ways by the establishment of a discount variety store with partially overlapping product

1. The first store appeared in 1939, and in 1955, they took the name Dollar General. Hence, it was not until 64 years had passed that they expanded into food products.

ranges. On the one hand, because the stores only compete on a subset of their product categories, the grocery store may gain new customers due to the increased quality of their location as a result of the complementarities across stores. We consider this as an increase in the extensive margin: as long as the entering store is differentiated with regard to product range, this co-location effect should be positive. This positive effect of establishment should be stronger the closer the new establishment is located to the incumbent grocery store, and maximised if co-location allows for one-stop shopping. On the other hand, the entry will increase competition for the product categories offered by both stores. This can be interpreted as a reduction in the intensive margin: some of the incumbent's existing customers may choose to purchase some products that they used to buy at the incumbent grocery store from the entrant discount variety store.[2] This effect will be negative, and it will be stronger the closer the new establishment is to the grocery store.

The net effect of the two effects outlined above is not clear. Furthermore, while we expect both the positive and negative effects to decrease in size with distance, they may do so at different rates. For example, it may be that the agglomeration effect is important only when the stores are fairly close, while the competition effect continues to be important also when the distance is relatively large.

In our empirical analysis, we find that one-stop shopping leads to positive agglomeration effects, increasing local demand when new stores enter. We also find clear evidence of a competitive effect that decreases with the distance between the stores. What we find particularly intriguing is that our results suggest that the interplay of the positive agglomeration effect and the negative competitive effect is such that the net effect of entry depends on the distance between the stores.

When we distinguish between the new entries that allow for one-stop shopping and those that require customers to stop twice, we see a distinct pattern: one-stop shopping increases sales by around 10%, indicating that the positive agglomeration effects dominate, whereas entries that require customers to stop twice have a negative impact on the incumbents' sales (close to -5%), suggesting that the negative competition effect dominates. To explore the effect of distance between the incumbent and the new entry in more detail, we proceed by splitting the two-stop shopping entries into different distance bins and re-estimating our models. Our results now indicate an interesting pattern. When we move away from one-stop shopping and up to a distance of two kilometres, we find a small negative effect on sales from new entries. For entries between two and five kilometres away the negative effect is the largest (around -9%), while it becomes smaller and ultimately fades away for entries even further away.

We attribute this pattern to the interplay of the two margins. Our results suggest that for the entries relatively close by (250 metres to two kilometers), the extensive margin effect of higher local demand still has a significant influence, though the intensive margin effect of fiercer competition dominates. As we move further away

2. The increased competition might also affect prices, but in our case the incumbent is already applying national prices, and thus the effect of the new store should come through changes in sales, something that is also confirmed by our analysis.

(in our case beyond two kilometers), the competition effect is significantly more important than the agglomeration effect, generating the maximum negative overall effect. Moving even further away in distance, the net effect goes towards zero, which is what we would expect given that both effects should taper off eventually. This suggests a non-linear pattern in the net effects from entry or relocation.

We provide a detailed assessment of our empirical strategy and perform a number of robustness checks. We also consider the recent literature on heteregenous treatment effects and variation in treatment timing. The assessment provide support for a causal interpretation of our empirical results.

To gain some additional insight into the mechanisms at play, we also develop a simple theoretical model that fits our empirical case closely. Using a framework inspired by Hotelling (1929), we consider how an incumbent store is affected by the entry of a competitor with partly overlapping product assortment in its vicinity. We find that the overall effect on sales may have a non-linear shape similar to what we observe in the empirical analysis.

Related literature. Stahl (1982) was one of the first to model the trend towards co-location and one-stop shopping behaviour theoretically. He models how the changes in the sellers' market demand influence location choices. In particular, he decomposes two effects: a negative substitution effect generated by competition for consumer demand and a positive market area effect generated by the joint location of sellers. If the increase in demand from joining the bigger market is higher than the effect of fiercer competition, co-location becomes the optimal choice. This will, in turn, become a positive externality for the incumbents already there. Stahl finds that co-location is an equilibrium outcome as long as customers are choosy enough about the variety of commodities.[3]

Our study relates to the empirical literature on store choices. Messinger and Narasimhan (1997) formulate and estimate a model on grocery data that aims to explain the growth in one-stop shopping. Using U.S. data, they find that increased income and reduced store operation costs have both increased supermarkets' assortments and the gains one-stop shopping entails. Bell et al. (1998) model store choice behaviour based on fixed and variable costs of shopping, attributing the former to the shopping list (products and quantities) and the latter to travel cost and store loyalty. They apply the model to data from a large U.S. city, and find support for fixed cost – shopping list heterogeneity being a major factor behind store choices. Fox et al. (2004) undertake an exploratory analysis, estimating a model based on consumer-reported data for purchases to understand how marketing policies affect shopping behaviour across retail store formats. Vitorino (2012) looks at how positive and negative spillovers between firms affect location choices and finds empirical

3. There are several theoretical studies modelling store choice and store location. Beggs (1994) looks at the rationale for malls rather than large department stores by modelling demand and pricing complementarities. Smith and Hay (2012) model competition between shopping centres, in particular how agglomeration effects between products are accommodated through different organisational structures and to what extent competition in prices and product quality is internalised.

support for firms co-locating despite potential business-stealing effects. Picone et al. (2009) suggest that, even if competitive forces make firms prefer distancing, they might end up co-locating because of few location options. Not surprisingly, this seems to be a more likely outcome among firms selling differentiated products. Related to the questions on store choices, Thomassen et al. (2017) study pricing in supermarkets. They estimate cross-category pricing effects, and find that the effects are stronger as more consumers prefer one-stop shopping. This has to do with these consumers being inclined to switch all their purchases to another store in response to a price change in one product category. Since supermarkets fully internalise the cross-category pricing effects (in contrast to specialised stores), one-stop shopping contributes to greater price competition.

Several empirical studies have analysed spatial competition between retail outlets more generally. Lindsey et al. (1991) analyse the video-cassette-retail market in Alberta to understand product variety and pricing. In a more recent study of the video-retail market, Seim (2006) finds empirical support for firms using spatial differentiation in order to reduce local competition. Smith (2004) estimates consumer choice in the UK supermarket sector using data on profit margins to deduce price parameters in consumer utility. Davis (2004) estimates a demand model where products are location-specific and consumers have preferences as regards geographic proximity and store/product characteristics, to understand substitution patterns between U.S. cinemas. Houde (2012) estimates a structural model of spatial competition using consumers' commuting paths as instruments for the consumers' locations in a Hotelling-like model, using data from the Quebec City retail gasoline market. Turolla (2016) estimates the intensity of competition in the French grocery retail sector. She builds a structural spatial competition model, where demand depends on both geography and heterogeneity of the customers' shopping lists. She recovers price-cost margins, and finds that the competitive pressure is very localised and depends on the presence of nearby competitors.

The paper is organised as follows. Section 2 discusses our empirical strategy, Section 3 presents the data and takes a first look at the market. In Section 4, we present and discuss our econometric results, and present a simple theory model. In Section 5, we assess our empirical strategy and perform a number of robustness checks. Section 6 concludes.

2. Empirical strategy

We want to explore how proximity to a discount variety store (in our case, Europris) affects grocery store sales. Our empirical strategy exploits the fact that, during our sample period, 69 Europris stores were established or relocated. Some grocery stores in our sample were affected by a Europris establishment in the sense that the distance to the closest Europris store changed after the establishment or relocation, while others were unaffected. This enables us to use a two-way fixed effects approach to estimate the effect on grocery store sales of having a discount variety store in the vicinity.

We refer to grocery stores that were affected by Europris establishments as treatment stores and to grocery stores that were unaffected as control stores. Among the grocery stores that experienced a change in the distance to the nearest Europris store, we focus on those that experienced a reduction in distance to the nearest Europris store.[4] We calculate the distance to the nearest Europris store for all the grocery stores in our data set in every week of our sample period. This means that, regardless of whether we look at relocations or new entries, we always consider a change from a given pre-distance. Hence, the estimated effect of relocation and new entry will be parallel, and we do not need to distinguish between them when evaluating the results. From now on, we will refer to both of them as establishments. Furthermore, while some of the treated grocery stores ended up with a Europris store next door after an establishment, other treated stores remained some distance away. This allows us to break down the effect of a Europris establishment by distance bins and to explore how the effect of having a discount variety store in the vicinity depends on the distance between the stores.

The main assumption allowing us to offer a causal interpretation of our results is that, conditional on time and store fixed-effects, the distance to the nearest Europris store is not correlated with unobservable factors that affect demand at the grocery stores. This assumption could be violated if, for example, Europris tended to locate their stores near grocery stores whose locations experienced increases in population or commercial activity, and further apart when there are other areas that experienced stronger development along these dimensions. Any systematic patterns along these lines would make a causal interpretation of the results difficult, since the distance between the grocery store and the nearest Europris store would be correlated with factors that also potentially affect grocery store demand. On the other hand, the location choices of the retail chains are also restricted by regulations and institutional features, arguably in an exogenous way. Specifically, the location of new stores is restricted by local zoning plans. These plans determine whether an area can be used for private housing and apartments, recreational or business purposes. Regulated by the Norwegian Plan and Building Act (2008), plans are made to facilitate local transport networks and the design of urban centres. [5] Relatively detailed plans for potential new retail store locations are part of this process, imposing restrictions on retail chains when choosing locations.

4. During our sample period, the locations of all grocery stores are fixed, implying that any changes in distance stem from Europris entries or relocations. Control stores are defined as any stores not experiencing a change in distance to the nearest Europris store during the sample period. In Section 5, we show that the results are robust to minor changes in the definition of the treatment and control group.

5. Additional regulations are formulated in the guidelines for coordinated housing, area and transport planning adopted in 2014 (Ministry of Local Government and Regional Development, 2014). More details on the precise content of the regulations, and the national intentions and policy can be found in guidelines from the Ministry of Local Government and Regional Development (2018 and 2019).

In Section 5, we provide a detailed assessment of our empirical strategy from several angles. The analysis in that section provides evidence in line with the main underlying assumption and lends support to a causal interpretation of our results.

In this article, we interpret changes in weekly sales turnover at the grocery stores following Europris establishments as increases in sales volume. It could also be the case that the grocery stores reacted by changing their prices when exposed to a new entry. This would imply that our results reflect both changes in volume and changes in prices. However, since both the grocery chain and Europris have national pricing strategies, we would not expect prices in local grocery stores to be affected by the distance to the closest Europris store.[6] To provide further evidence that price changes do not drive our results, we use price-quantity data to directly test whether the establishment of Europris stores influences prices at the grocery stores. As documented in Appendix F, we find no evidence of price effects, confirming that the effect of the establishment of Europris stores is manifested through changes in the sales volume and store visits in the grocery store (rather than in changes in prices).

3. Data and a first look at the market

3.1. Data

We combine data from several sources. The main data set used in our analyses consists of sales data received from the grocery chain. We have weekly sales data at the store-category level for all product categories, which means that we can both look at total weekly sales at the store level and consider sales of products that are also sold at Europris separately. In addition, we have data for weekly store visits. The sample period is from 11 January 2016 to 22 December 2019. For a subset of the product categories, we also have weekly price-quantity data at the product level. This allows us to investigate whether Europris establishments affect the price level at the grocery stores.

The next step is to compile geographical location data. We obtained data on the address, opening date and closing date (where applicable) of all Europris stores in Norway directly from the chain (Europris 2020). The data were received on 11 February 2019.[7] The sales data from the grocery chain also contain information about

6. Meile (2020) studies the price setting of Norwegian grocery retail chains and finds that the grocery chain we consider follows a national, uniform pricing strategy. Evidence of uniform national pricing is also found in Friberg, Steen and Ulsaker (2021). Regarding Europris, we examined the information on the chain's website. We find that the online prices (which at least apply to home delivery and in-store pickup) do not differ across stores and that weekly ads apply throughout the chain, suggesting that prices are decided centrally.

7. With updates on 2 July 2019 and 15 May 2020.

the grocery stores' addresses. The exact locations of the Europris and grocery stores were obtained through Google Maps Platform's Geocoding API.[8]

For a given grocery store in a given week, we aim to determine the distance and driving duration to the nearest Europris store. To calculate these metrics, we utilize the routing service of the Norwegian Public Roads Administration (NPRA).[9] For each grocery store and each week, we then record the distance and duration of the nearest Europris store that was open in the week in question.

We include a number of additional control variables in our regressions. Statistics Norway publishes quarterly municipality-level data on persons and land area, which we use to calculate population density (Statistics Norway 2021b). We also use yearly data at the municipality level for median after-tax income and the percentage of the population with higher education (Statistics Norway 2020a,b).

From Geodata (2021a), we have obtained a data set with yearly information on all grocery stores in Norway (from all chains), including information about store locations and store-level revenue. We use these data to calculate the Herfindahl-Hirschman Index (HHI) at the local market level, using market shares at both the store and the chain level. Because the municipality level may be too coarse to adequately capture retail competition, we calculate the HHIs at the 5 km $\times$ 5 km grid level. In Appendix N, we use several yearly grid-level variables to investigate whether Europris establishments are correlated with time-varying demand factors at the local level. In this analysis, yearly data on population and the number of buildings come from Statistics Norway (2021a), while yearly data on income and wealth were obtained from Geodata (2021b).

8. See https://developers.google.com/maps/documentation/geocoding/overview for documentation of this service. The locations were obtained on 15 October 2020.

9. See https://labs.vegdata.no/ruteplandoc/ for documentation of this routing service. The routing service was accessed on 15 October 2020, meaning that all durations and distances were calculated based on the road network as of that date. We also verify the locations on Google Maps to ensure that the driving distances are accurate representations of the actual distances between the stores. In some instances, manual adjustments are made to store assignments in different distance categories when they appear more fitting based on actual map distances.

