ECN431
Lab session 4
Market structure and local competition

Aud 12
Friday, March 2th 2017

Before the Lab

- Obtain the datasets and the example R notebook from the GitHub repository of the course

Purpose of this lab

Work with analysis of firm entry and retail outlet competition.

Learning goals

- Improve the skill of presenting data for own and other's understanding
- Understand the fundamental endogeneity problems in market structure analysis
- Be able to use visual inspection together with regression to understand data and endogeneity
- Know basic features of the Norwegian retail pharmacy market
- Be able to estimate regressions informing us about market structure
- Be able to explain, present and identify critical assumptions in your analysis

Practical information

The example R notebook (which you either pulled to your fork from the main repository or downloaded from GitHub) is opened in RStudio. This file contains several commands that you will need in order to answer the questions in these exercises, in addition to information on new commands. Remember: If/when you are uncertain about the purpose, functionality and extensions to/options for some of the commands, remember that you can get documentation for the command in RStudio by pressing F1. You should also try to google your particular problem.

Market structure: Pharmacies in Norway

The turnover of medical drugs in Norway was 23.4 billion NOK (approx. 2.5 billion EUR) in 2015. The retailing side of the Norwegian pharmaceutical market largely consists of three private chains who are vertically integrated with wholesale operations, all owned by international corporations, in addition to hospital pharmacies (publicly owned) and a few independent pharmacies. The owner-side of the market is deregulated, in the sense that anybody can own or establish a pharmacy, as long as they comply with the regulation for pharmacy operations and sale of medical drugs.[1] Advertisement for prescription drugs to the general public is forbidden in Norway.[2]

Prescription drugs are subject to price regulation, which means that pharmacies cannot raise prices above given price ceilings. The largest share of medical expenses is covered by the government (around 73% of total drug expenditures in 2015), meaning that consumers do not necessarily face the full price due to insurance.[3] Expenses for medical drugs are not a particularly large part of government expenditures, totalling 6.4% of public health expenditures and 0.5% of GDP in 2015. This is due to a combination of large public expenditures, high GDP, relatively tight price regulation and other public policies.

For the average pharmacy, almost 70% of revenue is due to prescription drug, while less than 10% is due to non-prescription drugs and more than 20% is due to other goods (could be materials for medical treatment, products for hygiene and personal care, as well as some consumables).

[1] This followed a deregulation in 2001.

[2] As in most other countries, with the US as a notable exception.

[3] See lab 3 part 2 for more details on this.

Data

You will work with two datasets for this exercise. The file *stores.csv* contains monthly sales of prescription drugs and other information on pharmacies from March 2004 to December 2011, while the file *markets.csv* contains yearly information on population within demographic groups and geographical size of designated markets in the same time period.[4] The markets are defined based on driving distances in minutes.

Variable list for *stores.csv*:

- date – monthly date in format `YMD`
- phid – pharmacy ID
- chain – chain affiliation ID
- market – geographic market ID
- storesize – physical size of the store in square meters
- sale – total revenue from prescription drugs in NOK
- packages – total number of packages sold
- nobs – number of customers

Variable list for *markets.csv*:

- year – year
- market – geographical market ID
- area – size of the market in square kilometers
- f0_19/m0_19 – females/males in market aged 0 to 20
- f20_66/m20_66 – females/males in market aged 20 to 67
- f67/m67 – females/males in market aged 67 and above

[4]The data is stored in *csv* format (*comma-separated values*). This is a very usual storage format for data, which has the benefit that all statistical and spreadsheet software can read and write such files.

Tasks

Data wrangling

1. Add market population information to the pharmacy store data. You will need to extract year from the date variable and join the data frames using year and market ID as keys.

2. Generate a variable showing the number of pharmacies within each market (for each pharmacy).

3. Generate total population and population density for each market.

Present and understand your data

4. Generate appropriate tables of summary statistics.

5. Make a graph showing the relationship between sales (in monetary terms), number of customers buying from the pharmacy and the number of pharmacies in the market.

6. Make a plot showing the relationship between sales (in monetary terms) and the number of packages sold separately for each pharmacy chain. We can reveal that one of the chain ID's specifies hospital pharmacies.

Estimation

7. Regress sale on the number of pharmacies in the market (i.e., the variable you generated in 2.) How do you interpret the result?

8. Generate a demeaned sales variable, that is, a variable showing the difference between sales and the mean sales (for the pharmacy). In addition, generate demeaned version of the variable showing number of pharmacies in the market.

9. Make a graph showing the relationship between the two variables you just created. Compare this graph to the corresponding graph from Task 5.

10. Regress demeaned sales on the demeaned number of pharmacies in the market.

11. Run a fixed-effects regression of sale on the number of pharmacies in the market. Compare the results with the regression in Task 10

12. Run a regression of sale on the number of pharmacies in the market, but with market, rather than pharmacy fixed-effects. Compare the results.

13. Extra:

    (a) Create a data frame with one observation (row) for each market per year. This frame should consist of one column with market ID, one with the number of pharmacies in the market in a given year and one with the total population in the market (for a later task, you can also include the population within demographic segments here as well).

(b) Make a plot showing the relationship between number of pharmacies and market size (measured by total population). Comment on the nature of the relationship – is it convex or concave? Does an ordinary linear regression line capture the relationship well?

(c) Demean the number of pharmacies and market size within each market and plot the relationship. Compare with the results above and discuss.

(d) Explore whether the number of pharmacies in a market is sensitive to local demographics (compared to an overall increase in population).

(e) Explore whether growth in population or particular demographics predicts entry of pharmacies.

(f) Take a look at Figure 1.1.1 at the following link: `http://www.apotek.no/fakta-og-ressurser/statistikk-for-2016/1--apotek/1-1-apotek-i-norge`. In light of your findings so far, discuss the implications for the development of the structure in this sector over time with an ageing population.