# Analysis of Entry and Market Structure in Swedish Retail Food Stores

David Hakula (23119), Christopher Hayes (41424), Filip Mellgren (23644)

Course 5321: Industrial Organization

2019-04-24

## Introduction

This report analyzes the link between competition and profits by investigating market structure and entry opportunities in the Swedish retail food market. We use cross-sectional municipality-level data from 2008, collected by DELFI Marknadspartner AB. The sample comprises 236 out of Sweden's total 290 municipalities.

## 1. Premises of the analysis

Our analysis focuses on big-box stores only, rather than all types of retail food stores. This approach has both pros and cons. Insofar as these smaller retailers are not perfect competitors to the big-boxes, excluding them has the benefit of stripping out extraneous noise. And the more homogenous our set of stores, the stronger the assumption of symmetry that underpins our probit model later on. Conversely, insofar as smaller stores do indeed compete with big-boxes, we potentially under-estimate the level of competition in the market and bias our results.

Data limitations preclude rigorous evaluation of the trade-off in our exact context, e.g. by price analysis, estimation of substitution effects or a hypothetical monopolist test like the SSNIP-test. However, based on intuitive qualitative evidence with regard to product offerings, and on studies of analogous settings, judge that the separate market assumption is not strictly plausible. Nonetheless, to assume otherwise is prohibitively restrictive analytically.

Firstly, natural experiments in the US show substitution between big-boxes and smaller retailers both on the demand side and in the labour market. Haltiwanger, Jarmin & Krizan (2009) find big-box entry reduces employment growth in both single unit and small chain retailers, but only within the immediate area and in the same detailed industry.[1] Similarly, Dube, Lester & Eidlin (2007) found lower average and aggregate retail worker earnings associated with Walmart's entry into a county. We find it likely that the Swedish market experiences similar characteristics.

Secondly, there is intuitively a clear overlap between the services offered by big-boxes and by smaller retailers, in terms of characteristics and intended use, suggesting plausible substitutability. But this overlap is inexhaustive, and itself comprises underlap between the different constituent small retailers (A hypermarket's product offering overlaps with both butchers and bakers, who nonetheless do not overlap with each other). Furthermore, big-boxes have distinct characteristics they don't share with smaller stores, offering convenience and scale economies. These complications suggest that the separate market assumption is analytically defensible.

[1] *https://www2.census.gov/ces/wp/2009/CES-WP-09-34.pdf*

## 2. Distribution of big-box stores

Table 1 shows that the distribution of the number of big-box stores across local markets is bimodal, with peaks at 2 and 5. More than half of municipalities in our data have 2 or fewer big box retailers, while 9% have none at all. 22% have the maximum observed value of 5. Observing 0 or 4 stores is rare, suggesting it is profitable to enter where there are no incumbents and that few want to be the fourth entrant, or that being the fifth entrant is beneficial.

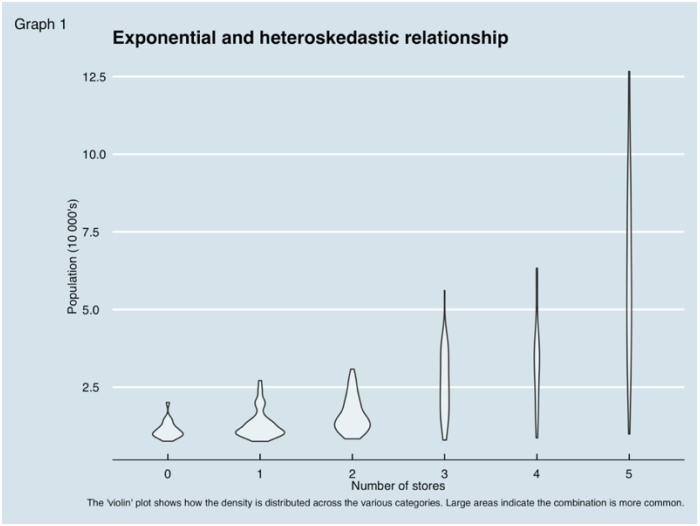
Table 1: Distribution of the stores variable

| Stores | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| Count | 21 | 46 | 57 | 38 | 22 | 51 |
| Frequency | 0.090 | 0.200 | 0.240 | 0.160 | 0.090 | 0.220 |
| Cumulative | 0.090 | 0.290 | 0.530 | 0.690 | 0.780 | 1 |