3.2. A first look at the market

Discount variety retail in Norway. Among discount variety retailers in Norway, Europris is the largest player, with a market share of about 30%.[10] Since its foundation in 1992, both its revenues and the number of stores have grown steadily, reaching more than NOK 6 billion in revenue and 264 stores in 2019. While the compound annual grown rate for total retail was about 3% for 2012-2017, variety retail grew almost twice as fast, suggesting that with overlapping product ranges, grocery chains were losing market shares to variety retail (Europris 2018). Few other retail segments can look back at a similar increase in revenues in recent years. In the period we consider, 2016 to 2019, Europris grew by 7.73% annually. During the same period, they opened 37 stores, and relocated another 32 stores (Europris 2017; Europris 2019).[11]

According to the latest Shopper Trend report (Nielsen 2020), more than 50% of the respondents answered that they had bought groceries from a discount variety retailer in the last six months, and the most frequently visited store was Europris. A comparison of the assortment in Europris and the grocery chain shows that the extent of product overlap is large: as much as 35% of the grocery chain's product categories are also sold in Europris stores, and these product categories amount to 25% of the grocery chain's turnover.[12]

The Norwegian producers are more concentrated than producers in comparable grocery markets, e.g., Sweden. Together with particularly high tariff-barriers, this has led to very strong national brands, and though increasing, private labels have a relatively low share of the Norwegian grocery market.[13] This implies that the same products are often found in different stores – even across grocery chains. This is also the case for many of the products that are sold by both Europris and the grocery chain.

The grocery stores. Our data sample consists of 190 distinct grocery stores. The stores are distributed all over Norway, but only stores located in municipalities where Europris establishments or relocations took place during our sample period are included. Because retail competition is likely to function differently in city centres than in suburban and rural areas, we drop observations from the municipality of Oslo, Bergen city centre and Trondheim city centre.[14] The number of active grocery stores

10. The first and second runners-up, Biltema and Clas Ohlson have approximately 20% and 15%, respectively.

11. Europris established (relocated) 11 (11) stores in 2016, 11 (7) stores in 2017, 9 (8) in 2018 and 6 (6) in 2019, respectively.

12. To find the product overlap, we first looked up all the product categories that Europris offers online (such as detergents, filter coffee, and pick-and-mix candy). We then compared this to the data set we obtained from the grocery chain, which includes information about product categories.

13. Even as late as in 2022, the Norwegian private label share was as low as 18%. Neighboring countries as Denmark and Sweden had private label shares in 2022 of 32% and 28%, respectively. Source: Presentation FOOD 2023, March 13th 2023, Marie-Louise Riewerts, Nielsen IQ)

14. In some cases, whether or not entry reduces the distance between the grocery store and the nearest Europris store depends on the direction of travel or the exact route chosen. In the main analysis, we exclude

in a given week ranges from 149 to 180. Figure 1 below shows the distribution of the stores by distance to the nearest Europris store in the first and last sample week.

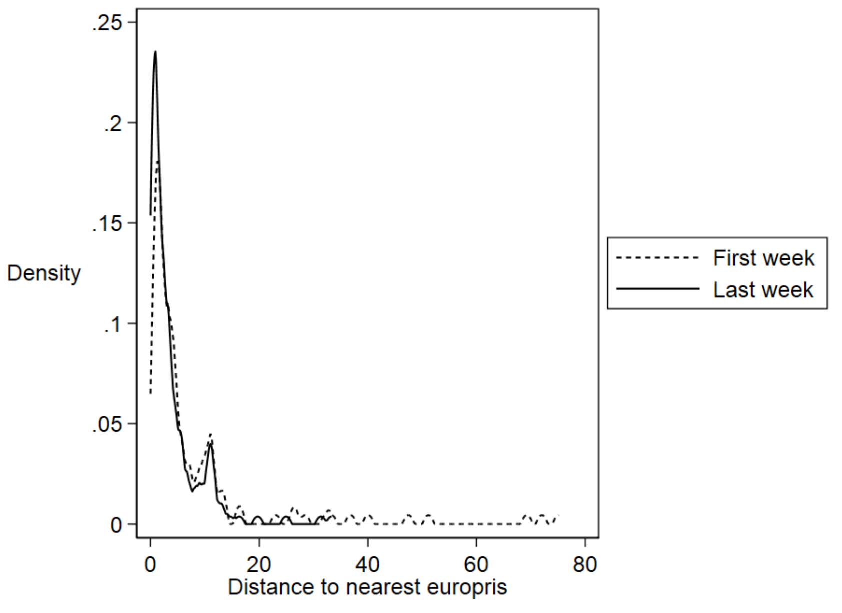


FIGURE 1. Density of grocery stores by distance to nearest Europris store

Most of the grocery stores are located within a few kilometres of a Europris store, and the distribution shifts slightly to the left over the period we consider. In the first week, 134 out of 149 grocery stores are closer than 15 km to Europris. In the last week, the same is true for 175 out of 180 stores.

A key distinction in our analysis is between one-stop and two-stop shopping. In some places, the grocery store and Europris are located close enough to one another for customers to reach both stores from the same parking area. We define one-stop shopping locations as those where the distance between the stores is 250 metres or less. In some of our analyses, we lump together all cases where the distance is above 250 metres as two-stop locations, while in other analyses we break down the two-stop locations into distance bins.

grocery stores that experience a Europris establishment in the vicinity but where it is unclear whether the distance to the closest Europris store was reduced or not after entry. In Section 5, we show that our results are robust to including grocery stores where there is some uncertainty about the effect of establishment on the distance to the closest Europris store. Furthermore, some grocery stores may be affected by more than one entry or relocation, which may complicate the interpretation of our results. In Section 5, we show that our main results are robust to excluding grocery stores that experience more than one establishment within a reasonable distance during the sample period.

The main variables of interest are activity indicators: sales and store traffic.[15] In our dataset, the average grocery store had on average NOK 1,232,886 in weekly sales, and 5,504 customers visiting per week, suggesting an annual turnover of USD 7.3 million and nearly 290,000 visits per year.[16]

Distance categories. As we argued above, the effect of establishment may depend on the distance between the stores. Hence, we define five distance categories. In Table 1, we summarise weekly sales and the store traffic across the five distance categories. The grocery stores in the "Same parking" category are located within 250 metres of a Europris store, which we define as close enough for the customers to visit both the grocery store and Europris in one stop. The other four categories require the customer to stop twice.

TABLE 1. Average store activity by distance category

	Weekly sales	Store traffic
Same parking	1 341 951	5613.08
250m-2km	1 275 654	5980.73
2km-5km	1 265 425	5516.03
5km-15km	1 157 070	5176.06
More than 15km	1 076 884	4071.88

Note: The numbers are weekly averages across all stores over the entire sample period by distance category. Stores that move from one distance category to another are weighted according to the fraction of weeks in each category.

The grocery stores that can be reached from the same parking area as a Europris store have the highest weekly sales and second highest store traffic, while the stores with the longest distance to a Europris store have the lowest turnover and store traffic. Overall, however, the differences are not large between the groups.

Store composition: Control and treatment groups. For descriptive purposes, we consider the 142 stores that are never affected by a Europris establishment as control stores and the 48 stores that at some point become affected as treatment stores.[17] As mentioned above, treatment stores are stores that experience a reduction in the distance to the nearest Europris during the sample period, while control stores are grocery stores that do not experience a change in the distance to the nearest Europris store during our sample period. Table C.1 summarises the distance statistics by treatment status.

15. Store traffic, as measured by the number of receipts, refers to the number of customers visiting per week.

16. The calculation is based on an average year of 52 weeks, and using the 2019 annual exchange rate of 8.81 (USD/NOK), Norges Bank: https://www.norges-bank.no/tema/Statistikk/Valutakurser/?tab=currency&id=USD

17. Table A.1 in Appendix A.1 shows the number of stores by treatment status and distance category.

TABLE 2. Distance statistics by treatment status

	Count	Mean	Sd	Min	Max	p25	p50	p75
Control group								
Distance	24738	4.43	4.76	0	40.07	1.30	2.96	6.04
Treatment group								
Pre-distance	9490	15.66	19.67	0.41	75.11	3.03	5.67	23.05
Post-distance	9490	3.11	6.32	0.00	31.37	0.31	1.01	2.20
Change	9490	12.55	17.87	0.34	71.76	1.95	3.61	13.18

Note: The table shows distance statistics for the control and treatment groups separately. For the control stores, the distance to the nearest Europris is constant, while for the treatment store, we show summary statistics both before and after the distance to the nearest Europris changes.

The average distance to Europris in the control group is 4.4 km, the shortest distance is 0.0 km and the longest distance is 40.1 km; 75% of the stores in the control group are located less than 6.0 km from a Europris store, 50% less than 3.0 km and 25% less than 1.3 km.

Compared to the control group stores, which have an average distance to the closest Europris store of 4.4 km, the treatment stores were on average less exposed to Europris prior to the establishments (15.7 km), but are on average more exposed to Europris in the post-establishment period (3.1 km).

Looking at the change within the treatment group, the relocations and new establishments led to an average reduction of approximately 12.6 km. In the pre-period, 50% of the grocery stores were located less than 5.7 km from a Europris store, while 25% were located less than 3.0 km away. In the post-period, 50% are located less than 1.0 km from Europris and 25% less than 310 metres. This shift is illustrated in Figure 2.

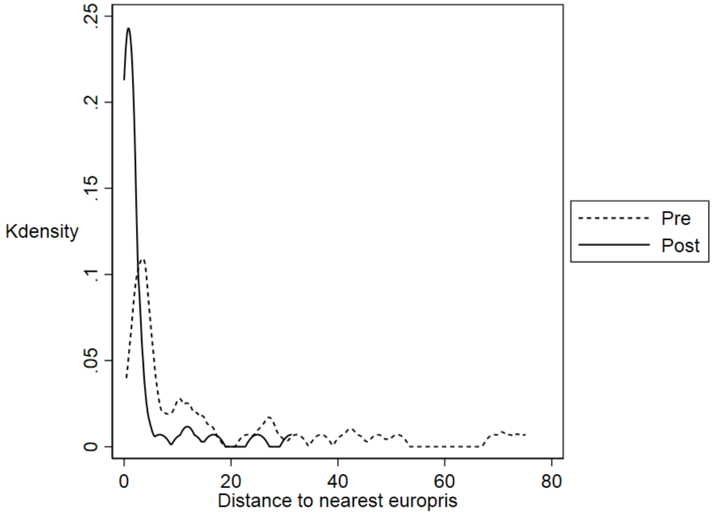


FIGURE 2. Density of grocery stores in the treatment group by distance

In the first week, 30 treatment stores were closer than 15 km from Europris, while 12 stores were further away. In the last week, 46 out of 48 treatment stores are located within 15 km of Europris. Table 3 presents the store activity measures by treatment status and whether the established Europris stores could be visited from the same parking area as the grocery stores in the post-period.

Store activity by treatment status. As many as 25% of our grocery stores ended up with a Europris store much closer than previously. As we saw from Figure 2, the shift was significant for most stores. To what extent does this shift result in a change in the activity level? In Table 3 below, we explore the change in our two measures of store activity.

TABLE 3. Average weekly sales and store traffic

	Pre-establishment	Post-establishment	Overall	Change
Weekly sales				
Control	-	-	1 196 403	-
Treatment One stop	1 381 962	1 622 909	1 546 380	17.4%
Treatment Two stops	1 189 723	1 302 516	1 264 461	9.5%
Store traffic				
Control	-	-	5508.14	-
Treatment One stop	6015.6	6552.3	6350.6	8.9%
Treatment Two stops	4975.3	5130.3	5170.7	3.1%

Note: The table shows before/after numbers, as well as overall numbers for stores in the treatment group for the entire sample period. The percentage figures measure change in sales between pre- and post-establishment. The table also shows overall figures for stores that are always control stores (since the control group consists of both grocery stores never affected by Europris entry and grocery stores not<w yet affected by entry, overall numbers across these two groups of control stores cannot be defined).

We find that, for both measures, activity increases after the change. There is also a distinct pattern where the effect is between two and three times higher for the one-stop establishments, as compared to cases where customers need to drive between the two stores. However, these figures only represent a before-after effect. Obviously, this change might be correlated with market growth stemming from other sources. The table also shows control group averages, and in the next section we will use a two-way fixed effects approach where we use activity development in the 142 non-affected stores to control for general market growth. Note that we will also account for the latter group's distance to the nearest Europris stores and store heterogeneity through store fixed effects.

In Tables B.1 and B.2 in Appendix A.2, we show sales and store traffic by treatment status and post-period distance categories. We observe that, generally, the change following a new establishment falls with distance. For weekly store traffic we even see negative numbers for the 2km-5km bin, or basically no change (0.93%) for the corresponding bin for weekly sales. In Table C.1, we present descriptive statistics for the control variables by treatment status.

4. A two-way fixed effects analysis of co-location effects

The descriptive analysis above suggested that the arrival of a new discount variety store close by affects the activity level of the grocery stores. In fact, to the extent that we could see some clear patterns, co-location – and, in particular, co-location enabling one-stop shopping – increased the incumbent grocery stores' traffic and sales. Now, we investigate these effects econometrically, controlling for both the development in these measures over time in grocery stores not affected by establishments, and the competitive environment faced by the different stores and the demographics of the area.

4.1. Two-way fixed effects analysis disregarding product heterogeneity

Our two-way fixed effects model includes several control variables for local competition and local demographics. We estimate the following generic model:

$$ln(y_{it}) = \alpha_i + \lambda_t + \eta X_{it} + \beta D_{it} + \varepsilon_{it} \tag{1}$$

Where y_{it} is a measure of activity, either weekly sales or store traffic. Subscript i refers to store and t refers to week. The matrix X_{it} consists of several local control variables. Municipality-level demographics are included through inhabitants per square kilometre (measured each quarter), and annual data on income and the share of the population with higher education. In addition, we include measures of retail competition. Because concentration measured at the municipality level may be too coarse to properly capture the spatial aspect of retail competition, we calculate HHI-measures at the 5 km $\times$ 5 km grid level. We calculate HHI both on store-, chain- and umbrella chain level. We also include population by store at the grid level. We include fixed effects for store (α_i) and week-year (λ_t), and cluster standard errors on the store level.

