Assignment – Matching Models

Instructions. The assignment contains two problems. In the first problem, you will estimate the Dupuy and Galichon (2014) model, as in the tutorial, on worker-firm data. For this, you will need to have the *affinitymatrix* R package installed. In the second problem, you will look at a simple application of the Choo and Siow (2006) model.

Please email your solutions to alexander.wintzeus@kuleuven.be by **Friday, 13 June 12:00 CEST** at the latest. Your solutions should consist of a single R file, possibly accompanied by a small text document with written answers to questions on interpretation. The latter can also be included as comments within the R file.

Problem 1. Load in the data set `P1_Data.csv`.[1] The data consist of a sample of matched workers and firms (or jobs). Workers differ in their cognitive (`cog_skill`) and manual (`man_skill`) skills. Jobs differ in the intensity with which they require these two types of skill (`cog_job` and `man_job`).

1. Infer the number of matches N from the data. Create two matrices i and j containing the observations for workers and jobs, respectively.
2. Estimate the affinity matrix using the `estimate.affinity.matrix()` function and print out the results using `show.affinity.matrix()`.
   - What can we learn about the matching patterns in our data from these estimates?
   - Provide a correct interpretation of the estimated coefficient related to the worker's manual skills and the job's cognitive demand.
3. Print out the results of the rank test using the `show.test()` function.
   - What does this test suggest about the rank of the affinity matrix?
   - How does the rank test relate to the singular value decomposition of the affinity matrix (*saliency analysis*)?

Problem 2. Load in the data sets `P2_Data_Couples.csv` and `P2_Data_Singles`.[2] The data contain information on the educational attainment of a sample of couples and a sample of single men and women, respectively. Education is discrete, and ranges from 1 (high school dropouts) to 4 (college education).

1. Recall the Choo and Siow (2006) matching function:
$$\Pi_{ij} = \frac{\mu_{ij}}{\sqrt{\mu_{i0}\mu_{0j}}}.$$
   Use this matching function to estimate the surplus Π_{ij} nonparametrically. You can do this efficiently by looping over all possible combinations (i, j).
2. What can we learn about the matching patterns in our data from these estimates? How does this relate to random matching?
3. Does the model allow us to infer anything about male (α_{ij}) and female (γ_{ij}) preferences?

[1] The data are taken from Lindenlaub (2017).
[2] The data are generated, based on Eika *et al.* (2019).

References

Choo, E. and Siow, A. (2006). Who marries whom and why. *Journal of Political Economy*, **114** (1), 175–201.

Dupuy, A. and Galichon, A. (2014). Personality traits and the marriage market. *Journal of Political Economy*, **122** (6), 1271–1319.

Eika, L., Mogstad, M. and Zafar, B. (2019). Educational assortative mating and household income inequality. *Journal of Political Economy*, **127** (6), 2795–2835.

Lindenlaub, I. (2017). Sorting multidimensional types: Theory and application. *Review of Economic Studies*, **84** (2), 718–789.