Empirical problem set, Industrial Organization, Richard Friberg 2022

The purpose of this exercise is to let you get a feel for how one can estimate demand for differentiated products. I want results in a nice-looking format as if you were to present them in a paper. On each point I want the results as well as some discussion. You can discuss together but I want you to run your own regressions and provide your own discussion.

The data are in a file "wine_ps2.dta" in stata format (also available as an excel file), which is available on the course home page. It is a subset of the data used for the paper Friberg, Paterson and Richardson (2011), "Why is there a home bias: A case study of wine". See that paper for a description of the source etc (it's up on my webpage, under "older papers" http://dokument.richardfriberg.se/fribergetal_final.pdf). It covers sales of wine in 20 New Hampshire liqor stores. The variables are:

Date: selfexplanatory
Storenum: each of the 20 stores in the sample has a unique number
Idcode: identifies the wine
numbot: number of bottles sold of that itcode sold in that store on that week
proof: alcohol content
sub_class the classification from the NH liqor commission
grape: which grape the wine is made of
bot_oth: sales of wine other than red in the respective store and week
national: country of origin
p: price per bottle (nominal)
variet: 0 if blended (or European that is not marketed according to grape): 1 if made of specified grape varietals (such as merlot)
doc: 1 if made according to European region of origin rules
cabsv, merlot, shiraz, zinf: grape dummies
fx: exchange rate vis-a vis source country

1) Do descriptive statistics as you would in a paper: number of products, volumes etc. Comment on these.
2) Choose some wines and estimate demand as a function of own price and price of competing products and possibly other factors (dependent variable is thus quantity sold of a particular wine). Discuss potential problems that you run into.
3) Now estimate the demand system using a discrete choice setting as in Berry (1994) using observable product characteristics. Assume that individual taste parameters are iid with an extreme value distribution such that the market share for each product is given by the "logit formula" as in eq (13) in Berry (1994). Estimate the equivalent of eq (14) in Berry (94) using OLS. Discuss the coefficients.
4) Assume that marginal costs are constant and that the price of each wine is optimised individually so that cross-price effects are not taken into account when optimizing. Using the first order condition (as in eq 31 in Berry) back out the implied marginal cost of each product. What is the implied markup - at the 5th percentile, 25th, median, 75th and 95^{th}?
5) Do the same as in 3) and 4) but use product fixed effects instead of observable product characteristics. Comment on how results differ.
6) Do the same as in 3) and 4) but also use BLP-style instruments for price. Comment on how results differ.

7) Repeat 3-4), but instead using a "nested logit" specification. (see eqs 28 and 33) in Berry. Use countries of origin to form nests. What are implied mark-ups now?