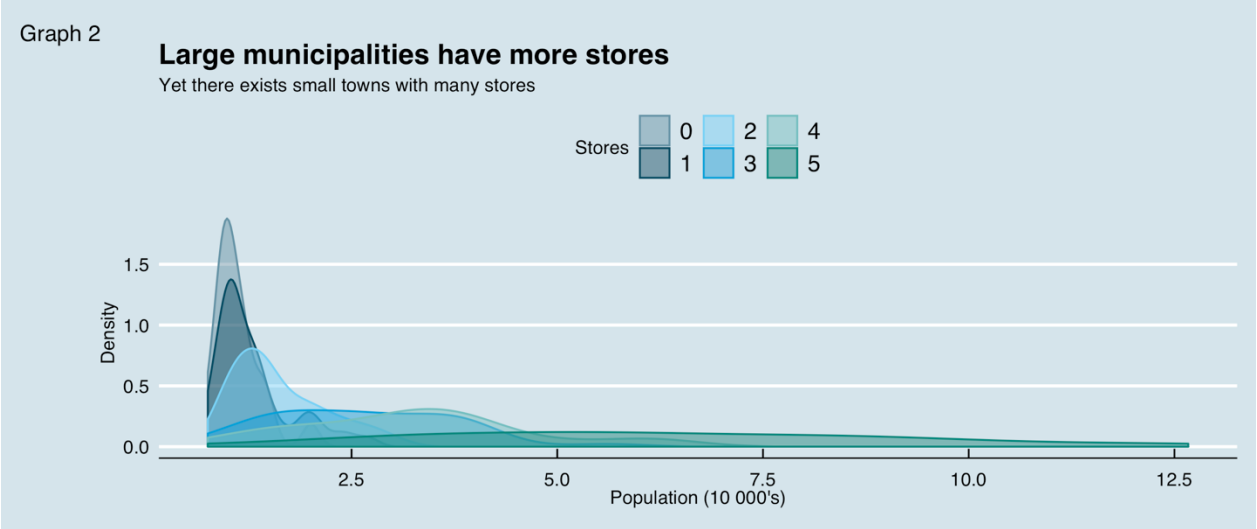
## 3. The relationship between population and number of stores

Graph 1, and particularly Graph 2, illustrate how smaller municipalities can sustain even up to five stores due to local characteristics. They also show that areas with large populations always need to be served by many big box stores. Seeing that local characteristics do play an important role in determining what number of big-box stores can be sustained, the remainder of this report tries to distinguish what these factors may be.

Graph 1




Graph 2



## 4. Descriptive summary statistics

Table 2 presents summary statistics as well as correlations between the number of stores and local market characteristics. Note that we include the variable consumption per capita, as the consumption variable appears to be an aggregate measure rather than per capita. We also include the following created measures: population (in number) living within two kilometers of a store (close_pop) and a measure of consumption as share of wage (to capture the saving behavior).