Our main parameter of interest is β, which measures the effect of a reduction in distance to the nearest Europris store for the stores in the treatment group. D_{it} is thus our treatment variable that, for store i in the treatment group, takes the value 0 prior to the Europris establishment, and 1 after.

In Table 4, we report estimates of the overall effect of a reduction in distance to Europris on the grocery stores for the two activity measures.

TABLE 4. Effect of establishment

	Log weekly sales	Log weekly store traffic
Establishment	-0.012	-0.0172
	(0.0215)	(0.0184)
Store FE	✓	✓
Week-year FE	✓	✓
Control variables	✓	✓
N	32204	34204
r2	0.835	0.839

Note: The table reports the coefficient on D_{it} from estimation of the model described in (1). In the first column, the dependent variable is the log of weekly sales at the store level. In the second column, the dependent variable is the log of the total number of weekly store visits. Clustered (store level) standard errors in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Using this approach, we find no significant effects of the establishment of Europris stores. However, this is an overall average effect that combines the effects from both nearby establishments and more distant ones. As we discussed in the introduction, it could be that the sign of the effect depends on the distance between the grocery store and the newly established Europris store. We could then fail to find an overall effect even if there are actually significant effects for the different co-location distance bins. Hence, we next differentiate the treatment effect into bins for different co-location distances, and extend the model to allow for more treatment dummy variables:

$$ln(y_{it}) = \alpha_i + \lambda_t + \eta X_{it} + \sum_b \beta_b D_{itb} + \varepsilon_{it} \tag{2}$$

Now, each β_b refers to a separate distance bin. We start by differentiating between one-stop and two-stop shopping: Comparing distance bin 1 to distance bins 2 to 5 (as defined in Table 1). In Table 5, we show the results.

TABLE 5. Effect of establishment by distance

	Log weekly sales	Log weekly store traffic
One stop	0.104**	0.0654
	(0.0489)	(0.0413)
Two stops	-0.0489**	-0.0435**
	(0.0199)	(0.0182)
Store FE	✓	✓
Week-year FE	✓	✓
Control variables	✓	✓
N	34204	34204
r2	0.838	0.842

Note: The table reports the coefficient on D_{itb} from the estimation of the model described in (2). In the first column, the dependent variable is the log of weekly sales at the store level. In the second column, the dependent variable is the log of the total number of weekly store visits. Clustered (store level) standard errors in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

In line with the descriptive figures in Table 3, we now obtain a very clear result. For both activity measures, we find that one-stop co-location increases the grocery stores' turnover and store traffic in the range of 7% to 10%. However, the results reported in Table 5 suggest that the net effect of establishment is negative for the grocery stores where one-stop shopping is not possible. This applies to both sales and traffic, with the reduced activity being in the order of -5% to -4%. Thus, when accounting for underlying time trends using a control group and when including control variables, the apparent positive effect observed in Table 3 only holds for grocery stores where one-stop shopping is possible after the establishment of a Europris store. For the other grocery stores, the estimated effect is negative.

That co-location can be beneficial for the incumbent grocery store is in line with most theoretical models considering one-stop shopping. The positive effect found for establishments that enable one-stop shopping suggests that the net effect of the positive agglomeration effect (what we refer to as the extensive margin) and the negative competition effect (what we refer to as the intensive margin) is positive for these stores. We expect that both effects are also present when two stops are required to visit both a grocery store and a Europris store, but that their relative magnitude may depend on the distance between the stores. Our next step is therefore to differentiate the treatment effects even further, allowing for different distance bins for the "Two stops" group of stores. Now we estimate separate effects for all our five distance bins. The results are shown in Table 6.

TABLE 6. Effect of establishment by distance

	Log weekly sales	Log weekly store traffic
Same parking	0.104**	0.0657
	(0.0490)	(0.0413)
250m - 2km	-0.0413	-0.0351
	(0.0274)	(0.0244)
2km-5km	-0.0882***	-0.0871***
	(0.0189)	(0.0238)
5km-15km	-0.0394*	-0.0343
	(0.0227)	(0.0218)
More than 15km	-0.0239	-0.0116
	(0.0176)	(0.0253)
Store FE	✓	✓
Week-year FE	✓	✓
Control variables	✓	✓
N	34204	34204
r2	0.839	0.842

Note: The table reports the coefficient on the variables D_{itb} from the estimation of the model described in (2). In the first column, the dependent variable is the log of weekly sales at the store level. In the second column, the dependent variable is the log of the total number of weekly store visits. Clustered (store level) standard errors in parentheses.
* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Now an interesting pattern emerges. The effect of an establishment is positive when the stores can be reached from the same parking area. When the stores are between 250 metres and two kilometres apart, there is no statistically significant effect. When the distance between the stores is between two and five kilometres, an establishment reduces sales by 9%. When the distance is even larger, the effect diminishes and becomes statistically insignificant for stores where the distance is more than 15 kilometres.[18] Figure 3 illustrates how the estimated effects on grocery stores' weekly sales and traffic vary non-monotonically with distance to the new Europris store, creating a non-linear "S-shaped" pattern for both activity measures.

18. Since the number of treatment stores varies across bins (see Table A.1 in the Appendix A), we estimate several models with alternative distance-bins definitions. The results are discussed in the robustness discussion in Section 5.3, and shown in Appendix I and in FigureI.1, and confirm the overall picture in Table 6.

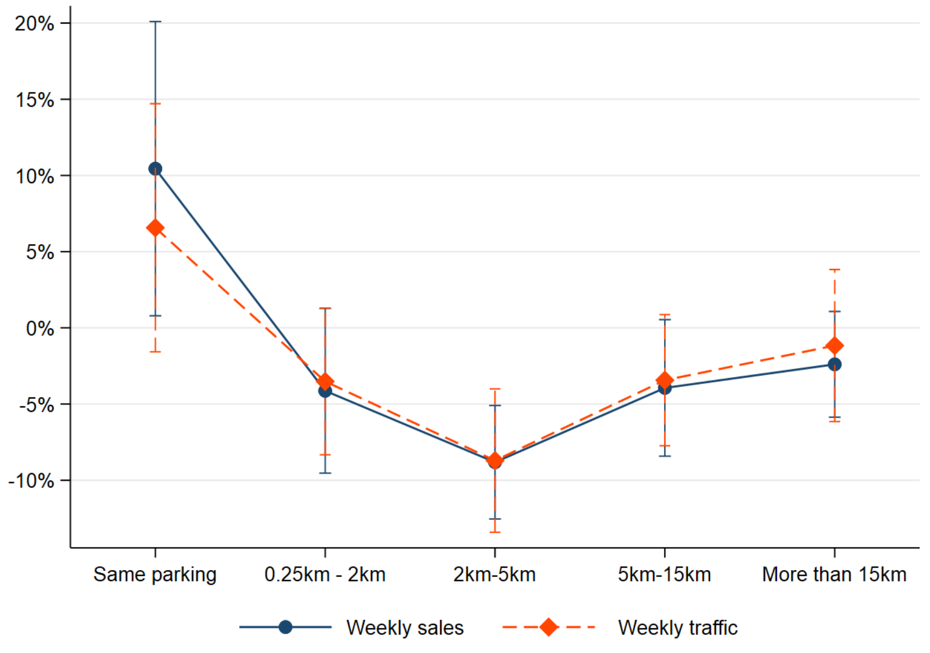


FIGURE 3. Illustration of estimated "S-shape".

Before exploring potential product heterogeneity in the next subsection, we briefly explore a potential heterogeneity within the two-stop entries besides different distances between our treatment stores and the entering Europris stores. In particular, if Europris co-locates with a rival grocery store from a competing chain, this might induce a local agglomeration effect with this competing store. In turn this can create an even stronger competition effect for grocery stores at a two-stop distance in our treatment group. To investigate this, we re-estimate the model, classifying Europris entries into two types: stand alone entries and entries next to a rival discount grocery store. The results are shown in table D.1 in Appendix D. As one would anticipate, the effect of Europris co-locating with a nearby discount grocery store magnify the competition effect somewhat, though only marginally and the difference is not significant.[19]

19. The discount category is a distinct segment and constitute close to two thirds of the Norwegian grocery market. The remaining one third constitutes supermarkets and local convenience stores. The three large grocery chains all offer one store concept in the discount segment and these compete on similar product ranges and prices. Hence, we narrow our attention to co-locations with competing discount grocery stores.

4.2. *Two-way fixed effects analysis accounting for product heterogeneity*

Clearly, we would expect to observe heterogeneous effects of Europris establishments depending on whether we look at competing or non-competing product categories. We now estimate our model where we allow the treatment effect to depend on the product type. We use the same distance bins, but now we estimate separate effects for all bins for the competing product categories (sold by both chains) and non-competing product categories (only sold by the grocery chain).[20] In Table 7, we show the results.

TABLE 7. Effect of establishment by product heterogeneity and distance category

	Non-competing	Competing
Same parking	0.111**	0.0869*
	(0.0481)	(0.0520)
250m-2km	-0.0349	-0.0603**
	(0.0278)	(0.0265)
2km-5km	-0.0818***	-0.107***
	(0.0182)	(0.0222)
5km-15km	-0.0378	-0.0430**
	(0.0236)	(0.0210)
More than 15km	-0.0212	-0.0303*
	(0.0178)	(0.0174)
Store FE	✓	✓
Week-year FE	✓	✓
Control variables	✓	✓
N	34204	34204
r2	0.833	0.846

Note: The dependent variable is the log of weekly sales at the store level. Clustered (store level) standard errors in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Separating competing and non-competing product categories, we find a similar pattern as we did for all products overall: one-stop shopping increases sales, suggesting that the extensive margin dominates. For other distance bins, the estimates are negative, suggesting that the competition effect prevails if the stores cannot be reached from the same parking area.

Furthermore, as we would intuitively anticipate, the competitive effect is stronger for products that are offered by both the incumbent grocery store and the entering discount variety store than for products that are only sold by the grocery store. While one-stop shopping increases sales also in the competing product categories, the effect is 2.4 percentage points higher for the non-competing product categories. For stores that have a new Europris store between two and five kilometres away,

20. The grocery store data contain information about category sales at different levels of aggregation. We consider an intermediate level of aggregation, which refers to categories such as ketchup, chocolate bars and detergents.

the effect of entry is -10.7% for competing product categories and -8.2% for non-competing categories, a difference of 2.5 percentage points. Additionally, we now find a negative and significant parameter for the competing product categories for the distance bin “250m-2km” which is both bigger (-6.0%) and now statistically significant, as opposed to what we found above for all products.

In Figure 4, we illustrate how the effects on weekly grocery sales vary with distance to the new Europris store for competing and non-competing product categories separately. We observe an “S-shaped” pattern similar to what we see in Figure 3.

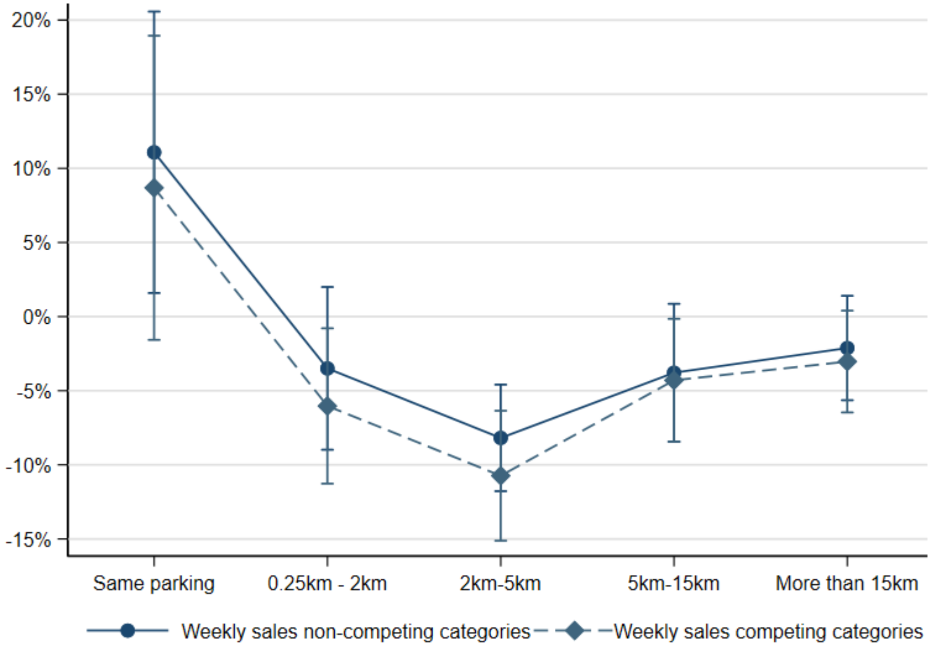


FIGURE 4. Illustration of estimated “S-shape” for competing and non-competing product categories.

In Table 7, we look at all competing and non-competing product categories. To explore the individual effects for some particularly relevant categories, we estimate the model for product categories where the grocery chain and Europris clearly compete, and for product categories where there is no competition. First, we estimate the effect for candy, coffee and detergent, categories that are known to be important in the Europris product portfolio, and where a number of strong national brands suggest that the products sold in Europris and the grocery chain really do compete. The results are reported in Table 8.

