Empirical Workshop 2:
Entry and market structure

Stockholm School of Economics

Matilda Orth
Research Institute of Industrial Economics IFN, Stockholm

April 16, 2019

You work as an analyst for a global retail firm. The firm plans to enter into the Swedish retail food market and is about to make decisions over entry into several local markets. The firm thus needs to analyze the link between competition and profits. You are responsible for this investigation. The firm is particularly interested in the market for large ("big-box") retailers and wants you to perform a detailed investigation of competition and profits of hypermarkets. The starting point for your analysis is the paper by Bresnahan and Reiss [BR] (1991).

To do the analysis, you have got cross-sectional data on Swedish retail food stores in 2008. The unit of observation is a municipality. The sample consists of 236 (out of 290) municipalities. The data set is collected by DELFI Marknadspartner AB.

You access the following variables: *stores* is the number of big-box stores (hypermarkets); *pop* is total population in 10,000s; *gpop* is the overall growth in population from 2001 to 2008; *pos_gpop* is positive population growth from 2001 to 2008; *neg_gpop* is negative population growth from 2001 to 2008; *s_kids* is the share of children (below 6yrs old); *s_pens* is the share of pensioners (above 65yrs old); *s_young* is the share of young people (20-29yrs old); *s_women* is the share of women; *consumption* is the per-capita purchasing power in 10,000s of SEK; *hprice* is the average price per square-meter of houses sold in 10,000s of SEK; *hsold* is the number of houses sold in 100s; *wage* is the average monthly wage for municipality employees in 10,000s of SEK; and *dist* is the share of population with a store within 2 kilometers.

The data can be downloaded from the course web page. The data set is named *Hypermarkets.csv* (in text format) and *Hypermarkets.dta* (in STATA format). You can use any software package you want to solve the exercise.

Bresnahan and Reiss (1991) propose an empirical framework to measure the effects of entry in concentrated markets. They assume symmetry among firms in each market. The profit for a firm in market m is:

$$\begin{aligned} \pi_{N_m} &= S(Y, \lambda) V_{N_m}(X, \alpha, \beta) - F_{N_m}(W, \gamma) + \varepsilon_m \\ &= \tilde{\pi}_{N_m} + \varepsilon_m, \end{aligned} \tag{1}$$

where N_m is the number of firms in market m; S is market size; V is variable profits; F is fixed costs; Y contain population variables; X is local market characteristics; W is cost shifters; ε_m is an error term assumed to be i.i.d. across local markets and distributed $N(0,1)$; and $\theta = (\lambda, \beta, \alpha, \gamma)$ are parameters to be estimated.

Market size is assumed to be a linear function of population variables, where the coefficient on population equals 1. The variable profits per-capita is assumed to have a linear form according to

$$V_{N_m} = \alpha_1 + \mathbf{X}\boldsymbol{\beta} - \Sigma_{n=2}^{N_m} \alpha_n, \tag{2}$$

where $\alpha_n > 0$ allows competition to increase in number of firms. Finally, fixed costs take the following form

$$F_{N_m} = \gamma_1 + \gamma_L W_L + \Sigma_{n=2}^{N_m} \gamma_n, \tag{3}$$

where $\gamma_n > 0$ allows later entrants to have higher fixed costs due to, e.g., higher barriers to entry. A detailed presentation of the model is given in Appendix and in the Lecture notes.

1. Explain to the retail firm what are the pros and cons of using only big-box stores rather than all stores in your analysis. Is it plausible to assume that there is a separate market only for big-box stores? Why or why not?

2. Show the distribution of the number of big-box stores across local markets (BR Table 2). Comment.

3. Graphically illustrate the relationship between population and the number of stores in local markets. Moreover, show a histogram of local market population across markets with different number of hypermarkets.

4. Present summary statistics and correlations of the number of stores and local market characteristics. Interpret your findings.

5. Based on the data set at hand, what variables would you include in each part of the profit function in equation (1)? Justify your answer.