Table 2: Descriptive statistics

| | Variable | Min. | Mean | Max. | Std. Dev | Store Correlation |
|---|---|---|---|---|---|---|
| 1 | Stores | 0.00 | 2.63 | 5.00 | 1.63 | 1.00 |
| 2 | Population | 0.75 | 2.84 | 12.67 | 2.46 | 0.72 |
| 3 | Pop. growth | -0.09 | 0.02 | 0.23 | 0.05 | 0.28 |
| 4 | Pop. growth > 0 | 0.00 | 0.03 | 0.23 | 0.04 | 0.21 |
| 5 | Pop. growth <0 | -0.09 | -0.01 | 0.00 | 0.02 | 0.30 |
| 6 | Fraction kids | 0.08 | 0.11 | 0.16 | 0.02 | 0.12 |
| 7 | Fraction pensioner | 0.12 | 0.20 | 0.27 | 0.03 | -0.24 |
| 8 | Fraction young | 0.06 | 0.10 | 0.22 | 0.02 | 0.54 |
| 9 | Fraction women | 0.48 | 0.50 | 0.53 | 0.01 | 0.51 |
| 10 | Consumption | 1.68 | 6.76 | 30.02 | 5.91 | 0.72 |
| 11 | House prices | 0.25 | 1.21 | 4.71 | 0.76 | 0.30 |
| 12 | Houses sold | 0.13 | 2.20 | 9.68 | 1.61 | 0.68 |
| 13 | Wage | 2.17 | 2.30 | 2.59 | 0.07 | 0.32 |
| 14 | Fraction close to store | 0.45 | 0.75 | 1.00 | 0.11 | 0.38 |
| 15 | Pop clsoe to store | 0.40 | 2.26 | 11.73 | 2.20 | 0.71 |
| 16 | Consumption/capita | 2.15 | 2.35 | 3.55 | 0.13 | 0.18 |
| 17 | Consumption wage share | 0.92 | 1.02 | 1.47 | 0.05 | 0.02 |

We note that the average number of stores 2.63 and that the maximum is 5. The latter implies either: that no Swedish municipality has more than five big-box stores; that larger number have been coded as five in the dataset; or that all municipalities with six or more stores belong to the 54 excluded from the dataset.

Another interesting note is that there is be large variation in consumption between municipalities: The mean value is 6.76 (in 10,000s of SEK), while the minimum and maximum value is 1.68 and 30.02 respectively. The Consumption/capita measure has a lower variance, however.

The variables that correlates strongest with number of stores is population (0.72) (and overall consumption which captures the same information). Hence, the population size is by itself able to account for 52% of how many stores there are in a municipality. Other variables that correlate strongly with number of stores are the share of population living close the stores (0.71), which is determined by the number of stores and how dispersed they are; the number of houses sold (0.68), reflecting the population size; and the share of young people (0.54), which is mostly driven by population since young people tend to live in larger cities.

The variables that have the least correlations with number of stores are share of pensioners (-0.24), reflecting that pensioners tend to live in more rural communities but may also reflect their consumption patterns; consumption wage share (0.02); and share of kids (0.12). Hence the municipalities with many pensioners appears to have a few number of stores, while the number of stores does not vary with consumption per capita and only correlates weakly with the share of kids in the municipality.

## 5. Defining variable categories included in the analysis

Our preferred specification is obtained through following the type of reasoning used by Bresnahan & Reiss (1991). The model constructs a profit formula by relating the available variables into three separate models – market size, variable profits, and fixed costs. Variable profits and fixed costs are both parametrized such that margins are falling in the number of firms. In general, we choose to

include rather than exclude variables since the descriptive analysis above shows that population needs to be supplemented.

To determine the market size, S, we include the following variables in the population variables vector Y: population (pop), positive population growth (pos_gpop) and negative population growth (neg_gpop). We choose to include both positive and negative growth, as there may be a non-linearity present as hinted in Table 2, but hard to discern graphically. We leave out gpop to avoid perfect collinearity.

To determine the variable profits, V, we include the following variables in the local market characteristics vector X: share of kids (s_kids), share of pensioners (s_pens), share of young people (s_young), share of women (s_women) and consumption per capita (cons_pc). We choose to include these as the variable profits are likely to depend on the respective groups' consumption patterns – some groups might tend to buy higher margin products. We also include the consumption per capita, the idea being that margins can be affected by the willingness of consumers to spend. We do not include wages as a variable cost; instead, it enters as a fixed cost and we still want to be somewhat parsimonious in our specification since the data only has 235 complete cases.

To determine the fixed costs, F, we include the following variables in the (fixed) cost shifters vector W: average price per square-meter of houses sold (hprice), and the average monthly wage (wage). We include the average price per square-meter of houses sold as it functions as a proxy that captures the cost of buying or renting a facility. The average monthly wage captures the fixed costs associated with staffing the stores.

Note that we choose to not use the share of population with a store within 2 kilometers (dist), as it might classify as a bad control: it is likely to be determined by the number of stores.