TABLE 8. Log weekly sales - competing product categories

	Detergent	Coffee	Candy
Same parking	0.0687	0.0946*	0.0219
	(0.0643)	(0.0500)	(0.0970)
250m - 2km	-0.131***	-0.00707	-0.174***
	(0.0329)	(0.0328)	(0.0418)
2km-5km	-0.183***	-0.119***	-0.211***
	(0.0382)	(0.0228)	(0.0640)
5km-15km	-0.159***	-0.0973***	-0.147*
	(0.0408)	(0.0345)	(0.0771)
More than 15km	-0.109**	-0.0546**	0.108***
	(0.0474)	(0.0246)	(0.0386)
Store FE	✓	✓	✓
Week-year FE	✓	✓	✓
Control variables	✓	✓	✓
N	34201	34173	34112
r2	0.833	0.635	0.819

Note: The table reports the coefficient on the variables D_{itb} from the estimation of the model described in (2). The dependent variable is the log of weekly sales at the category-store level. Clustered (store level) standard errors in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

We still find a positive agglomeration effect for same-parking establishments, though only weakly significant for coffee. More noteworthy, we find much stronger competition effects. Already for establishments as close as 250m-2km, we see strong competition effects and, for the second category ("2km-5km"), the competition effects are strong (between -12% and -21%) and significant for all three product groups.

Turning now to product groups that are not sold in Europris stores, we estimate the effect for bread, fresh chicken and milk, and present the results in Table 9.

TABLE 9. Log weekly sales - non-competing product categories

	Bread	Fresh chicken	Milk
Same parking	0.110**	0.145**	0.124**
	(0.0500)	(0.0655)	(0.0552)
250m - 2km	-0.0339	-0.0426	-0.0284
	(0.0274)	(0.0408)	(0.0313)
2km-5km	-0.0897***	-0.0537*	-0.0944***
	(0.0288)	(0.0298)	(0.0219)
5km-15km	-0.0122	-0.0364	-0.0198
	(0.0245)	(0.0300)	(0.0199)
More than 15km	-0.0514	-0.00341	-0.00009
	(0.0512)	(0.0274)	(0.0165)
Store FE	✓	✓	✓
Week-year FE	✓	✓	✓
Control variables	✓	✓	✓
N	34201	34198	34201
r2	0.884	0.857	0.879

Note: The table reports the coefficient on the variables D_{itb} from the estimation of the model described in (2). The dependent variable is the log of weekly sales at the category-store level. Clustered (store level) standard errors in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

The results from Table 7 are amplified, and the local agglomeration effect in Table 9 now varies between 11% and 15%, as compared to 11% for the overall effect for non-competing product categories in Table 7. However, there is still evidence of a competition effect from establishments further away.

In sum, we find that when looking at all competing product categories, as well as at some central individual products, we see some evidence of agglomeration effects dominating the competitive effect for same parking establishments. This suggests that even with a relatively large product overlap as here – one third product overlap – positive agglomeration effects can dominate the competitive effect of co-location also within categories sold by both stores.

4.3. A simple theory model on the interplay of intensive and the extensive margins and co-location

We have now established empirically that whether the grocery store ends up being better or worse off after Europris's entry depends on the distance between the two stores. We also find an "S-shaped" pattern, where the effect depends non-monotonically on the distance between the two stores. If Europris ends up sufficiently close, the grocery store tends to benefit. In contrast, an establishment that does not bring Europris close enough appears to be harmful. We attribute these findings to the interplay of the extensive margin (increased local demand) and the intensive margin (fiercer competition and reduced purchases by existing customers). In Appendix O we develop a simple theoretical model that shows how decomposing the effect of entry into an extensive and an intensive margin provides an intuitive explanation of the results.

To see how this model predicts the interplay of the extensive and the intensive margin we provide an numerical illustration. Figure O.3 shows the effect of a Europris establishment on the grocery store's revenues. It shows the effects from the extensive the intensive margins and the total effect.[21]

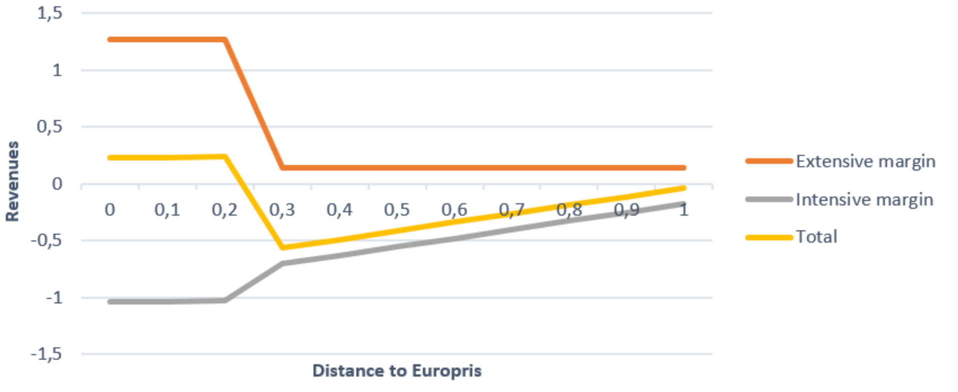


FIGURE 5. Intensive vs extensive margin

The effect of the extensive margin dominates when the distance between Europris and the grocery store is short. There are two main reasons for this. First, the customers do not have to make an additional stop to visit Europris, which attracts more customers. Second, the gain from attracting a new customer is greater than the loss from an exclusive customer turning into a shared customer. However, as the distance between the grocery store and Europris increases, the effect of the intensive margin becomes dominant. When shopping at both stores requires two stops, a Europris establishment might not attract sufficiently many customers for the grocery store to benefit from it. Eventually, the competition effect also fades away and the total effect approaches zero.

While the predictions from our modelling framework will be sensitive to the parameters chosen, we do find in Figure O.3 a very similar pattern to the pattern observed in our empirical analysis, as illustrated in, e.g., Figure 3. The observed and estimated relationship between treatment effect and distance to entry is thus consistent with the simple theoretical framework in Appendix O .

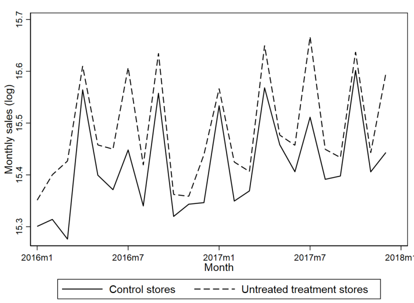
21. The illustration is made using reasonable parameter values, which, together with a more thorough discussion, can be found in Appendix O.

5. Assessing the empirical strategy

As already mentioned, the main assumption behind our empirical strategy is that, conditional on time and store fixed-effects, the distance to the closest Europris store is not related to other factors that affect demand at the grocery stores. In this section, we explore this underlying assumption from several angles. First, we take a closer look at store activity in treatment and control stores prior to treatment taking place, then we explore whether variation in treatment timing and heterogeneous effects affect our results, before we run a series of specification checks.

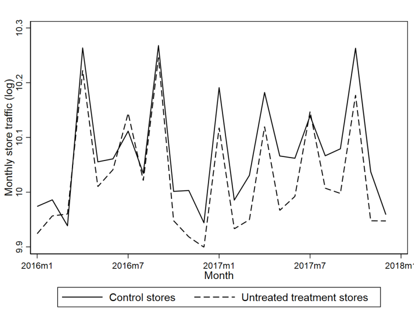
5.1. Assessing pre-trends in treatment and control stores

We now take a closer look at how sales and store traffic evolve over time in the treatment and control stores prior to the treatment taking place. We plot average monthly sales and store traffic in Figure 6. The panels *a* and *c* show the raw data, while panels *b* and *d* plot the trends after we have removed the effect of seasonality (by adding monthly dummies).

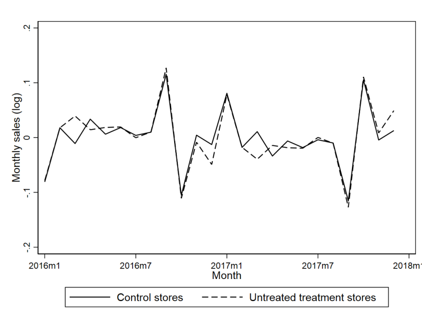


(A) Sales: Raw pre-trend

(B) Sales: Pre-trend controlled for seasonality



(C) Traffic: Raw pre-trend



(D) Traffic: Pre-trend controlled for seasonality

FIGURE 6. Pre-trends

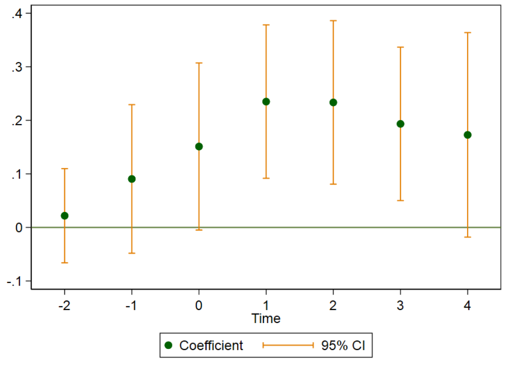
The dashed lines represent the average monthly sales and traffic in stores that never receive treatment, while the solid line shows the average monthly sales and traffic in treatment stores that have not yet received treatment. The trends in sales and traffic in the two groups share dynamics. This picture is particularly clear when controlling for seasonality, suggesting that the activity changes in the control and treatment groups display a common trend.[22]

Since treatments occur at different times for different stores, we also perform a Granger causality test. Following the approach used by Autor (2003), we now estimate:

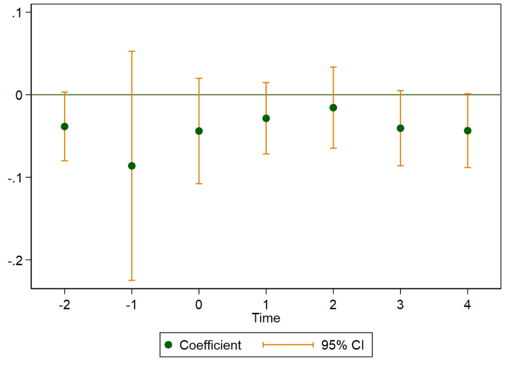
$$ln(y_{ti}) = \alpha_i + \lambda_t + \eta X_{it} + \sum_{\tau=-2}^{-1} \varphi_\tau D_i 1(t - T_i^* = \tau) + \sum_{\tau=0}^{4} \varphi_\tau D_i 1(t - T_i^* = \tau) + \varepsilon_{it} \quad (3)$$

The binary indicator D_i equals one if a store received treatment during the period we consider. We interact D_i with event-time dummies, $1(t - T_i^* = \tau)$. The dummies take on the value one when the time of observation (t) is $\tau \in [-2, 4]$ months from the treatment month (T_i^*). Earlier pre-months $(t - T_i^* \leq -2)$ serve as the baseline. Observations more than four months after a treatment are included through the dummy $1(t - T_i^* \geq 4)$. The coefficients on leads and lags of establishment are represented by φ_τ and φ_τ, respectively. If it is indeed the case that entries affect store activity, and not the other way around, we expect non-significant leads and significant lags. The results of the estimation are plotted below in Figure 7.

22. In Appendix E, we plot pre-trends separately for competing and non-competing categories. These plots reinforce the impression that pre-trends are parallel.



(A) One stop



(B) Two stops

FIGURE 7. Granger plot

Overall, the panels are consistent with what we observe in our econometric analysis. Neither the one-stop results (Panel A) nor the two-stop results (Panel B) show significant leads. This suggests that there are no anticipatory effects of establishments. In Panel a, we notice a higher point estimate in the month of establishment, which is also sustained in the subsequent months. The lags provide evidence of increased store activity in the post-periods. In Panel B, we do not observe clear shift at the month of establishment and the lags are individually statistically insignificant at the 5% level. Still, they are all below zero indicating a negative effect of establishment for this groups of stores reflecting the result reported in Table 5. The lack of a clearly visible treatment effect for the two-stops stores reflects the results we obtain when we divide the two-stop stores into more detailed distance bins. As reported in Table 6, only two-stop stores located between 2 km and 15 km from the closest Europris stores have significant treatment effects.

5.2. Staggered treatment and heterogeneous treatment effects

Recently, several methodological papers have shown that the two-way fixed effects regressions we use in this paper can yield biased results when there is variation in treatment timing and the treatment effects are heterogeneous (see, e.g., De Chaisemartin and d'Haultfoeuille (2020), Callaway and Sant'Anna (2021), and Sun and Abraham (2021)).

Our empirical case is an example of a "staggered design": The grocery stores are "treated" on different dates, and are treated at most once and remain treated thereafter. De Chaisemartin and d'Haultfoeuille (2020) show that, when treatment is staggered, the treatment effect estimated by a two-way fixed-effects regression, $\hat{\beta}_b$ in our case, is the weighted sum of average treatment effects (ATEs) across groups and time periods, where the groups are defined by the time treatment occurs. If the treatment effects are heterogeneous across stores and time, this weighted sum does not result in an unbiased estimate of the average treatment effect on the treated (ATT). Sun and Abraham (2021) propose an alternative to two-way fixed effects that is robust to staggered treatment and heterogeneous treatment effects. Their method is based on estimation of treatment effects at the cohort-time level, where cohorts are defined by the time period in which they received treatment. These treatment effects can then be aggregated to obtain the average treatment for the treated (ATT) for each (relative) period.

