**Computational Assignment**

**Empirical Industrial Organization and Market Design**

**Due by midnight of December 18, 2019**

This assignment will require you to estimate a random coefficient logit model (i.e., BLP) for the app market that you analyzed in your problem sets 1 and 2. Please use the dataset enclosed in the Computational Assignment folder as it contains downloads performed at different points in time, so that you can exploit the time variability for identification. In the same folder you also have a second dataset containing demographic characteristics of the different markets.

Your task consists of performing the estimation of a baseline BLP model and to evaluate a few alternative specifications. To simplify your work, instead of starting from scratch, you are welcome to begin from my BLP codes enclosed in the folder. Note that these codes (there is both a Matlab and a Python version) perform the analysis for a different dataset/market (that of Medicare Part D), so you will need to operate the needed changes.[[1]](#footnote-1)

**Premise**: some notation used in my code is as follows. The master file is called “r1\_medicare.m” and, within this file, the crucial objects are:

- X1: the observed characteristics

- X2: the subset of X1 that gets random coefficients by interacting with the demographic and noise distributions

- IV: the set of instruments (check both line 151 and 253, in the latter we extend the set instruments to have some IV that move with the demographics)

- the dataset with the demographics contains several variables on features like age and income of the various markets at various points in time. *Note that for the app market the dataset with the demographics is isomorphic to this Medicare demographics dataset: the variables are exactly the same (but, clearly, the number of periods and the number of markets differs relative to the Medicare case).*

- initial guesses: “mvalold” and “theta2”

- [theta2,fval,exitflag,output] = fminsearch('gmmobjg',theta2, options): this is the function launching the minimization routine of the GMM objective function 'gmmobjg'. If you open 'gmmobjg', you can see how all other subfunctions are invoked.

Once you run the code, among the interesting output produced in your workspace:

- theta1: the estimates of the linear parameters of the characteristics in X1, in the order in which they appear in X1 and rescaled by 100

- theta2: the estimates of the linear parameters of the characteristics in X1, in the order in which they appear in X1 and rescaled by 100

- se: the standard errors of the theta parameters

- fval: the value of the objective function at the optimum

- hist(alfa\_i, 25): the histogram describing the distribution of the alphas

Now that you have familiarized with the code, let’s start to play!

**Detailed list of tasks:**

1) Using the same instrument as in the simple Berry logit of your problem set 2, estimate the BLP model. In particular, use as characteristics of the products: price, average score, in app purchases. Allow for the random coefficients only on price, but not on the other product characteristics, and use income as the only demographic to be interacted with price.

2) I want you to make small edits to the existing code to see how estimates respond to them. That’s similar to what you did for the logit in your problem set 2. Thus, create this table:

- In the first columns, report the estimates of the theta 1 and 2 parameters, their standard errors, the value of the objective function at the optimum. More in detail, I am asking you just to run the code as it is after you implemented point 1 above.

- In the second column, obtained after adding one more characteristic to your specification: dummy variables for the genre (more precisely: use the variable “nest” in the dataset so that you have less genres).

- In the third column report results like those of column 1 but change the set of instruments. Do this by using as IV for price an Hausman style instrument built as the average price of the same app in all other countries.

- In the fourth column, add one more characteristic into X2. Try adding “averagescore.” *Note: now you need to make additional guesses for the interaction terms.*

1. Note that there are many different codes out there estimating BLP. My codes are based on: <http://faculty.wcas.northwestern.edu/~ane686/supplements/rc_dc_code.htm>

   <http://www.rasmusen.org/zg604/lectures/blp/frontpage.htm>

   <http://emlab.berkeley.edu/users/bhhall/e220c/rc_dc_code.htm>

   <http://web.mit.edu/knittel/www/KM_website.html>

   <http://faculty.chicagobooth.edu/che-lin.su/research/code.html>

   There are also several more codes for Stata or Python: just search online (I am not familiar with them, so cannot recommend one). For this problem set, you are welcome to start from my code or from one of the others. If you start form my code, I suggest that you begin by simply running it on the Medicare data and look at all the objects that are created in order to familiarize with its working. [↑](#footnote-ref-1)