## 6. Estimation of model parameters

Table 3 shows the full specification and the estimates that maximizes the likelihood of observing the number of stores observed given the data and the parameter estimates. First, all estimates are not to be interpreted as marginal effects, as they are sometimes interacted with other variables. However, the estimates do have an economic meaning in the sense that they enter into the profit equation provided in the task description.

Table 3: ML estimates

| | Variable | Coef. | s.e. | Type |
|---|---|---|---|---|
| 1 | a1 | -3.62 | 2.93 | Marginal variable profit penalty |
| 2 | a2 | 0.04 | 0.36 | Marginal variable profit penalty |
| 3 | a3 | 0.09 | 0.23 | Marginal variable profit penalty |
| 4 | a4 | 0.42*** | 0.14 | Marginal variable profit penalty |
| 5 | a5 | 0.1 | 0.09 | Marginal variable profit penalty |
| 6 | g1 | 8.93*** | 3.83 | Marginal fixed cost premium |
| 7 | g2 | 1.08*** | 0.44 | Marginal fixed cost premium |
| 8 | g3 | 1.05*** | 0.36 | Marginal fixed cost premium |
| 9 | g4 | 0.07 | 0.27 | Marginal fixed cost premium |
| 10 | g5 | 0.36 | 0.28 | Marginal fixed cost premium |
| 11 | House prices | 0.37* | 0.23 | Fixed cost |
| 12 | Wages | -3.81 | 1.69 | Fixed cost |
| 13 | Fraction kids | -6.86 | 6.26 | Variable profit |
| 14 | Fraction young | -3.19 | 2.44 | Variable profit |
| 15 | Fraction pensioner | 0.21 | 3.34 | Variable profit |
| 16 | Fraction Women | 12.52** | 6.50 | Variable profit |
| 17 | Consumption/capita | -0.09 | 0.34 | Variable profit |
| 18 | Positive growth | -1.62 | 2.30 | Market size component |
| 19 | Negative growth | 3.03 | 3.22 | Market size component |

'***' indicates $p$ <1%, '**' indicates $p$ <5%, '*' indicates $p$ < 10% level.

We note that only six variables are significant at any level, which is disappointing but reflects the limited power in our analysis. All of our variable profit penalty and fixed cost premium parameters are of the expected sign; firm entry both tightens variable profits and inflates fixed costs through competition for property and workers.

Apparently, more women is significantly related to more big box stores which indicates that the variable profit on female consumers is higher relative to men. It is counterintuitive that positive

population growth, i.e. a higher future market size, is negatively related to the number of stores. However, that relationship is insignificant and, again, points to the limits of the data at hands.

## 7. Entry thresholds: Population requirements

The first row of Table 4 shows the market size required to sustain the number of businesses as indicated by the respective column (note that market size is population adjusted for positive and negative population growth and should not be confused with raw population.) These values were calculated using the full set of coefficients from the maximum likelihood specification, including insignificant estimates despite the fact that these may mostly contain noise following Bresnahan and Reiss (1991).

The required market size per firm needs to increase to sustain profitability. This reflects the mechanism that margins decrease with increased competition and need to be offset by volumes to maintain profits. This relationship is observed in Table 4.

## 8. Per firm entry threshold

Table 4:

| Number of incumbents: | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Entry threshold | 0.43 | 1.24 | 2.14 | 3.25 | 4.16 |
| per firm entry threshold | 0.43 | 0.62 | 0.71 | 0.81 | 0.83 |
| Entry threshold ratios | 1.42 | 1.15 | 1.14 | 1.02 | 1.00 |
| Marginal absolute threshold | 0.43 | 0.80 | 0.90 | 1.11 | 0.91 |

The per-firm entry thresholds (showed in the second row of Table 4) are increasing in the number of firms at a diminishing rate, i.e. competition intensity is increasing and concave in the number of firms. As the threshold ratios (showed in the third row of Table 4) show, the per-firm duopoly threshold is 42% greater than the monopoly threshold, while the per-firm triopoly threshold is only 15% greater than the duopoly threshold. Competition intensity increases by the same amount with the fourth entrant but is more or less flat thereafter, with a per-firm threshold ratio barely above unity at 1.02, indicating near perfect competition. This suggests that collusion is not tenable between any more than four firms. A corollary is that perfect competition has been approximated on both demand- and supply-side: entry no longer compresses variable profits, nor inflates fixed costs.