6. Specify and estimate the entry model by BR (1991) using the available data. The do-file “BR_Estim.do” is available on the course web page. Discuss your results (BR Table 4).

7. Calculate entry thresholds based on your estimated coefficients and the average values of the exogenous variables. The entry thresholds are given by

$$S_N = \frac{\hat{\gamma}_1 + \hat{\gamma}_L \overline{W}_L + \Sigma_{n=2}^{N} \hat{\gamma}_n}{\hat{\alpha}_1 + \overline{\mathbf{X}}\hat{\boldsymbol{\beta}} - \Sigma_{n=2}^{N} \hat{\alpha}_n} \tag{4}$$

   where the bars indicate average values. How many individuals are required for a monopolist to operate? How many individuals are required for two, three, four and five stores, respectively, to operate? Explain. (BR Table 5, left panel A)

8. Construct per-firm entry threshold ratios (BR Table 5, right panel A). How does the intensity of competition change when there are two big-box stores compared to one? How does the intensity of competition change when there are three big-box stores compared to two, four big-box stores compared to three, and five big-box stores compared to four? Interpret and discuss your results.

9. What types of local markets would you recommend the international retail firm to enter? Why?

10. The performed analysis relies on several assumptions. Explain the limitations of your analysis to the retail firm.

The report needs to be **no more than 6 pages**. The report must be e-mailed to **matilda.orth@hhs.se** by 23:59, on Wednesday **April 24, 2019,** at the latest. I will let you know that I have received the report.

References

1. Bresnahan T., P. Reiss, ”Entry and Competition in Concentrated Markets”, *Journal of Political Economy*, Vol. 99, No. 5, pp. 977-1009.

Appendix: Bresnahan and Reiss (1991)
Firm profits are specified according to equation (1). To make the presentation tractable, we use the notation that observed profits equal $\tilde{\pi}(N_m, Y_m, X_m, W_m; \theta)$. The two necessary conditions for a Nash equilibrium in pure strategies are:

a. Firms that enter make positive profits

$$\tilde{\pi}(N_m, Y_m, X_m, W_m; \theta) + \epsilon_m > 0, \quad \text{whenever} \quad N_m > 0$$

b. Further entrants would earn negative profits

$$\tilde{\pi}(N_m + 1, Y_m, X_m, W_m; \theta) + \epsilon_m < 0.$$

where N_m is the number of firms, Y_m, X_m, W_m are local market characteristics in local market m and $\theta = (\lambda, \beta, \alpha, \gamma)$ are the parameters to be estimated. The error term ϵ_m are assumed to be i.i.d across local markets and distributed $N(0, 1)$. Firms are identical and are assumed to have the same unobserved profits.

These assumptions help us to predict a unique number of firms in each market equilibrium. The probability to observe $N_m = n$ firms in market m is

$$\begin{aligned} Pr(N_m = n | Y_m, X_m, W_m; \theta) &= Pr(\pi_n > 0, \pi_{n+1} < 0) \\ &= Pr(\tilde{\pi}_n + \epsilon > 0, \tilde{\pi}_{n+1} + \epsilon < 0) \\ &= Pr(\epsilon > -\tilde{\pi}_n, \epsilon < -\tilde{\pi}_{n+1}) \\ &= Pr(\tilde{\pi}_{n+1} < \epsilon < \tilde{\pi}_n) \\ &= \Phi(\tilde{\pi}_n) - \Phi(\tilde{\pi}_{n+1}) \end{aligned}$$

where $\Phi(\cdot)$ is the normal cumulative distribution function (cdf).[1] The parameters θ can be estimated by maximum likelihood.[2]

---

[1]Since $\epsilon_m \sim N(0, 1)$, we have

$$Pr(\epsilon_m < y) = \Phi(y) = \frac{1}{\sqrt{2\pi}} \int_0^y e^{-\frac{1}{2}y^2}$$