Our case is further complicated by the fact that our treatment variable is interacted with a discrete variable measuring the distance to the closest Europris after entry, which means that the two-way fixed effects regression incorporate comparisons not only between treatment and control stores, but also between stores in different treatment groups (that is, stores that vary with respect to the distance to the closest Europris after entry), which implies that also the different treatment groups must have the same evolution in counterfactual outcomes.[23]

As a first robustness measure, we estimate the model like the one reported in Table 5, but where we estimate separate models for each treatment group. That is, when estimating the effect of entry on stores in, e.g., the "One stop" treatment group, we only keep observations from stores in this treatment group and the control stores (i.e., the the stores that are never affected by Europris entries). This prevents stores in the two treatment groups to implicitly function as control stores for each other. The results are reported in Table G.1 in Appendix G. Reassuringly, the results from estimating two-way fixed-effects models one treatment group at a time yields results that are very similar to the results reported in Table 5. Because the methods described in De Chaisemartin and d'Haultfoeuille (2020) and Sun and Abraham (2021) are derived for the case of binary treatments, estimating separate models for each treatment group

23. Callaway, Goodman-Bacon, et al. (2021) discuss both multi-valued and continuous treatments in situations with staggered adoption and the implicit parallel trend assumptions in these cases. The causal effect they focus on is the effect of a marginal (in case of continuous treatments) or discrete (in the case of multi-valued treatments) increase in the treatment intensity.

in this way also facilitates an assessment of the robustness our results using these methods.

First, we calculate the weights associated with the coefficient reported in Table G.1 using the method described in De Chaisemartin and d'Haultfoeuille (2020). We see that in the "One stop" treatment group all the weights are positive, and that in "Two stops" treatment group bin the vast majority of the weights are positive and that the sum of the negative weights is close to zero. We furthermore report the metric $\underline{\hat{\sigma}}$ introduced in De Chaisemartin and d'Haultfoeuille (2020). This metric is the hypothetical standard deviation of the ATEs under which the average treatment on the treated (ATT) may actually have the opposite sign than that estimated coefficient. We see that for both treatment groups, this critical value is very large relative to the magnitude of the estimated coefficient, indicating that it is unlikely that the true ATTs have the opposite sign of the estimated coefficients. We also report the metric $\underline{\underline{\hat{\sigma}}}$, which is the hypothetical standard deviation of the ATEs under which *all* the treatment effects may have the opposite sign than that estimated coefficient. This metric is only relevant for the "Two stops" treatment group (because all weights are positive in the "One stop" treatment group), and we see that the metric is large.

TABLE 10. Weights associated with two-way fixed effects regressions results

Metric	One stop	Two stops
Number of weights	1846	3770
Number of positive weights	1846	3432
Number of negative weights	0	338
Sum of positive weights	1.0	1.0102
Sum of negative weights	0.0	-0.0102
$\underline{\hat{\sigma}}$	0.134	0.054
$\underline{\underline{\hat{\sigma}}}$	-	1.043

Note: This table reports metrics on the weights underlying the two-way fixed effects results reported in Table G.1. The metrics are calculated using the methods described in De Chaisemartin and d'Haultfoeuille (2020).

Second, in Figure 8, we use the method derived in Sun and Abraham (2021) to construct Granger plots of the same type as in Figure 7. Reassuringly, we find that the results from Sun and Abraham (2021)'s method are very similar to the results obtained by two-way fixed-effects shown in Figure 7.[24] Crucially, there are no significant leads. In fact, the coefficients in the pre-treatment periods are closer to zero with Sun and Abraham (2021)'s method than with two-way fixed effects.

24. The two-way fixed effects results reported in Figure 8 are derived from separate regression for the One stop and Two stops treatment groups, as in Table G.1.

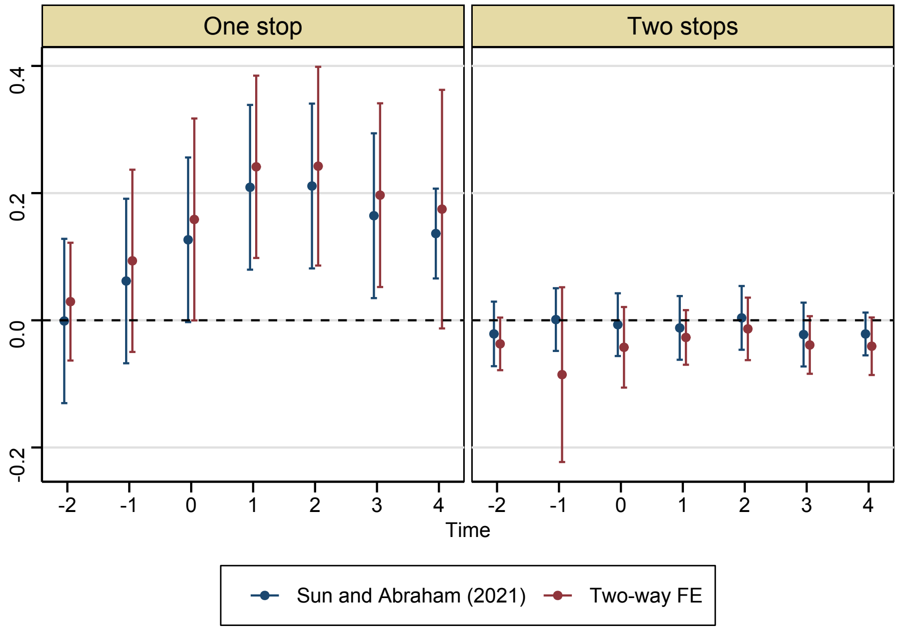


FIGURE 8. Granger plots - Two-way fixed effects and Sun and Abraham (2021)

Taken together, the analysis in this subsection suggests that variation in treatment timing and heterogeneous treatment effects do not seem to be a major threat to a causal interpretation of our empirical results.

5.3. Specification checks

In this section, we report results indicating that our results are robust to a series of specification checks.

Model without control variables. Considering that we have very detailed local control variables, it might be useful to know how much they affect our main results. We therefore re-estimate our main empirical specification excluding our control variables. These results are presented in Appendix H. Generally, we obtain the same results for the whole set of models. There are some marginal changes in significance levels but, in general, all our results are robust to these alternative specifications and data sets.

Models with alternative definitions of distance bins. A possible concern is that the treatment stores are not evenly distributed across the distance bins we use to investigate the relationship between distance and treatment effects. As can be seen from Table A.1 in Appendix A, most treatment stores are in the "Same parking" and "250m-2km" bins, and there are only three treatment stores in each of the "5km-15km" and "More than 15km" bins. In Appendix I we report results from re-estimating our model using several alternative definitions of the distance bins.

We begin by re-estimating the model reported in Table 6 but now with the two most distant bins ("5km-15km" and "More than 15km") merged. We see that the coefficient for the merged bin "More than 5km" lies between the coefficients of the two merged bins from the main specification. We also note that the coefficient is more precisely estimated and is significant at the five percent level. We also re-estimate a model where we merge the three most distant distance bins. We obtain a coefficient for the merged bin ("More than 2km") that lies between the coefficients for the three most distance bins in the main specification. The coefficient is significant at the five percent level. Finally, we estimate a six-bin model where we divide the "250m-2km" bin into "250m-1km" and "1km-2km". The former group then consists of 13 treatment stores, while the latter consists of 10 treatment stores. The results confirm that the overall effect is only positive for establishments very close to the grocery stores.

We illustrate the three models from the Appendix I and Table 6 in Figure I.1. The four different models confirm jointly the non-linear "S-shaped" pattern. We find it reassuring that these robustness checks corroborate what we see as the main results from our analyses: Establishments sufficiently close to the grocery store increases demand, while establishments further away reduce demand.

Model including separate linear trends by treatment group. We can also investigate the underlying common trends assumption by adding separate linear time trends for the stores in the control group and for each of the bins separating the treatment stores. If this produces significantly different results from the results reported in Table 6, our results could reflect diverging underlying trends rather than the effect of Europris establishment (Angrist and Pischke 2009, p. 238). Appendix J reports results from a model with separate time trends. Reassuringly, we see that the results are similar to the results in Table 6. We note that the point estimate for the positive effect at one-stop locations is higher when including linear trends by treatment group.

Model excluding stores that experience multiple establishments. In our models, we account for the changing distance to the nearest Europris. A potential problem is that some stores may be within a reasonable range of several entries over time. Hence, we may mix distance effects with multiple stores entries. To see whether our results are affected by multiple establishments, we exclude all grocery stores that experience two or more entries within 20 kilometres. This reduces our sample by approximately one third. Most stores that drop out of the sample are control stores. The control group is reduced by 56 stores and ends up consisting of 86 stores. In comparison, the treatment group is reduced by 9 and ends up with 39 stores. Hence, the great majority of treatment stores are only affected by a single entry. Moreover, the two most distant bins, with the fewest treatment stores, are not affected. In Appendix K, we show the results for this reduced sample, and the main conclusions stay the same. While all coefficients have the same sign as in the main specification, we note that the magnitude of the estimated effects is slightly larger than in the main specification and that the statistical significance is higher for several of the coefficients.

Model including stores with unclear treatment status. As mentioned in Footnote 14, we drop grocery stores from the main analysis where the effect of Europris establishment on travel distance is not entirely clear because the distance between the grocery store and the nearest Europris store depends on the direction of travel or the exact route chosen. As a further robustness check, we re-estimate our main specification including these stores. The results are presented in Appendix L. They are very much in line with the results from the main specification.

Model without control stores that are never treated. If the control stores that are never treated are inherently different from the treatment stores, they might not constitute an appropriate control group. Even though the assessment of the pre-trends in Section 5.1 suggests that the control and treatment stores display a common trend, we investigate this further by estimating a version of the model without using stores that are never treated as control stores. The results, which are reported in Appendix M are similar to our main result, indication that our results are not driven by differences in the control and treatment group.

5.4. Modelling Europris establishment

A causal interpretation of our empirical results hinges on the assumption that stores experiencing a reduction in the distance to the nearest Europris store have different underlying trends in store activity than stores in the control group. In Appendix N, we provide additional evidence in support of this assumption, complementing the analysis in Section 5.1. Specifically, we estimate models where we directly model the location choice of the Europris stores. The aim of the models is to investigate the extent to which the location choice of the Europris stores (and, by extension, the distance between the grocery stores and the nearest Europris store) can be explained by a set of time-varying factors that may also affect demand at the grocery stores. Both the dependent and independent variables we consider in this section are defined for geographical grids (of different resolutions) covering the municipalities where the grocery stores in our sample are located. Overall, we find no indication that the location choice of Europris stores co-varies with the time-varying demand factors at the grid cell level.

6. Conclusion

We analyse a number of entries and relocations by the Norwegian discount variety chain Europris during the period 2016 to 2019. We measure how location choices affect local grocery stores' sales and traffic, using a two-way fixed effects strategy and data from a large Norwegian grocery chain. We combine detailed data on traveling distance between new entries and grocery stores, and data on grocery store activity to measure the entry effects. The granularity of the data enables us to estimate separate effects for competing and non-competing product categories.

We find significant effects from entries and relocations. Moreover, our findings suggest an non-monotonic relationship between distance and store activity: sufficiently close entries increase local demand because more customers are attracted to the market, but, as the distance increases, the competitive effect of a new discount variety store dominates, and local grocery sales and traffic are reduced. As we move further away, the entry effect is gradually reduced to zero. We show that this empirical finding can be squared with a simple location theory model, showing a similar pattern. We explore our findings from several angles. We investigate the common trends assumption, explore whether staggered treatment and heterogenous treatment effects affect our results, before we run a series of specification checks. We find that our predictions are robust.

Most of the empirical literature accommodating local competition in retail markets treats local competition as a linear effect: the closer a competitor is located, the fiercer the competition (see, e.g., Seim (2006) and Picone et al. (2009)). In line with existing literature, we find that a competitive effect is present, but our results also suggest that this competition effect is dominated by a local and positive agglomeration effect, leading to more demand if the distance between stores stores is short enough. However, the agglomeration effect seems to be very local: as soon as the consumer has to travel even short distances between the stores, the agglomeration effect tapers off and the negative competition effect dominates.

Our results clearly support some of the insights from theory, such as Stahl's (1982) conjectures that, depending on product overlap and demand heterogeneity, co-location can be positive. Moreover, our findings are in line with what others have found, such as Vitorini (2012), who finds empirical support for firms co-locating despite potential business-stealing effects. Picone et al. (2009) find that co-location is more likely if the firms sell differentiated products. However, this does not necessarily imply that co-location requires maximal differentiation. Our results suggest that even a relatively large product overlap is compatible with co-location. We complement existing literature by providing evidence that the net effect of agglomeration forces and competitive pressure depends on the distance between the stores.

Our results are also relevant to the ongoing public debate about store location policies in several countries. Some countries (e.g., Denmark and Sweden) have imposed local competition regulations for new store locations to maximise local competition. Our results seem to support the development of larger areas where several shops can be established (e.g., in shopping centres) and share joint parking areas rather than regulating areas for single store establishments. The stores can anticipate higher local demand, although they will be exposed to a competitive effect from stores offering competing products. The first effect is obviously positive for the retail firms. The latter effect is not, but it is positive for consumers.

References

Angrist, Joshua D. and Jörn-Steffen Pischke (2009). *Mostly harmless econometrics: An empiricist's companion.* Princeton University Press.

Autor, David (2003). "Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing." *Journal of Labor Economics*, 21(1), 1-42

Beggs, Alan W. (1994). "Mergers and malls." *The Journal of Industrial Economics*, 42(4), 419-428

Bell, David R., Teck-Hua Ho and Christopher S. Tang (1998). "Determining where to shop: Fixed and variable cost of shopping." *Journal of Marketing Research*, 35(3), 352-369

Callaway, Brantly, and Pedro HC Sant'Anna (2021). "Difference-in-differences with multiple time periods." *Journal of Econometrics*, 225(2), 200-230.

Callaway, Brantly, Andrew Goodman-Bacon and Pedro HC Sant'Anna (2021). "Difference-in-differences with a continuous treatment." Working paper.

Davis, Peter (2006). "Competition in retail markets: Movie theaters." *The RAND Journal of Economics*, 37(4), 964-982

De Chaisemartin, Clément, and Xavier d'Haultfoeuille (2020). "Two-way fixed effects estimators with heterogeneous treatment effects."*The American Economic Review*, 110(9), 2964-96.