Moreover, while the threshold ratios flatten, the absolute marginal entry threshold increases fairly reliably until it peaks at the 4th firm (11.12k), falling to 9.06k for the 5th firm. The absolute entry threshold is therefore not convex beyond 4 firms.

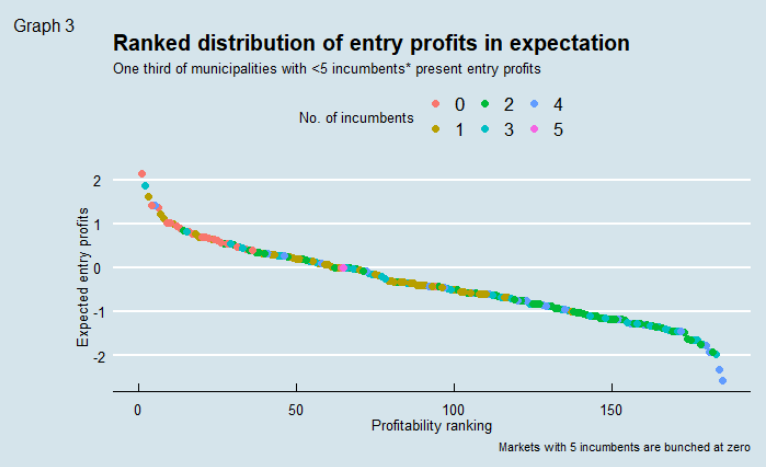
This analysis boils down to the very intuitive insight that a market is more preferable (in terms of lower competition and hence higher potential profit, from a firm's point of view) the fewer firms that are competing in it.

## 9. Recommendations

We recommend that further analysis is undertaken whenever the estimated would-be-profits under $n + 1$ firms is positive, as this means that an entering firm would be profitable. To obtain this information, we simply applied our estimated profit equation to each municipality with the data observed in that municipality, having the number of stores increased by one. That way, we obtain markets for which we can't explain why there is not an additional firm present with the current specification.

We find that 64 markets present positive entry profits in expectation, out of 235 in our dataset, and out of 184 with 4 incumbents or fewer. (Markets with 5 incumbents have projected profits set to zero under this specification.) Figure 3 ranks the markets by expected entry profits and plots their distribution, colour-coding by the number of incumbents. We observe that the vast majority of the most profitable entry opportunities are in markets that are currently completely unoccupied, as well as a smattering of monopoly and duopoly markets. Even in most monopolised markets, the prospects of entering as a duopolist are mildly negative. The majority of markets with more incumbents are very unattractive entry prospects.

Graph 3




## 10. Limitations

The above performed analysis relies on several assumptions of the characteristics of the big-box stores market. First, the firms in the market share identical profit functions as we have specified it, save a normally distributed error term. 'Identical' may stretch reality, but the firms will likely face at least similar profit functions, so this assumption is hardly heroic. However, as argued above, independence of the error term may be problematic, since it is likely endogenously determined on a municipal level.

Second, there should be no product differentiation. As big-box stores are likely to offer similar ranges of products, we consider this assumption to stretch reality, but not to be too strong.

Finally, we assume no spillovers in the sense that all shoppers are locals and that all locals shop locally. This is a very strong assumption that could be mitigated by controlling for neighboring populations, tourism, transport links and outbound shoppers (unfortunately, we lack data on these aspects for this analysis).

Note that the analysis also relies on the assumption that big-box stores make up a separate market, as more thoroughly discussed under topic *1. Premises of the analysis*. Another weakness is the limited data set which makes the estimates subject to random noise as evident in the lack of significance in Table 3.

Following from the points made above, it is clear that the analysis is limited. There are other aspects of the 'real' market situation that should be considered before making an entry decision. For example, the used approach does not take the risk of future entries of an (from this model's point of view) irrational actor into account.

## Concluding words

In this report, we have analyzed the link between competition and profits by investigating market structure and entry opportunities in the Swedish retail food market. The key take-away is that 64 markets (municipalities) offer positive entry profits due to under-utilized market characteristics.