Industrial Organization and Data Science

Instructors: Justin Rao, Affiliate Professor & Senior Researcher, Microsoft

Jacob LaRiviere, Affiliate Professor & Senior Researcher, Microsoft

Emails: justin.rao@microsoft.com

jlariv@microsoft.com

Course Assignments & Reading

Course assignments should be printed (code, output and descriptive answers) and turned in at the start of class unless otherwise noted. Feel free to work in groups but everyone is required to turn in their own work with answers written in your own words. In both calculations and complex ideas, write down each step of logic used in reaching your conclusion. Keep in mind that in most cases a good answer is one precise sentence; quality is heavily favored over quantity. This will be graded on a full credit, half credit and no credit basis. All work must be typed

Discussion questions do not need be written out ahead of time. At the beginning of each class the professors will lead a discussion around these questions. Students will be called on, potentially at random, to add their insight. This part of class will contribute heavily to your course participation grade.

Week 8, due May 25

Discussion questions

In McMillan we learned about Dutch auctions as a price that ticks down from a high starting point and whoever says "I'll take it first" gets it. In Klemperer, he discusses Dutch auctions as sealed bid first price auctions. What are the similarities? Do you think they are basically the same?

Assignment to be turned in. Please turn in your R output and answers to the questions.

In this assignment we're going to do two things. First, we're going to take the output from the week 7 assignment, interpret the coefficients and turn it into an excel spreadsheet to create a pricing and revenue "model". Second, we're going to try to put the data into a panel format and first difference the prices.

1) Now work a bit more with the intertemporal dynamics from last time, as a reminder, here is the function we used then.

$$\ln(Q_{ijt}) = \alpha + \ln(Q_{ijt-1})\beta_1 + \ln(Q_{ijt-2})\beta_2 + [\ln(P_{ijt}) - \ln(P_{ijt-1})]\gamma_1 + [\ln(P_{ijt-1}) - \ln(P_{ijt-2})]\gamma_2 + \epsilon_{ijt}$$

   a. Let's play around with sorting the data. Load the oj data and then use the command: ojsorted = oj[order(oj$store, oj$week),].
      i. The order command is sorted the data frame by the first argument, and then breaking ties with the second argument. Look at the dataframe, in this case you should see the stores are grouped and as you scroll down, you advance by weeks (3 brands for each week).

    ii. Use a similar command to create a dataframe where we order by store, then brand, then weeks. Inspect the data, you should see “chunks” where it is the same store and the same brand and week is advancing as you scroll down (i.e. now brands are chunked together)

  b. With the dataframe from (ii), let’s create some lagged terms. We’ll use this package: install.packages("DataCombine"), library(DataCombine). This package relies on the data being sorted properly. Right now we have arranged our data by store and then by brand and then by week. So if we look within a store, we have all the weeks for a given brand, and then after the end of the time period, we move to the next brand. This tells us that we want to lag “by brand”.

    i. Run: ojsorted <- slide(ojsorted, Var = "logmove", GroupVar = "brand",slideBy = -1). What variable was created?

    ii. Run the same command, by slide by -2, this should give you logmove for 2 weeks back. Inspect the data frame, notice that it creates NAs for the first two weeks of the sample for each brand (e.g. look at row 111)

    iii. Run: ojsorted <- slide(ojsorted, Var = "logmove", GroupVar = "store",slideBy = -1). You will notice this almost works, except look at row 111 in your data, here we have switched from Dominicks in the last week of the sample to MinuteMaid in the first week, so we should get an NA here, but the logmove-1 variable puts a value for lagged Dominicks quantity (from the last week).

    iv. Run the same command on the unsorted data (oj data frame). What goes wrong?

  c. Using this method, create logmove-1 (lagged 1 week), logmove-2,price-1,price-2.

    i. It turns out the DataCombine package gives the variables names that are annoying for R (since -1 is a R specific command), so let’s rename them, using the reshape package (install package and load library)
    ojsorted <- rename(ojsorted, c(`logmove-1`="lag1logmove"))
    Notice I had to put the quotes around logmove-1.
    Do this for all your lagged variables (you can do it one command)

    ii. Run a regression of logmove on your lagged terms (log the price terms) and brand. Print your results. What is the r-squared?

    iii. Now include “today’s price” (log(price)) in the regression. How do the results change? What does the positive coefficient on lagged price indicate (when price was high in the past, what happens to sales today?) What does the positive coefficient on lagged quantities indicate?

    iv. Now include as.factor(week) in the regression to allow for seasonality. Do the results qualitatively change?

    v. Now run the model with the formula as given above? What do you learn about how price changes impact sales?

2) According the Klemperer article, what are the main problems that go wrong with auction design? (1 paragraph answer)