Europris (2017). "Annual Report 2017." https://s22.q4cdn.com/579442476/files/doc_downloads/general_meetings/2018/Europris-ASA-annual-report-2017.pdf

Europris (2018). "Capital markets day presentation." https://s22.q4cdn.com/579442476/files/doc_presentations/2018/11/Europris-Capital-markets-day-presentation-web.pdf

Europris (2019). "Annual Report 2019." https://s22.q4cdn.com/579442476/files/doc_financials/annual/Europris-ASA-annual-report-2019.pdf

Europris (2020). "Store data base."

Fox, Edward J., Alan L. Montgomery and Leonard M. Lodish (2004). "Consumer shopping and spending across retail formats." *The Journal of Business*, 77(S2), 25-60

Friberg, Richard, Frode Steen and Simen A. Ulsaker (2022). "Hump-shaped cross-price effects and the extensive margin in cross-border shopping." *American Economic Journal: Microeconomics*, 14(2), 408-438

Geodata, (2021a). "Grocery store data base." geodata.no

Geodata, (2021b). "Grid level demographics." geodata.no

Hotelling, Harold (1929). "Stability in competition." *The Economic Journal*, 39(153), 41-57

Houde, Jean-Francois (2012). "Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline." *The American Economic Review*, 102(5), 2147-2182

Lindsey, Robin, Balder V. Hohenbalken and Douglas S. West (1991). "Spatial price equilibrium with product variety, chain stores, and integer pricing: an empirical analysis." *The Canadian Journal of Economics*, 24(4), 900-922

Meile, Nathanael G., (2020). "Uniform pricing in Norwegian grocery retail; an empirical study on pricing strategies at Norwegian grocery retail chains". Master thesis, Norwegian School of Economics.

Messinger, Paul R. and Chakravarthi Narasimhan (1997). "A model of retail formats based on consumers' economizing on shopping time." *Marketing Science*, 16(1), 1-23

Ministry of Local Government and Regional Development (2014). "Statlige planretningslinjer for samordnet bolig-, areal- og transportplanlegging." https://www.regjeringen.no/no/dokumenter/Statlige-planretningslinjer-for-samordnet-bolig--areal--og-transportplanlegging/id2001539/

Ministry of Local Government and Regional Development (2018). "Reguleringsplan-veileder." https://www.regjeringen.no/contentassets/b1752a6a42f84a88a9595a4061956b43/no/pdfs/reguleringsplanveileder_sept_2018.pdf

Ministry of Local Government and Regional Development (2019). "Om hensyn i planlegging og kommunens myndighet og plikter." https://www.regjeringen.no/no/tema/plan-bygg-ogeiendom/plan--og-bygningsloven/plan/fagtema-i-planlegging1/hensyn-i-plan/id2412347/

Nielsen Norge (2020). "Dagligvarerapporten 2020, press release." https://www.nhosh.no/contentassets/1b23f7130e87488fab7aab5c7b8db5f5/nielsen-norge-pressemelding-dagligvarerapporten-2020_13022020.pdf

Picone, Gabriel A., Ridley, David B. and Zandbergen Paul A. (2009). "Distance decreases with differentiation: Strategic agglomeration by retailers." *International Journal of Industrial Organization* 27(3), 463-473

Seim, Katja (2006). "An empirical model of firm entry with endogenous product-type choices." The RAND Journal of Economics, 37(3), 619-640

Smith, Howard (2004). "Supermarket choice and supermarket competition in market equilibrium." *Review of Economic Studies*, 71(1), 235-263

Smith, Howard and Donald Hay (2012). "Streets, malls and supermarkets." *Journal of Economics and Management Strategy*, 14(1), 29-59

Stahl, Konrad (1982). "Differentiated products, consumer search, and location oligopoly." *The Journal of Industrial Economics*, 31, 97-113

Statistics Norway (2020a). "Income and wealth statistics for households – table 06944." https://www.ssb.no/en/statbank/table/06944/

Statistics Norway (2020b). "Educational attainment of the population – table 09429." https://www.ssb.no/en/statbank/table/09429

Statistics Norway (2021a). "Grid level statistics." https://www.ssb.no/natur-og-miljo/areal/artikler/kart-og-geodata-fra-ssb#rutenett

Statistics Norway (2021b). "Population and changes during the quarter – table 01222." https://www.ssb.no/en/statbank/table/01222

Sun, Liyang, and Sarah Abraham (2021). "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*, 225(2), 175-199.

The Norwegian Plan and Building Act 2008. "Lov om planlegging og byggesaksbehandling (LOV-2008-06-27-71)." https://lovdata.no/dokument/NL/lov/2008-06-27-71

Thomassen, Øyvind, Howard Smith, Stephen Seiler and Pasquale Schiraldi (2017). "Multi-category shopping and market power: A model of supermarket pricing." *The American Economic Review*, 117(8), 2308-2351

Turolla, Stephane (2016). "Spatial competition in the French supermarket industry." *Annals of Economics and Statistics*, 121/122, 213-259

Vitorino, Maria Ana (2012). "Empirical entry games with complementarities: An application to the shopping center industry." *Journal of Marketing Research*, 49(2), 175-191

Appendix A: Store number by distance category and treatment status

TABLE A.1. Number of stores

	Number of stores
Control stores	142
Treatment stores	
• Same parking	13
• 250m-1km	13
• 1km-2km	10
• 2km-5km	6
• 5km-15km	3
• More than 15km	3

Appendix B: Store activity distance bins and treatment status

TABLE B.1. Average weekly sales

	Pre-establishment	Post-establishment	Overall	Change
Control	-	-	1 196 403	-
Same parking	1 381.962	1 622 909	1 546.38	17.44 %
250m-2km	1 247.753	1 402 328	1 343.209	12.39 %
2km-5km	1 227.563	1 239 016	1 251.796	0.93 %
5km-15km	814 134	848 312	842 879	4.20 %
More than 15km	1 044 728	1 118 496	1 107 640	7.06 %

TABLE B.2. Average weekly store traffic

	Pre-establishment	Post-establishment	Overall	Change
Control	-	-	5508.14	-
Same parking	6015.57	6552.26	6350.64	8.92 %
250m-2km	5214.69	5481.83	5490.48	5.12 %
2km-5km	5421.41	5189.43	5382.92	-4.28 %
5km-15km	3636.14	3665.31	3687.39	0.80 %
More than 15km	3586.40	3781.66	3778.39	5.44 %

Appendix C: Descriptive statistics for control variables

TABLE C.1. Descriptive statistics overall and by treatment status

	Count	Mean	Sd	Min	Max	p25	p50	p75
Population density								
Overall	34228	420.36	529.34	2.70	1987.03	61.40	263.03	629.67
Control group	24738	510.44	541.23	2.70	1987.03	76.08	385.35	633.66
Treatment group	9490	185.55	413.24	2.70	1987.03	14.94	50.06	150.81
Income after tax								
Overall	34228	537017	49919	444000	689000	498000	528000	572000
Control group	24738	539904	51251	446000	689000	499000	528000	575000
Treatment group	9490	529492	45414	444000	660000	493000	523000	560000
Higher education								
Overall	34228	0.34	0.09	0.18	0.53	0.27	0.31	0.42
Control group	24738	0.36	0.09	0.18	0.53	0.29	0.35	0.42
Treatment group	9490	0.28	0.07	0.18	0.53	0.22	0.26	0.30
Population by store								
Overall	34228	1608.11	806.55	147.00	5618.00	1120.75	1559.50	1906.14
Control group	24738	1772.78	830.45	147.00	5618.00	1294.50	1647.59	2126.60
Treatment group	9490	1179.27	542.37	262.25	2376.50	690.60	1089.33	1596.52
HHI store level								
Overall	34228	0.27	0.24	0.00	1.00	0.09	0.20	0.35
Control group	24738	0.24	0.23	0.00	1.00	0.08	0.15	0.31
Treatment group	9490	0.35	0.24	0.03	1.00	0.22	0.32	0.45
HHI chain level								
Overall	34228	0.37	0.21	0.00	1.00	0.23	0.30	0.42
Control group	24738	0.35	0.20	0.00	1.00	0.22	0.29	0.38
Treatment group	9490	0.41	0.22	0.15	1.00	0.26	0.35	0.50
HHI umbrella level								
Overall	34228	0.51	0.21	0.00	1.00	0.36	0.43	0.56
Control group	24738	0.50	0.21	0.00	1.00	0.36	0.42	0.54
Treatment group	9490	0.54	0.22	0.30	1.00	0.38	0.50	0.58

Note: Population density, income after tax and higher education are on the municipality level, while population by store and the HHI-measures are for a 5 km $\times$ 5 km grid. The control variables are annual, except for population density, which is quarterly.

Appendix D: Exploring heterogeneity within two-stop entries

We extend our main model (equation 2) to impose a finer bin separation for two-stop stores:

$$ln(y_{it}) = \alpha_i + \lambda_t + \eta X_{it} + \beta_1 D_{it1} + \beta_2 \theta_{i1} D_{it2} + \beta_3 \theta_{i2} D_{it2} + \varepsilon_{it} \tag{D.1}$$

Now, β_1 captures the effects of being in the one-stop bin (D_{it1}). The indicator θ_{i1} is equal to one for two-stop stores (D_{it2}) that are exposed to stand-alone entries, and θ_{i2}

indicates accordingly entries in the vicinity (one-stop) of a rival grocery store. Hence, β_2 measures the effect of a two stop establishment for a stand-alone entry, and β_3 the effect of a an entry close to a rival store.

TABLE D.1. Effect of establishment by distance and entry type

	Log weekly sales
One stop	0.104**
	(0.0489)
Two stops, stand alone	-0.0407
	(0.0386)
Two stops, co-location	-0.0508**
	(0.0224)
N	34204
r2	0.838

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Appendix E: Pre-trends for competing and non-competing product categories

Pre-trends competing product categories

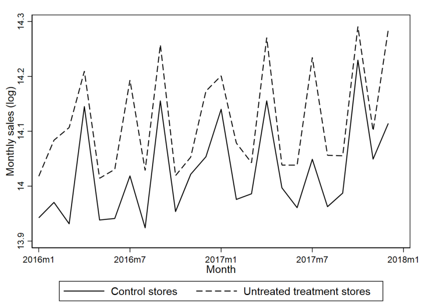


FIGURE E.1. Raw pre-trend

FIGURE E.2. Pre-trend controlled for seasonality

Pre-trends non-competing product categories

FIGURE E.3. Raw pre-trend

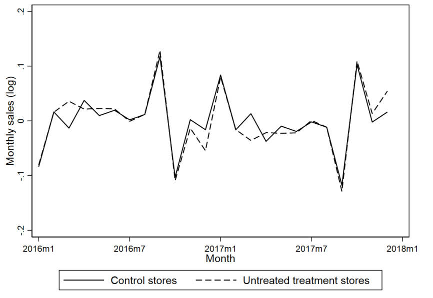


FIGURE E.4. Pre-trend controlled for seasonality

Appendix F: Price regression

In this article, we interpret changes in the weekly sales turnover in the grocery stores following Europris establishments as increases in sales volume. However, it could also be the case that the grocery stores reacted by changing their prices when exposed to a new entry, which would imply that our results reflect both changes in volume and changes in prices. To investigate this issue further, we have used weekly price-quantity data from the grocery chain to directly test whether the entry of a Europris store nearby affects the price level in the grocery stores. More specifically, we calculate a weekly measure of the price level at each grocery store and use this price level as the dependent variable in the two-way fixed effects model described in Equation N.1. We use data from the same categories analysed in Tables 8 and 9: Detergent, Coffee, Candy, Bread, Fresh chicken and Milk.

To obtain a measure of the price level at a given store in a given week, we do the following. First, we calculate for each product j in category c the log-difference, denoted $R_{scj,t}$, between the average price at store s, calculated as total sales amount divided by the total quantity bought, and the median price of the product across all grocery stores in that time week. The average relative price of store s is given by $R_{s,t} = \sum_c \Omega_{c,t} \sum_j \omega_{scj,t} R_{scj,t}$, where $\omega_{scj,t}$ is product j's share of the expenditure in the category in the given week, and $\Omega_{c,t}$ is category c's share of the total expenditure in the week. We calculate $R_{s,t}$ using only products that are sold by all grocery stores in the given week. In Table A5, we report results of the two-way fixed effects model with $R_{s,t}$ as the dependent variable. As can be seen from the table, we find no indication that the price levels at the grocery stores are affected by the establishment of Europris stores.