[2]The likelihood is

$$\begin{aligned} L(\theta | \cdot) &= \Pi_{m=1}^{M} Pr(N_m = n | \pi(\cdot; \theta)) \\ &= \Pi_{m=1}^{M} \Phi(\tilde{\pi}(N_m, Y, X, W; \theta)) - \mathbf{1}_{\{n>0\}} \Phi(\tilde{\pi}(N_m + 1, X_m; \theta)). \end{aligned}$$

To be more precise, we consider a sample of markets where the total number of firms N_m varies between 0 and 5. We follow the specification of the profit function in equation (1) and denote observed profits $\tilde{\pi}(N_m = 1, Z_m; \theta) + \varepsilon_m$. In addition, let $Z_m = [Y_m, X_m, W_m]$. We then have the following cases

- if the number of firms is $n = 0$, then

$$\tilde{\pi}(N_m = 1, Z_m; \theta) + \varepsilon_m < 0 \tag{5}$$

- if the number of firms is $n = 1$, then

$$\begin{aligned} &\tilde{\pi}(N_m = 1, Z_m; \theta) + \varepsilon_m > 0 \quad \text{and} \\ &\tilde{\pi}(N_m = 2, Z_m; \theta) + \varepsilon_m < 0. \end{aligned} \tag{6}$$

- if the number of firms is $n = 2$, then

$$\begin{aligned} &\tilde{\pi}(N_m = 2, Z_m; \theta) + \varepsilon_m > 0 \quad \text{and} \\ &\tilde{\pi}(N_m = 3, Z_m; \theta) + \varepsilon_m < 0. \end{aligned} \tag{7}$$

- if the number of firms is $n = 3$, then

$$\begin{aligned} &\tilde{\pi}(N_m = 3, Z_m; \theta) + \varepsilon_m > 0 \quad \text{and} \\ &\tilde{\pi}(N_m = 4, Z_m; \theta) + \varepsilon_m < 0. \end{aligned} \tag{8}$$

- if the number of firms is $n = 4$, then

$$\begin{aligned} &\tilde{\pi}(N_m = 4, Z_m; \theta) + \varepsilon_m > 0 \quad \text{and} \\ &\tilde{\pi}(N_m = 5, Z_m; \theta) + \varepsilon_m < 0. \end{aligned} \tag{9}$$

- if the number of firms is $n = 5$, then

$$\tilde{\pi}(N_m = 5, Z_m; \theta) + \varepsilon_m > 0 \tag{10}$$

The probabilities to observe a given number of firms in market m are as follows:

$$
\begin{aligned}
&Pr(N_m = 0) = 1 - \Phi(\tilde{\pi}(N_m = 1, Z_m; \theta)),\\
&Pr(N_m = 1) = \Phi(\tilde{\pi}(N_m = 1, Z_m; \theta)) - \Phi(\tilde{\pi}(N_m = 2, Z_m; \theta)),\\
&Pr(N_m = 2) = \Phi(\tilde{\pi}(N_m = 2, Z_m; \theta)) - \Phi(\tilde{\pi}(N_m = 3, Z_m; \theta)),\\
&Pr(N_m = 3) = \Phi(\tilde{\pi}(N_m = 3, Z_m; \theta)) - \Phi(\tilde{\pi}(N_m = 4, Z_m; \theta)),\\
&Pr(N_m = 4) = \Phi(\tilde{\pi}(N_m = 4, Z_m; \theta)) - \Phi(\tilde{\pi}(N_m = 5, Z_m; \theta)),\\
&Pr(N_m = 5) = \Phi(\tilde{\pi}(N_m = 5, Z_m; \theta)).
\end{aligned}
$$

Those are in fact the ordered probit probabilities, where the probability to observe a given number of firm is a function of market characteristics and thresholds (α_n, γ_n). Our profit function specification in equation (1) tells us that $\tilde{\pi}_{N_m} = S(Y_m, \lambda)V_{N_m}(X_m, \alpha, \beta) - F_{N_m}(W_m, \gamma)$, which we use in the empirical implementation.