TABLE F.1. Effect of establishment on prices

	Log weekly sales
Same parking	0.000375
	(0.000467)
250m - 2km	-0.0000340
	(0.000274)
2km - 5km	0.000680
	(0.000703)
5km - 15km	-0.000701
	(0.00121)
More than 15km	-0.000224
	(0.000611)
N	28544
r2	0.151

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Appendix G: Estimating separate regressions for each treatment group

TABLE G.1. Effect of establishment by distance

	Log weekly sales	Log weekly sales
One stop	0.103**	
	(0.0488)	
Two stops		-0.0464**
		(0.0200)
Store FE	✓	✓
Week-year FE	✓	✓
Control variables	✓	✓
N	27236	31682
r2	0.823	0.836

Note: The table reports the coefficient on the variables D_{itb} from the estimation of the model described in (2), but where the effect for each treatment group is estimated in separate regressions, that is, keeping only the observations from this treatment group and the control group. Clustered (store level) standard errors in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Appendix H: Model without control variables

TABLE H.1. Effect of establishment by distance

	Log weekly sales	Log weekly store traffic
Same parking	0.100**	0.0633
	(0.0474)	(0.0398)
250m - 2km	-0.0390	-0.0345
	(0.0270)	(0.0241)
2km-5km	-0.0830***	-0.0850***
	(0.0192)	(0.0241)
5-15km	-0.0520***	-0.0426***
	(0.0170)	(0.0161)
More than 15km	-0.0179	-0.0105
	(0.0129)	(0.0229)
Store FE	✓	✓
Week-year FE	✓	✓
N	34228	34228
r2	0.838	0.842

Clustered (by store) standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Appendix I: Models with alternative distance bins

TABLE I.1. Effect of establishment by distance, four bins

	Log weekly sales	Log weekly store traffic
Same parking	0.105**	0.0657
	(0.0490)	(0.0413)
250m - 2km	-0.0413	-0.0351
	(0.0274)	(0.0244)
2km - 5km	-0.0883***	-0.0873***
	(0.0188)	(0.0238)
More than 5km	-0.0341**	-0.0266
	(0.0172)	(0.0174)
Store FE	✓	✓
Week-year FE	✓	✓
Control variables	✓	✓
N	34204	34204
r2	0.839	0.842

Clustered (by store) standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

TABLE I.2. Effect of establishment by distance, three bins

	Log weekly sales	Log weekly store traffic
Same parking	0.104**	0.0652
	(0.0489)	(0.0413)
250m - 2km	-0.0414	-0.0353
	(0.0274)	(0.0244)
More than 2km	-0.0642***	-0.0603***
	(0.0165)	(0.0183)
Store FE	✓	✓
Week-year FE	✓	✓
Control variables	✓	✓
N	34204	34204
r2	0.838	0.842

Clustered (by store) standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

TABLE I.3. Effect of establishment by distance, six bins

	Log weekly sales	Log average basket
Same parking	0.105**	0.0663
	(0.0489)	(0.0412)
250m - 1km	-0.00761	-0.00878
	(0.0274)	(0.0238)
1km-2km	-0.0793*	-0.0649
	(0.0440)	(0.0405)
2km-5km	-0.0878***	-0.0868***
	(0.0187)	(0.0237)
5km-15km	-0.0389*	-0.0339
	(0.0226)	(0.0217)
More than 15km	-0.0224	-0.0104
	(0.0172)	(0.0249)
Store FE	✓	✓
Week-year FE	✓	✓
Control variables	✓	✓
N	34204	34204
r2	0.839	0.842

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

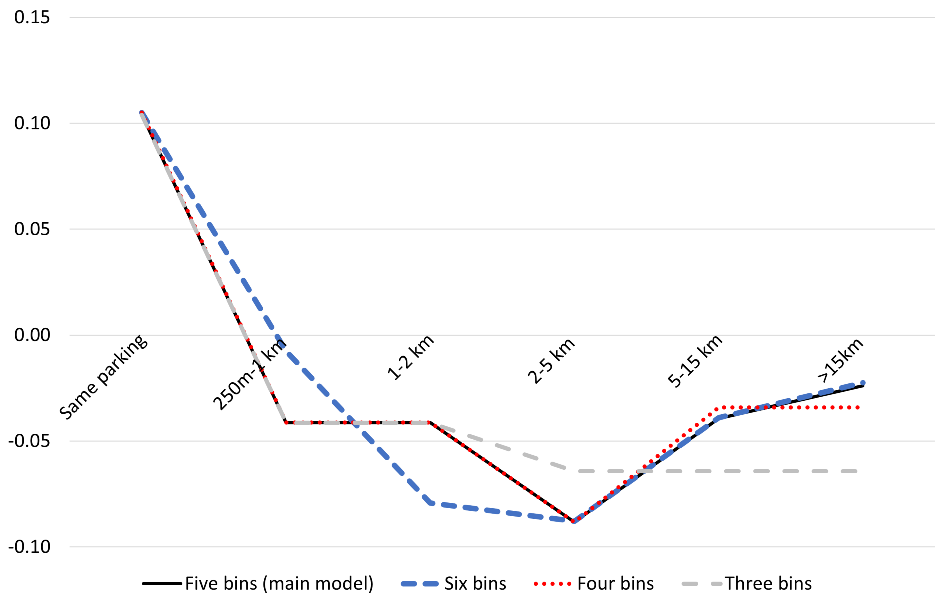


FIGURE I.1. Illustration of models with different bin specifications. Note that in the Figure we have for illustrative purposes imposed the number estimated for the first part of the distance bin also for the remaining parts. For instance, for the three-bin model we impose the number -0.04 estimated for the 250m-2km for both the 0.25-1km bin and the 1-2km bin. Likewise, the number estimated for more than 2km (-0.06) is imposed both for the 5-15 km and the more than 15 km bins.

Appendix J: Model that includes linear trends by treatment group

TABLE J.1. Effect of establishment by distance

	Log weekly sales	Log weekly store traffic
Same parking	0.196***	0.0831
	(0.0555)	(0.0537)
250m - 2km	0.0409	-0.0225
	(0.0463)	(0.0323)
2km-5km	-0.0834**	-0.104***
	(0.0331)	(0.0229)
5km-15km	-0.101	-0.105
	(0.105)	(0.0635)
More than 15km	-0.0133	-0.0592
	(0.0867)	(0.0603)
N	34204	34204
r2	0.844	0.848

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Appendix K: Model excluding stores that experience multiple establishments

TABLE K.1. Effect of establishment by distance

	Log weekly sales	Log weekly store traffic
Same parking	0.107*	0.0769
	(0.0604)	(0.0496)
250m - 2km	-0.0774***	-0.0640**
	(0.0289)	(0.0260)
2km - 5km	-0.102***	-0.0874***
	(0.0188)	(0.0187)
5km - 15km	-0.0580**	-0.0444*
	(0.0250)	(0.0239)
More than 15km	-0.0400**	-0.0225
	(0.0174)	(0.0258)
N	23098	23098
r2	0.848	0.853

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Appendix L: Model including stores with unclear treatment status

TABLE L.1. Effect of establishment by distance

	Log weekly sales	Log weekly store traffic
Same parking	0.0900**	0.0553
	(0.0418)	(0.0351)
250m - 2km	-0.0325	-0.0257
	(0.0240)	(0.0215)
2km-5km	-0.0855***	-0.0738***
	(0.0182)	(0.0229)
5km-15km	-0.0771**	-0.0715*
	(0.0377)	(0.0391)
More than 15km	-0.0231	-0.0102
	(0.0164)	(0.0238)
Store FE	✓	✓
Week-year FE	✓	✓
Control variables	✓	✓
N	36915	36915
r2	0.846	0.849

Clustered (by store) standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Appendix M: Model without control stores that are never treated

TABLE M.1. Effect of establishment by distance

	Log weekly sales	Log weekly store traffic
Same parking	0.0965*	0.0752
	(0.0535)	(0.0449)
250m - 2km	-0.0447	-0.0239
	(0.0273)	(0.0252)
2km-5km	-0.0907***	-0.0748***
	(0.0217)	(0.0258)
5km-15km	-0.0415	-0.0258
	(0.0357)	(0.0305)
More than 15km	-0.0256	0.000785
	(0.0224)	(0.0273)
Store FE	✓	✓
Week-year FE	✓	✓
Control variables	✓	✓
N	9490	9490
r2	0.895	0.873

Clustered (by store) standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Appendix N: Modelling Europris establishment

While the pre-trend and event study plots included in Section 5 do not indicate that stores experiencing a reduction in the distance to the nearest Europris store had different underlying trends in store activity than stores in the control group, some concerns could still persist. First, parallel pre-trends do not preclude that the trends of the treatment and control stores could diverge in the post-treatment period, even in the absence of treatment: the location choice of the Europris stores is presumably based on forward-looking reasoning, and a particular location may be chosen based on a belief that this location will develop positively in the coming years. Second, the pre-trend analyses compare treated stores (i.e. stores experiencing a reduction in the distance to the nearest Europris store) with control stores (i.e. stores never experiencing a reduction in the distance to the nearest Europris store). While these are indeed crucial comparisons, there may still be concern that the treatment *intensity* (e.g., whether a Europris store is established within 250 metres or between 2 and 5 kilometres from a grocery store) depends on characteristics of the potential Europris locations that may be correlated with grocery store demand.

In this subsection, we therefore estimate models where we directly model the location choice of the Europris stores. The aim of the models is to investigate the extent to which the location choice of the Europris stores (and, by extension, the distance between the grocery stores and the nearest Europris store) can be explained by a set of time-varying factors that may also affect demand at the grocery stores.

Both the dependent and independent variables we consider in this section are defined for geographical grids (of different resolutions) covering the municipalities where the grocery stores in our sample are located.

We estimate a set of models where the dependent variable indicates whether Europris is present in a given grid cell or not. The models we estimate are in the following form.

$$Europris^{s}_{gt} = \alpha_g + X'_{gt}\beta + \lambda_t + \varepsilon_{gt} \tag{N.1}$$

The dependent variable $Europris^{s}_{gt}$ is a binary variable that is one if there is at least one Europris store present in a grid cell g of size s at the end of year t. The explanatory variables collected in X_{gt} are the demand factors that could influence both demand at the grocery store and the profitability of a Europris establishment. The explanatory variables we consider are population, the number of buildings, the number of grocery stores, mean income and mean wealth. The variables are measured at the grid cell-year level, giving us, e.g., the number of grocery stores in a given 5 km $\times$ 5 km grid in a given year. We consider the grid sizes 5 km $\times$ 5 km and 1 km $\times$ 1 km. For each grid size, we estimate one model with mean income and mean wealth, and one model without. This is because income and wealth are missing in any grid with zero population, so that including these variables substantially reduces the number of observations. Since we include fixed effects at the grid cell and year level, the explanatory variables will indicate whether it is more likely that a Europris store is present in a given grid cell in years where, e.g., there are more grocery stores or higher mean wealth in the grid cell than in other years. We also estimate models with only the fixed effects in order to investigate how much of the variation in the dependent variable can be attributed to the time-varying grid characteristics. Figure N.1 illustrates the richness of the data and the degree of local (*i.e.*, within-municipality) variation in store locations and demographics.

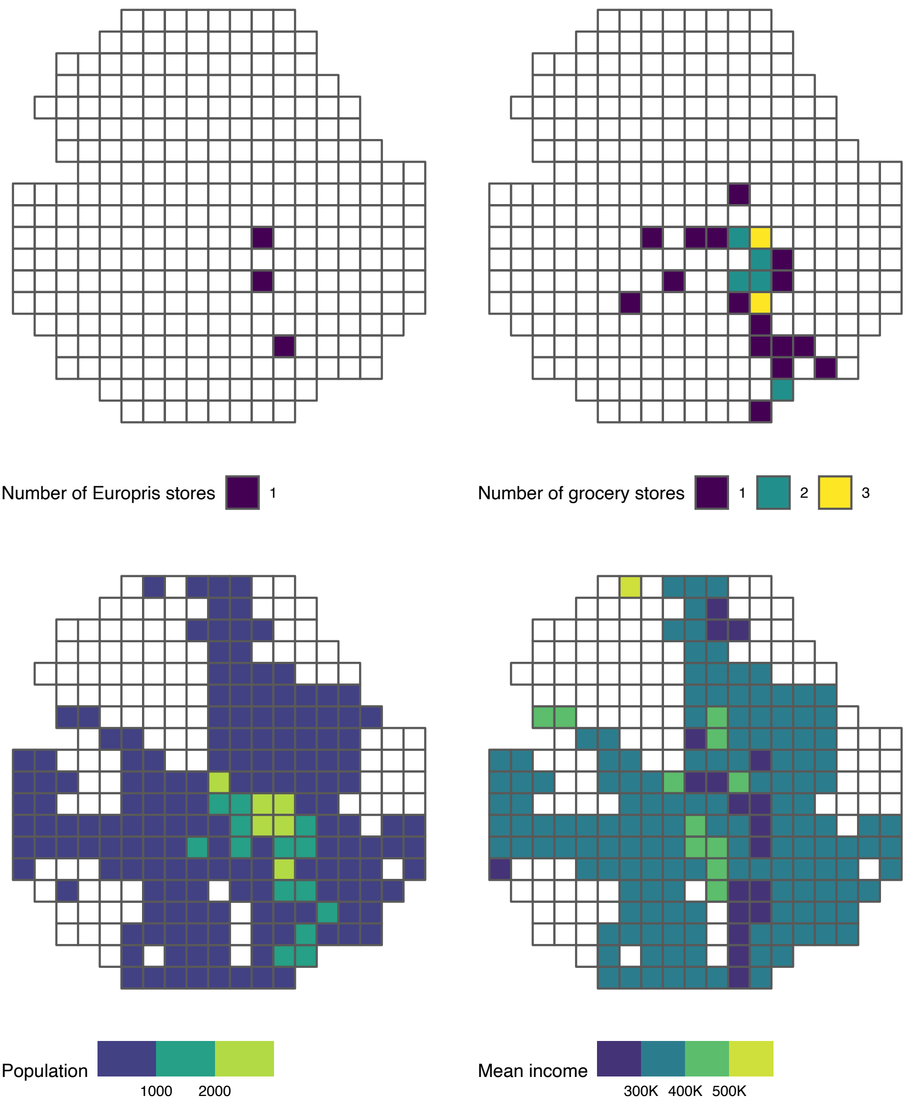


FIGURE N.1. Illustrative example of statistical grids. The figures show central parts of Skien municipality, broken down into 1 km × 1km grids. The four panels provide information about different grid-level measures in 2018. The top-left panel shows the number of Europris stores (at the end of the year). The top-right panel shows the number of grocery stores (at the end of the year) in each grid cell. The bottom-left panel shows the population in each grid cell, while the bottom-right panel shows the mean income. For the bottom-right panel, white grid cells indicate that mean income is missing (either because there are no persons living in the grid cell, or because there are so few that the mean income is not reported due to privacy concerns). For the other panels, white grid cells indicate zero values.

The results are reported in Table N.1. We find no indication that the location choice of Europris stores co-varies with the time-varying demand factors at the grid cell level. In all models, the R^2s with only the fixed effects and with fixed effects and time-varying grid cell characteristics variables are nearly identical, indicating that the time-varying characteristics do not explain any substantial variation in the presence of Europris at the grid cell level. Furthermore, the explanatory variables are neither jointly nor individually statistically significant.

TABLE N.1. Effect of control variables on the presence of Europris

	(1)	(2)	(3)	(4)
Population in 1km×1km grid	0.029	0.029		
	(0.043)	(0.043)		
Buildings in 1km×1km grid	-0.000	-0.000		
	(0.000)	(0.000)		
Grocery stores in 1km×1km grid	0.036	0.036		
	(0.028)	(0.028)		
Mean wealth in 1km×1km grid		-0.001		
		(0.001)		
Mean income in 1km×1km grid		0.003		
		(0.004)		
Population in 5km×5km grid			-0.077	-0.077
			(0.048)	(0.049)
Buildings in 5km×5km grid			0.000	0.000
			(0.000)	(0.000)
Grocery stores in 5km×5km grid			0.019	0.019
			(0.022)	(0.022)
Mean wealth in 5km×5km grid				-0.007
				(0.005)
Mean income in 5km×5km grid				-0.000
				(0.015)
Observations	219728	39405	8948	4412
R2	0.876	0.875	0.926	0.924
R2, fixed effects only	0.876	0.875	0.925	0.923
p value, joint test	0.461	0.472	0.299	0.329

Note: This table reports the results from the estimation of Equation N.1. In columns (1) and (2), all variables are measured on the 1 km × 1 km-grid-year level. In columns (3) and (4), all variables are measured at the 5 km × 5 km-grid-year level. In all columns, the dependent variable is one if there is at least one Europris store located in the grid, and zero otherwise.

In Table N.1, we consider both 1 km × 1 km and 5 km × 5 km. As a robustness check, we have also estimated a model where the dependent variable is the presence of a Europris in 1 km × 1 km grid, but where we use a distance weighted average of the independent variables from all 1 km × 1 km within 5 km of the grid in questions. This specification captures the idea that location choices are based on a wider area than a 1 km × 1 km grid but that closer locations are more important than more distant locations. The results are fully in line with the results reported in Table N.1: The independent variables have non discernible effect of the presence of Europris stores.

TABLE N.2. Effect of control variables on the presence of Europris

	(1)	(2)
Population	0.096	0.096
	(0.099)	(0.099)
Number of buildings	0.000	0.000
	(0.000)	(0.000)
Number of grocery stores	0.140	0.140
	(0.103)	(0.103)
Mean wealth		-0.000
		(0.000)
Mean income		0.000
		(0.001)
Observations	218780	137461
R2	0.874	0.874
R2, fixed effects only	0.874	0.874
p value, joint test	0.165	0.178

Note: This table reports the results from the estimation of Equation N.1. In both columns, the dependent variable is one if there is at least one Europris store located in a 1 km $\times$ 1 km-grid cell, and zero otherwise. The dependent variables are weighted averages of the values on all 1 km $\times$ 1 km grids within 5 km of the grid in question, with weights inversely related to the distance.

Another way of investigating whether Europris store locations co-vary with local demand and competition factors is to replace our left-hand-side variable in our main model with the control variables used in this model. Hence, we estimate seven versions of the model described in Equation N.1 using our control variables as left-hand-side variables. While there are some statistically significant coefficients in the municipality-level variables, we find no systematic patterns across these models that are correlated with the difference between Europris establishments close to the grocery stores and establishments further away. The results are presented in Tables N.3 and N.4.

TABLE N.3. Control variable regressions part 1

	Population density	Higher education	Median income after tax
Same parking	16.25**	-0.00180**	3733.7***
	(6.297)	(0.000786)	(970.0)
250m - 2km	-6.816	-0.000427	-229.2
	(6.403)	(0.000647)	(884.1)
2km - 5km	11.61	-0.00186	-2004.3
	(9.434)	(0.00153)	(1782.4)
5km - 15km	2.520	0.0000410	8101.0***
	(6.242)	(0.000945)	(843.2)
More than 15km	32.26**	-0.00601***	-1937.2*
	(13.02)	(0.00215)	(1137.6)
Store FE	✓	✓	✓
Week-year FE	✓	✓	✓
Control variables	✓	✓	✓
N	34204	34204	34204
r2	0.999	1.000	0.997

Clustered (by store) standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

TABLE N.4. Control variable regressions part 2

	Population by store	HHI umbrella	HHI chain	HHI store
Same parking	13.86	0.0113***	-0.00452	-0.00103
	(31.18)	(0.00431)	(0.00420)	(0.00676)
250m - 2km	-24.70*	0.000781	-0.00419	0.00527
	(14.61)	(0.00738)	(0.00351)	(0.00395)
2km - 5km	0.826	0.0131	0.00362	-0.00969
	(18.53)	(0.00880)	(0.00461)	(0.00789)
5km - 15km	-145.0*	0.00868	-0.0568	0.0621
	(78.57)	(0.00763)	(0.0422)	(0.0463)
More than 15km	21.05	-0.00531	-0.00111	-0.000279
	(29.25)	(0.00530)	(0.00485)	(0.00629)
Store FE	✓	✓	✓	✓
Week-year FE	✓	✓	✓	✓
Control variables	✓	✓	✓	✓
N	34204	34204	34204	34204
r2	0.984	0.987	0.993	0.994

Clustered (by store) standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Appendix O: A simple theory model on the interplay of intensive and the extensive margins and co-location

In the empirical analysis, we find that whether the grocery store ends up being better or worse off after Europris's entry depends on the distance between the two stores. We also find a non-linear "S-shaped" pattern. The effect ultimately depends on the distance between the two stores. If Europris ends up sufficiently close, the grocery store tends to benefit. In contrast, an establishment that does not bring Europris close enough appears to be harmful. We attribute these findings to the interplay of the extensive margin (increased localised demand) and the intensive margin (fiercer competition and reduced purchases by existing customers). In this section, we develop a simple theoretical example that shows how decomposing the effect into an extensive and an intensive margin provides an intuitive explanation of the results.

Suppose that the market is represented by a line that starts at 0 and ends at an indefinite point. the grocery store is located at $x_G = 0$. It sells n products at a common price p. The customers are uniformly distributed at discrete intervals along the line. They value store proximity, and face travel costs (t) that increase with the distance to the grocery store. Hence, the utility a customer located at x obtains from shopping at the grocery store is given by

$$u_G = nv - tx - np$$

Where v is the customer's gross willingness to pay per product. Note that the customers only shop at the grocery store if the utility exceeds their reservation utility u_R [25].

Pre-Europris establishment

Consider first a market without a Europris store located close enough to affect the grocery store's demand. The consumer who is indifferent about shopping or not shopping at the grocery store is located at

$$\hat{x} = \frac{nv - np - u_R}{t}.$$

The location of the grocery store and the indifferent consumer is illustrated in Figure O.1.

25. The reservation utility reflects the attractiveness of the customers' outside options, such as rival grocery stores

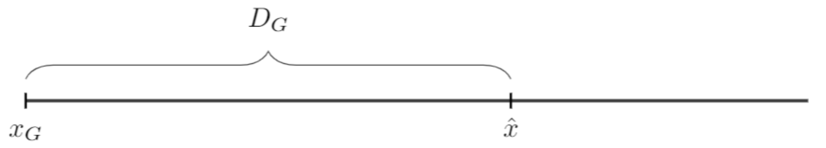


FIGURE O.1. Pre-establishment

The figure also shows that the grocery store's demand before a Europris establishment is given by

$$D_G = \hat{x} = \frac{nv - np - u_R}{t}.$$

Post-Europris establishment

Suppose then that Europris establishes a store at $x_E \in [0, \hat{x}]$. Europris offers one of the products sold by the grocery store, but at a lower price αp, where $\alpha \in (0,1)$. The utility of just shopping at the grocery store is unchanged, but the customers might obtain an additional value by purchasing the cheaper product from Europris. Visiting both stores provides a utility equal to

$$u_{E,G} = nv - (n-1)p - \alpha p - tx - t(x_E - x) - F$$

for customers located at $x \in (0, x_E)$, and

$$u_{E,G} = nv - (n-1)p - \alpha p - tx - F$$

for customers located at $x > x_E$ [26]. The parameter F denotes the additional cost that customers face if the stores cannot be visited in one stop, i.e., unless $x_E \leq 250$m. We find that the location of the consumer who is indifferent about just shopping at the grocery store or shopping at both the grocery store and Europris is given by

$$\tilde{x} = x_E - \frac{p(1-\alpha) - F}{t}.$$

The shorter the distance between the grocery store and Europris, the more customers prefer shopping in both stores. The customer who is indifferent about shopping at both stores or none of them is located at

$$\hat{x}' = \frac{nv - p(n-1) - \alpha p - u_R - F}{t}.$$

Consequently, customers who only shop at the grocery store are located to the left of $\tilde{x}$, while customers who shop at both stores are located between $\hat{x}'$ and $\tilde{x}$. Figure O.2 outlines the grocery store's exclusive demand (D_G) and shared demand ($D_{E,G}$).

26. These customers pass Europris on their way to the grocery store and no extra travel costs are incurred. We assume that the customers do not care where on the way Europris is located, only about whether they have to stop once or twice.

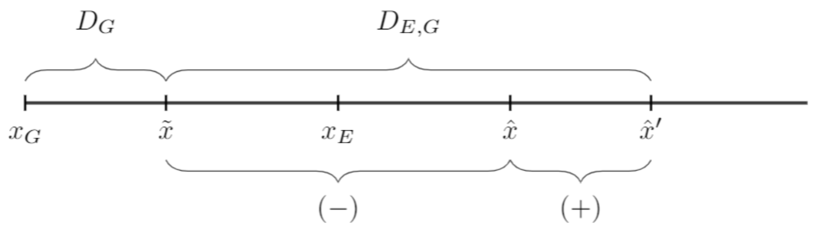


FIGURE O.2. Post-establishment

The extensive margin

For customers to the right of x_E, the presence of Europris increases the utility of travelling to the left on the line. As a result, some of the customers who previously did not shop at the grocery store change their mind now that they can visit Europris during the same trip. This effect is what we refer to as *the extensive margin* in response to a Europris establishment. Graphically, the extensive margin is captured by $\hat{x}'$ being located further to the right than $\hat{x}$. New grocery store customers are given by

$$\hat{x}' - \hat{x} = \frac{1}{t}(p(1-\alpha) - F)$$

Since the new customers purchase $(n-1)$ products from the grocery store and 1 product from Europris, the increase in the grocery store's revenue equals

$$(\hat{x}' - \hat{x})p(n-1)$$

The intensive margin

After the Europris establishment, some of the customers who previously purchased all n products from the grocery store decide to purchase the discounted product from Europris. This response to the increased competition is called *the intensive margin*. For the grocery store, this effect is always negative as it implies lower demand. A comparison of Figure O.1 and Figure O.2 shows how the customers located between $\hat{x}$ and $\tilde{x}$ went from being exclusive grocery store customers to becoming shared customers in the wake of the establishment. Formally, we see that

$$\hat{x} - \tilde{x} = \frac{nv - u_R - p(n-1) - \alpha p - F}{t} - x_E$$

customers purchase less at the grocery store. This corresponds to a revenue loss equal to

$$(\hat{x} - \tilde{x})p$$

The total effect is simply the sum of the gained revenues due to the extensive margin and the lost revenues due to the intensive margin.

Numerical illustration of the co-location forces

Figure O.3 shows the effect of a Europris establishment on the grocery store's revenues. It shows the effects from the extensive margin, the intensive margin and the total. The parameter values are set to $v = 1$, $t = 2$, $\alpha = 0.5$, $n = 10$, $p = 0.75$, $u_R = 0,1$ and $F = 0.33$.

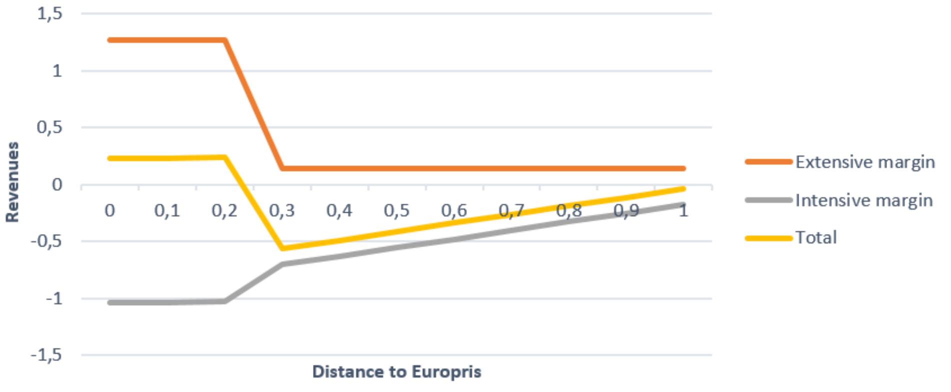


FIGURE O.3. Intensive vs extensive margin

Note that the effect of the extensive margin dominates when the distance between Europris and the grocery store is short. There are two main reasons for this. First, the customers do not have to make an additional stop to visit Europris, which attracts more customers. Second, the gain from attracting a new customer is greater than the loss from an exclusive customer turning into a shared customer. Recall that new customers purchase $(n-1)$ products, while shared customers only purchase one product fewer than before the Europris establishment. However, as the distance between the grocery store and Europris increases, the effect of the intensive margin becomes dominant. When shopping at both stores requires two stops, a Europris establishment might not attract sufficiently many customers for the grocery store to benefit from it. Eventually, the competition effect also fades away and the total effect approaches zero.

While the predictions from our modelling framework will be sensitive to the parameters chosen, we do find in Figure O.3 a very similar pattern to the "S-shape" observed in our empirical analysis, as illustrated in, e.g., Figure 3.